Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

_____ _____

Amanda Pierce Date

Determining the effects of animal migration and range expansion on population genetics

By

Amanda Pierce
Doctor of Philosophy

Graduate Division of Biology and Biomedical Sciences
Population Biology, Ecology, and Evolution

Jacobus de Roode, Ph.D.
Advisor

Nicole Gerardo, Ph.D.
Committee Member

Timothy Read, Ph.D.
Committee Member

Leslie Real, Ph.D.
Committee Member

Michael Zwick, Ph.D.
Committee Member

Accepted:

Lisa A. Tedesco, Ph.D.
Dean of the James T. Laney School of Graduate Studies

Date

Determining the effects of animal migration and range expansion on population genetics

By

Amanda Pierce
B.S. University of Kansas, 2010

Advisor: Jacobus C. de Roode

An abstract of
A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies at Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
Graduate Division of Biological and Biomedical Sciences
Population Biology, Ecology, and Evolution
2015

Abstract

Determining the effects of animal migration and range expansion on population genetics

By Amanda Pierce

Animal movement does not only affect the individual, but can also have profound implications for population dynamics and species distributions. In this thesis, I use the monarch butterfly, *Danaus plexippus*, as a model system to understand the connectivity and genetic differentiation among populations with different migratory strategies, as well as the effects of range expansion on population genetics and structure. Monarchs are known for their fall migration from eastern North America to their overwintering sites in Mexico, but also occur west of the Rocky Mountains, from which they migrate to the California coast. Using microsatellite markers, I found that despite differences in migration destination and the Rocky Mountains to serve as a potential barrier, eastern and western North American monarchs are genetically indistinguishable. This indicates that monarchs are able to maintain divergent migratory pathways despite high genetic similarity. I expanded this study to include additional sampling sites located south of the Mexican overwintering sites. It has long been believed that monarchs have a two-way migration in which all monarchs return north after the overwintering period. However, I found that monarchs in Costa Rica and Belize are not genetically differentiated from their northern counterparts. A hypothesis is that rather than a strict two-way migration, some percentage of monarchs instead radiate outwards from the overwintering sites. Moreover, the monarchs in Costa Rica and Belize are non-migratory, which again demonstrates how different migratory strategies are maintained despite high similarity among neutral genetic sites. In addition to the monarchs mentioned, there are also non-migratory monarchs located around the world wherever temperature and larval food plant abundance allow. Despite monarchs colonizing these locations fairly recently, my research found high levels of genetic structure and differentiation across areas and continents that are separated by seas and oceans. My work also suggests that despite the high proclivity for dispersal, genetic drift still plays a major role in shaping allele frequencies in these newly colonized areas, through multiple and serial founder effects. Future work coupling population genetic theory with next generation technologies will lead to additional breakthroughs in this field.

Determining the effects of animal migration and range expansion on population genetics

By

Amanda Pierce
B.S. University of Kansas, 2010

Advisor: Jacobus C. de Roode

A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies at Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
Graduate Division of Biological and Biomedical Sciences
Population Biology, Ecology, and Evolution
2015

Acknowledgements

I would first like to thank my advisor, Jaap de Roode, for being an amazing mentor. His guidance and support throughout my time in graduate school have been invaluable and I would not be the scientist I am today without his help.

I would also like to thank my committee for their guidance both scientifically and professionally. I thank Nicole Gerardo for not only being willing to read drafts of manuscripts and applications, but also for taking a master's intern into her lab who would eventually become my husband. Aside from their roles on my committee, I thank Mike Zwick and Tim Read for their leadership as PBEE program director and DGS, respectively, during the first years of my graduate career. Finally, I thank Les Real for always asking the big questions and for introducing me to classic film. Outside of my committee, I would like to thank Marcus Kronforst (U. Chicago) and Sonia Altizer (U. of Georgia) for their advice and input. I have greatly benefited from their scientific expertise. Additionally, I thank Jennifer Gleason (U. of Kansas) for mentoring me while I was an undergraduate and giving me my first introduction to research.

Thank you to my de Roode "lab family" for creating a fun, supportive work environment. I could not ask for better colleagues or people to hang out with every day. I would also like to thank the members of my PBEE cohort aka BAGS- Brooke Bozick, Gayatri Sekar, and Sarah Guagliardo. I feel incredibly fortunate to be a member of BAGS and to have made multiple life-long friends.

Finally, I would like to thank my husband, Tom de Man, as well as the rest of my family. The last five years have been full of ups and downs and I could not have made it through without their support and love.

This work was supported by Emory University, an NSF grants DEB-1019746 and DEB-1257160 to Jacobus de Roode, and an NIH Training Grant in Population Biology of Infectious Diseases (5T32AI055404-10).

Table of Contents

Chapter 3: Unraveling the mysteries of monarch migration and global dispersal through molecular genetic techniques

Chapter 4: Serial founder effects and genetic differentiation during worldwide range expansion of monarch butterflies

Chapter 1

Introduction

This dissertation focuses on the effects of animal movement on population structure and genetics. Animal movement can have immediate and long term effects by impacting reproduction and survival. Movement does not only affect the individual, but can also have profound implications for population dynamics and species distributions (Birand *et al.* 2012; Johst & Brandl 1997; Slatkin 1987). Additionally, animal movement affects local community ecologies through changing trophic interactions, bringing in or removing vectors of disease, or by moving genetic material (Holdo *et al.* 2011). Two forms of movement that I focus on, and which differ in directionality and function, are dispersal and migration.

Dispersal is the permanent movement of an individual away from a source population, and affects the range in which genetic mixing occurs. Consequently, dispersal has effects on inbreeding and adaptation based on the associated levels of gene flow (Ebert *et al.* 2002; Johst & Brandl 1997; Pusey & Wolf 1996). Animals can disperse actively as a means to find mates or to escape deteriorating conditions when resources become scarce, or passively through weather events such as strong winds or hurricanes. In addition to dispersal events, each year, millions of animals undertake long-distance migrations to escape changing habitats, competition, and predation (Alerstam 2006; Alerstam *et al.* 2003; Altizer *et al.* 2011; Dingle 1972, 1996; Fricke *et al.* 2011; McKinnon *et al.* 2010). Migratory behavior, which involves synchronized and directional movement, requires navigational abilities (Liedvogel *et al.* 2011), and can have profound effects on species interactions and habitat ecology (Birand *et al.* 2012; Johst & Brandl 1997; Slatkin 1987). Some of the most impressive animal migrations involve the directed movement of millions of individuals across distances that span whole continents or hemispheres (Alerstam *et al.* 2003).

Dispersal and migration can result in speciation by creating geographic, behavioral, or temporal reproductive barriers. Conversely, movement can impede speciation by increasing gene flow amongst populations, which consequently hampers adaptation (Lenormand 2002). Gene flow can also have important consequences for the spread, persistence and evolution of infectious diseases (Cronin 2009; Ostfeld *et al.* 2005; Riley 2007; Thrall & Antonovics 1995; Thrall & Burdon 1997). Host movement can dramatically affect host-parasite interactions by allowing animals to escape from parasitized locations (Hassell *et al.* 1991). In addition, it can also affect host-parasite interactions by resulting in the spread of host resistance genes and parasites across landscapes (Best *et al.* 2011; Carlsson-Graner 2006; Carlsson-Graner & Thrall 2002). Due to the large effects that animal movement has on a multitude of processes, uncovering its patterns and mechanisms is critical for our understanding of population genetics and evolution.

1.1 Animal dispersal and its effects on population genetics

Animal dispersal can have profound effects on population structure and genetics through founder effects and alterations in gene flow. These effects can have lasting impacts on the evolution of a species, as founder effects and other forms of genetic drift can leave a persistent signal over time even in the face of strong natural selection (Kolbe *et al.* 2012). Similarly, the degree of gene flow can either assist in speciation through a reduction in interbreeding or impede local adaptation through an influx of non-adapted alleles. Due to these effects, there is considerable interest in understanding how animal dispersal further affects genetic and evolutionary processes. For my thesis, I focus on the specific effects of range expansion, as well as understanding the role of current gene flow, on population genetic makeup.

Typically, range expansions are characterized by a decrease in genetic diversity and an increase in among-population differentiation with increasing geographic distance from the source population (Eckert *et al.* 2008; Peter & Slatkin 2013; Schulte *et al.* 2013). Theoretical

and empirical studies have shown that in the case of a range expansion, the highest levels of genetic diversity tend to be located in the oldest portions of the range (Francois *et al.* 2008; Slatkin & Excoffier 2012; Taberlet *et al.* 1998), and that genetic diversity decreases with increasing distance from the source range due to serial founder events (Austerlitz *et al.* 1997). One way in which this has been detected is through measures of heterozygosity. In *Homo sapiens*, migration out of Africa has resulted in a number of serial founder effects, resulting in a decrease in heterozygosity with increasing distance from the African origin in Ethiopia (Deshpande *et al.* 2009; Li *et al.* 2008; Prugnolle *et al.* 2005; Ramachandran *et al.* 2005). Theory dictates that genetic drift should play a large role in the genetic structuring of newly formed populations along a range expansion front due to low population density and the potential for high growth rate (Edmonds *et al.* 2004; Klopfstein *et al.* 2006).

One genetic phenomenon that results from this is gene surfing (Figure 1), in which neutral, or even deleterious alleles, reach higher than expected frequencies along the front of an expansion wave and can result in genetic differentiation among populations (Biek *et al.* 2007; Edmonds *et al.* 2004; Flaxman 2013; Hallatschek & Nelson 2008; Klopfstein *et al.* 2006; Slatkin & Excoffier 2012). An experimental study using fluorescently labeled bacteria demonstrated that genetic drift can strongly change allele frequencies during range expansion, letting some go to fixation and others to extinction (Hallatschek *et al.* 2007). In fact, genetic drift can be such a powerful force during a range expansion that another study on budding yeast found that drift was able to override selection for mutualistic cooperation (Muller *et al.* 2014). As a result of gene surfing, clinal patterns in allelic frequencies are typically observed, in which the frequency of certain alleles increase and become more dominant along the wave front (Slatkin & Excoffier 2012), resulting in an increase in genetic differentiation among spatial groups (Excoffier & Ray 2008). The changing of allelic frequencies in combination with an increase in differentiation can even give the appearance of natural selection (Excoffier *et al.* 2009).

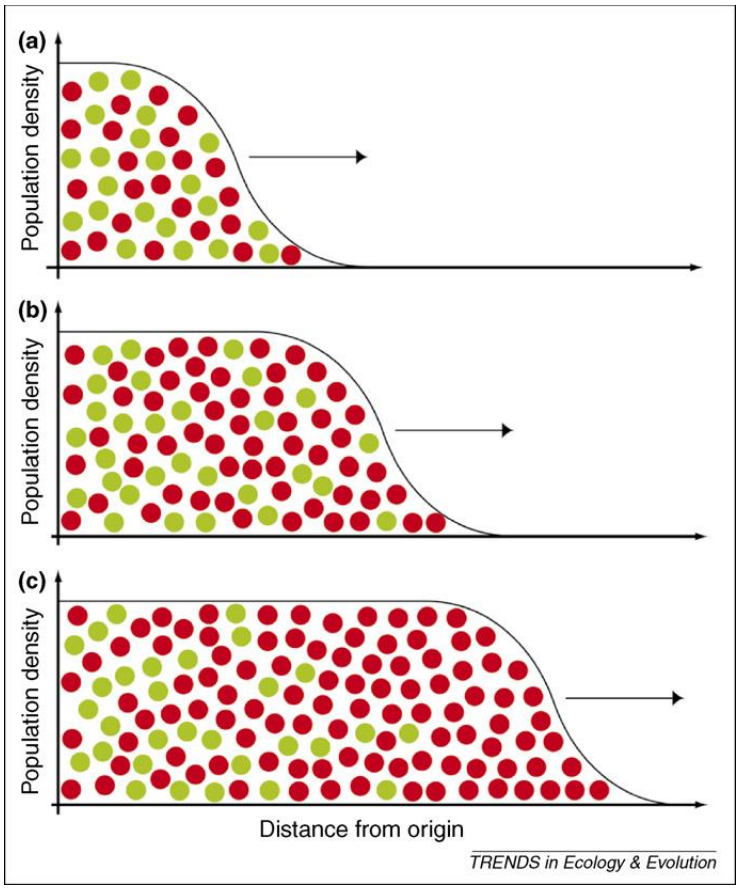

Figure 1.1. Genetic drift occurring at the wave front of an expanding population, leading to changes in allele frequencies and surfing. (a) Initial conditions show an equal proportion of two alleles (red and green). **(b)** The red allele found by chance at the tip of the wave front in (a) increases in frequency. **(c)** The red allele has become fixed by drift at the wave front. *Reprinted from Trends in Ecology and Evolution, 23, Excoffier & Ray, Surfing during population expansions promotes genetic revolutions and structuration, 347-351, 2008, with permission from Elsevier*

Gene surfing has been observed after range expansions in multiple species, including the bank vole, *Myodes glareolus*, in Ireland (White *et al.* 2013) and the tortoise *Testudo graeca* in Northern Africa and south-eastern Spain (Gracia *et al.* 2013), both of which have relatively low rates of dispersal. While these studies have demonstrated the powerful effects of genetic drift in shaping population genetics due to range expansions in both natural and invasive species, few studies have been carried out on organisms with great dispersal ability. In these situations, geographic distance among populations may not be as important, and consequently, gene flow may override genetic drift. In simulations investigating the effect of long range dispersal, it was

indeed found that while serial founder events tend to result in a loss of genetic diversity, diversity can be preserved through high rates of long-distance dispersal (Fayard *et al.* 2009). These long-distance dispersal events can also result in a reduction of genetic differentiation between populations, thus lessening many of the effects of gene surfing (Bialozyt *et al.* 2006; Ray & Excoffier 2010). In fact, the disappearance or reduction of an isolation-by-distance pattern has been shown in a handful of instances involving long-range dispersal, as is the case with invasive European starlings in South Africa (Berthouly-Salazar *et al.* 2013) and peat mosses in the Stockholm archipelago in Sweden (Szovenyi *et al.* 2012).

Similar to species with long distance dispersal, a highly mobile migrating species may also experience sufficient gene flow between populations to override the loss of genetic diversity. Each year, millions of animals undertake long-distance migrations to escape changing habitats, competition, and predation (Alerstam 2006; Alerstam *et al.* 2003; Altizer *et al.* 2011; Dingle 1972, 1996; Fricke *et al.* 2011; McKinnon *et al.* 2010). During migration, some individuals may be either blown off course, or drop out along the way, resulting in the formation of small colonies and subsequent range expansion. In species with limited dispersal, this may result in isolation by distance; however, with the high mobility of migrating species, higher levels of gene flow may exist among populations along the expansion wave front.

1.2 Animal migration genetics and its effects on population structure

Migration can affect levels of gene flow within and genetic differentiation among populations, especially when species occupy multiple breeding grounds and migration destinations (Haig *et al.* 1997). Divergent migratory pathways and destinations could lower opportunities for genetic mixing, and hence result in genetic divergence. For example, the Old World noctule bats, *Nyctalus noctula*, which migrate between hibernating and summer nursing sites, are genetically differentiated with respect to overwintering sites and migration flyways (Petit & Mayer 2000). Similarly, beluga whales, *Delphinapterus leucas*, migrate between wintering sites in arctic pack

ice and summering grounds in arctic and subarctic offshore waters, and there are considerable levels of differentiation between belugas using different summering grounds (O'Corry-Crowe *et al.* 1997).

In contrast, the use of common migratory flyways, breeding grounds, or overwintering areas can lead to high levels of genetic mixing, even when populations experience different selection pressures or population sub-structuring at other points in their migratory cycle. For example, red-billed quelea birds, *Quelea quelea*, in southern Africa undergo long-distance migrations in response to seasonal patterns of rainfall and grass seed production. Although different groups of birds move in north-westerly or south-easterly directions (Dallimer & Jones 2002), genetic analysis shows high levels of mixing between these groups, which probably occurs when birds re-colonize the same areas in the following season (Dallimer *et al.* 2003).

As with other species, the use of widely dispersed breeding grounds, distinct wintering sites, and different migratory flyways could cause local genetic differences in migratory monarch butterfly populations, the focus of this thesis. The earliest published population genetic study on monarch butterflies examined this issue using allozymes. In a seminal paper, Eanes and Koehn (1978) examined the population structure of eastern North American monarchs by collecting 30 geographic samples throughout the monarchs' summer breeding grounds and along their fall migration routes in the eastern United States. Using six allozyme loci, Eanes and Koehn found differentiation between monarch groups during the summer breeding season. They hypothesized that this differentiation may result from genetic drift. Because monarchs are more regionally contained during the summer than during the migration, random drift could result in differences in allele frequencies between sub-populations. It is also possible that differential selection could cause allele frequencies to vary among sampling locations. Such differential selection may act on allozyme markers, some of which have been linked with flight metabolism (Hughes & Zalucki 1993; Solensky & Oberhauser 2009; Zalucki *et al.* 1993). Eanes & Koehn also found that the annual migration erased the genetic differentiation detected across summer

breeding sites, by mixing monarchs from different breeding regions. Thus, as with red-billed quelea birds (Dallimer *et al.* 2003), monarchs originating from a range of breeding sites appear to mix randomly during the migration season.

Migration has also been deemed a "magic trait", or a trait that drives speciation, as differences in migration destination or timing could result in reproductive isolation and eventually species divergence (Irwin & Irwin 2005; Servedio *et al.* 2011). In addition to understanding its effects on population genetics and structure, it is also crucial to uncover the underlying genetic architecture of migratory ability. Although migration plays a large role in evolution, the mechanisms by which migrating animals navigate are not fully understood (Alerstam 2006). It appears circadian rhythm, (Froy *et al.* 2003), geomagnetic fields (Benhamou *et al.* 2011), polarized light (Reppert *et al.* 2004), and use of a celestial compass are some of the potential mechanisms involved in various animal migrations (Åkesson *et al.* 2001; Merlin *et al.* 2009; Reppert *et al.* 2010). However, the underlying genetic mechanisms of navigation remain a puzzle. Studies using neutral markers to investigate intraspecific divergent migratory phenotypes have found that differences in traits either do not, or only weakly, correlate with large scale genetic differentiation (Bensch *et al.* 2002; Bensch *et al.* 1999; Buerkle 1999; Wolf *et al.* 2010). It is becoming clear that differences in migratory phenotypes may instead be caused by differences in relatively few genomic regions, resulting in the alteration of gene networks and the expression levels of many genes (Liedvogel *et al.* 2011).

1.3 Using monarch butterflies as a model system to test the effects of dispersal and migration on population genetics

To investigate the genetic effects of migration and long-distance dispersal, I use the monarch butterfly as a model system. The monarch butterfly, *Danaus plexippus*, is an ideal study system for these objectives. Monarchs are found in populations worldwide with highly variable between-population geographic distances and dispersal barriers, and occur wherever

temperature and larval food plant distribution allow, as the caterpillars are specialist herbivores feeding only on milkweed plants in the family *Asclepiadaceae*. Monarchs have only recently formed stable colonies in the majority of these locations, providing a unique look at the effects of recent dispersal events on gene flow and genetic diversity. Monarch butterflies also have a naturally occurring protozoan parasite, *Ophryocystis elektroscirrha*, which allows for the investigation of the effects of host movement on parasite prevalence. Finally, monarchs are well known for their autumnal migration from eastern North America to Mexico; hundreds of millions of monarchs escape freezing temperatures and dying larval food plants in eastern North America, travel upwards of 4000 km to overwinter in the Oyamel fir forests of the Mexican Transverse Neovolcanic Range, then re-migrate to eastern North America in the spring (Urquhart & Urquhart 1978, 1979). Each migrating butterfly makes the journey only once, and 3-4 generations separate each migration. Therefore, the navigation and homing mechanisms involved must be innate. This makes the monarch butterfly an exceptional organism to study the effects of migratory strategy on genetics.

Figure 1.2. Overwintering sites of the eastern monarchs in Mexico (left) and the western monarchs in California (right).

Aside from this eastern group, monarchs also occur west of the Rocky Mountains, from which they migrate to the California coast (Tuskes & Brower 1978; Figure 2). Conventional wisdom suggested that these monarchs were geographically separated by the Rocky Mountains, thus preventing gene flow from occurring, but studies with adequate genetic markers and samples have been lacking (Figure 3).

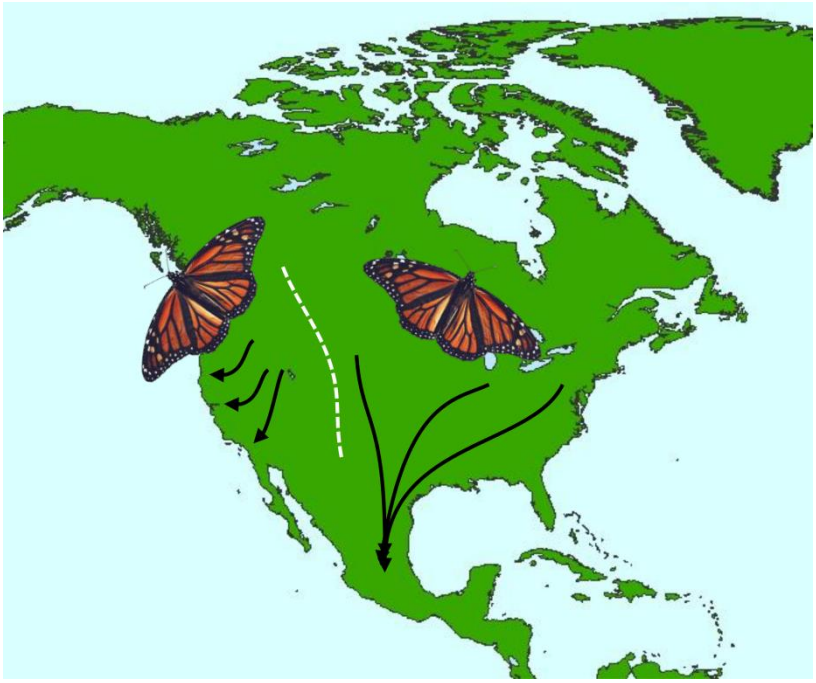

Figure 1.3. Map showing the location and migratory patterns of eastern and western North American monarchs. Eastern North American monarchs migrate between eastern North America and Mexican overwintering sites. Western North American monarchs migrate to the California Pacific Coast and are believed to be geographically and genetically separated from the eastern monarchs by the Rocky Mountains (indicated in the white dashed line).

Importantly, monarchs also occupy locations ranging from the New World tropics to more recently-colonized Pacific islands (Ackery & Vane-Wright 1984) to Europe (Figure 4), and most of these populations are non-migratory (Altizer *et al.* 2000; James 1993). A population is deemed non-migratory if its habitat range does not significantly differ throughout the year. The wide distribution of monarchs raises important questions regarding the genetic differences between and interconnectedness among existing populations. For example, is genetic separation a prerequisite for monarch populations to maintain different migration strategies and

destinations; and have non-migratory populations repeatedly arisen from migratory ancestors? The occurrence of monarchs around the world raises the additional question of where all these monarchs came from. Which ancestral populations served as sources for more recent colonization events, and in what patterns did monarchs spread across the Atlantic and Pacific Oceans from the New World (Vane-Wright 1993; Zalucki & Clarke 2004)?

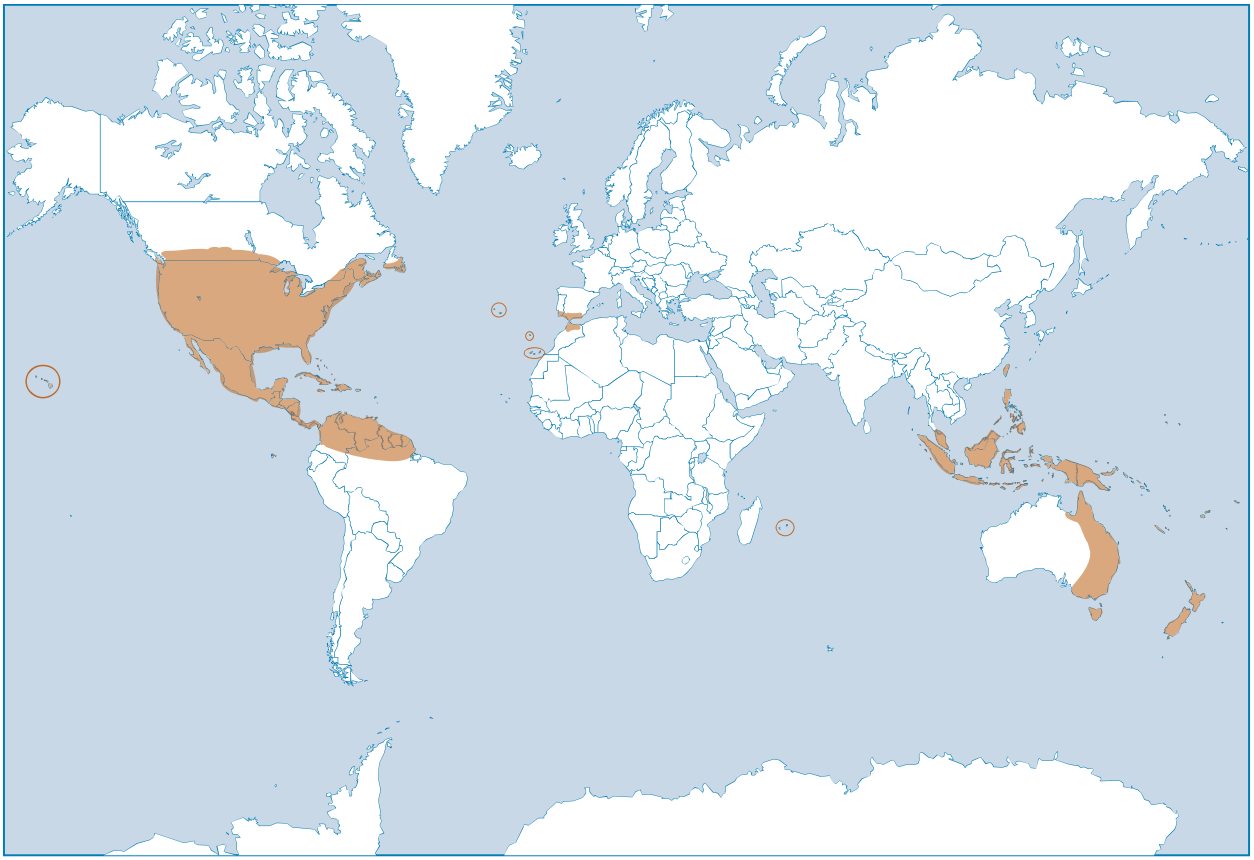

Figure 1.4. Worldwide distribution of monarchs. Orange shading and circles indicate known monarch range, following Ackery and Vane-Wright (1984) and updated following Neves et al. (2001) and personal communications.

Gene flow can also have important consequences for the spread, persistence and evolution of infectious diseases (Cronin 2009; Ostfeld *et al.* 2005; Riley 2007; Thrall & Antonovics 1995; Thrall & Burdon 1997). Host movement can allow animals to escape from parasitized locations (Hassell *et al.* 1991) as well as result in the spread of host resistance genes

and parasites across landscapes (Best *et al.* 2011; Carlsson-Graner 2006; Carlsson-Graner & Thrall 2002).

Monarchs are affected by a naturally occurring protozoan parasite, *O. elektroscirrha*, which is transmitted when infected adults scatter parasite spores onto their eggs and surrounding milkweed leaves. Larvae ingest the spores, parasites replicate within larval and pupal tissues, and butterflies emerge with dormant spores on the outsides of their bodies (Leong *et al.* 1997a; McLaughlin & Myers 1970). Parasites can be transferred vertically, from infected adults to their progeny, and horizontally, when unrelated larvae ingest spores (Altizer 2004; de Roode *et al.* 2009). Previous work has shown that parasite infection reduces host fitness by causing reduced pre-adult survival, smaller adult body mass and wing size, and shorter adult fecundity and lifespan (Altizer & Oberhauser 1999; de Roode *et al.* 2009; de Roode *et al.* 2007).

All monarch populations examined to date are parasitized by *O. elektroscirrha,* and prevalence is highly variable among regions (Altizer *et al.* 2011; Altizer *et al.* 2000). Monarchs in southern Florida and Hawaii that breed year-round and are non-migratory populations have the highest average parasite prevalence rate with up to 85% heavily infected. Roughly 30% of western North American monarchs are heavily infected whereas less than 8% of eastern North American monarchs are heavily infected (Altizer *et al.* 2011; Altizer *et al.* 2000). These patterns suggest a negative relationship between migratory distance and parasite prevalence, which could be due to losses of infected individuals along migratory routes and escape from infected habitats (Bartel *et al.* 2011).

Given the popularity of monarchs and the long history of scientific study focused on them, it is perhaps surprising that many questions regarding their migration and evolutionary history remain unanswered. The recent development and use of microsatellite markers (Lyons *et al.* 2012; Pierce *et al.* 2015; Pierce *et al.* 2014a; Pierce *et al.* 2014b), the publication of the monarch's genome sequence (Zhan *et al.* 2011), and the genomic work that preceded it (Zhu *et al.* 2008) now offer the potential to explore many aspects of monarch biology from a genetic

perspective. The monarch genome is 273 Mb, contains 29-30 chromosomes, and has a GC content of 31.6% (Zhan *et al.* 2011). The genome and EST library have already been used to identify genes potentially important in the navigational mechanisms involved in monarch migration (Zhu *et al.* 2009; Zhu *et al.* 2008). Further use of these tools will allow for additional investigation of animal migration and other aspects of monarch ecology and evolution. In this thesis, I address some of these topics and discuss answers offered by recent analyses. I also describe how the use of microsatellite markers provides insight into the genetic connectedness between monarch populations and worldwide monarch dispersal, and how these studies may change our thinking on monarch migration.

On a broader scale, I examine questions such as: Is wide-scale genomic differentiation necessary in order to maintain divergent migratory strategies? What effect does the migration of one population have on other non-migratory populations? Does the proclivity for long distance movement override the effects of genetic drift after a global range expansion? Finally, do host movement and population structure drive differences in disease patterns?

1.4 Summary of Chapters 2, 3, 4, 5, and 6

Monarch butterfly migration and range expansion provide an excellent opportunity to investigate the effects of both of these types of movement on underlying population genetics. In Chapter 2, I first examine the effect of divergent migratory pathways and overwintering sites on population genetics and structure. This work was published in *Molecular Ecology* in a study entitled “Lack of genetic differentiation between monarch butterflies with divergent migration destinations” (Lyons, Pierce *et al.* 2012). It has long been assumed that eastern and western North American monarchs, which are separated by the Rocky Mountains and travel to different overwintering sites, must be genetically distinct; however, my work has dispelled this notion and shown that based on neutral genetic markers, the two groups are genetically indistinguishable.

This demonstrates that alternate migratory strategies are able to be maintained despite extensive gene flow.

Chapter 3 expands on this study by further adding sample sites in the Americas: Bermuda, Mexico, Puerto Rico, Belize, Costa Rica, and Ecuador. This work has been published as a chapter entitled "Unraveling the mysteries of monarch migration and global dispersal through molecular genetic techniques" in the book Monarchs in a Changing World: Biology and Conservation of an Iconic Butterfly (Pierce *et al.* 2015). Monarchs from Belize and Costa Rica, which are populations south of the Mexican overwintering site, show minimal genetic differentiation from the eastern migratory population and even cluster with the migratory population as one panmictic group. This work suggests that rather than a strict two-way migration, monarchs may instead radiate outwards from the Mexican overwintering sites in search of larval food plants.

In Chapter 4, I added further sample sites from around the globe in order to determine the route by which monarchs expanded their range as well as the effect of this range expansion on population genetics. The results of this study were published in the *Proceedings of the Royal Society B* in a publication entitled "Serial founder effects and genetic differentiation during worldwide range expansion of monarch butterflies" (Pierce *et al.* 2014). Across the Pacific, monarchs appear to have colonized in a serial stepwise fashion with North America serving as the source population; whereas across the Atlantic, I observed signatures of multiple colonization events. Additionally, I found that genetic drift seems to have played the major role in shaping current allele frequencies and population genetics in these colonized populations.

In Chapter 5, I investigated the role of host population structure and genetics on infection status among monarchs distributed throughout the Hawaiian Islands. This work was published in *Plos ONE* and was entitled "Extreme heterogeneity in parasitism despite low population genetic structure among monarch butterflies inhabiting the Hawaiian Islands" (Pierce *et al.* 2014b). In this study, I found high heterogeneity in parasite prevalence both within

and between islands but found no evidence of host population structure. I also did not find evidence of host heterozygosity correlating with infection status. These results indicate that particular host genes, parasite genetics, or environmental factors may instead play a determining role in infection status and prevalence patterns.

In my final chapter, Chapter 6, I summarize my findings as well as describe limitations to my studies. The monarch butterfly genome has recently been sequenced (Zhan *et al.* 2011), but further exploration of the genome is needed to fully understand the migratory behavior. While some genes have been identified, like those associated with flight muscle (Zhan *et al.* 2014), we still lack a functional understanding of how migration genes ultimately enable the spectacular flight of several thousand kilometers (Ffrench-Constant 2014). Indeed, whether monarch butterflies are true navigators (meaning they are able to correct their directionality to reach a specific target), or instead employ vector navigation (meaning they simply fly in a southerly direction), is still hotly debated (Mouritsen *et al.* 2013a; Mouritsen *et al.* 2013b; Oberhauser *et al.* 2013). In this chapter I provide future directions in which questions such as these may be resolved.

Chapter 2

Lack of genetic differentiation between monarch butterflies with divergent migration destinations

Reprinted material from: J. I. Lyons*, **A. A. Pierce***, S. M. Barribeau, E. D. Sternberg, A. J. Mongue, & J. C. de Roode. *Molecular Ecology* 21, 14(Jul, 2012). *Co-first authorship listed in alphabetical order. Used by permission of the publisher, John Wiley and Sons.

Introduction

Each year, a wide variety of animals – including mammals, birds, fish and insects – undergo long-distance seasonal migrations to escape deteriorating habitats, colonize new resources and avoid predation, competition and parasitism (Alerstam 2006; Alerstam *et al.* 2003; Altizer *et al.* 2011; Dingle 1972, 1996; Fricke *et al.* 2011; McKinnon *et al.* 2010). Some of the most spectacular animal migrations involve the directed movement of millions of individuals across distances that span whole continents or hemispheres (Alerstam *et al.* 2003). Animal migration has great relevance to species persistence, ecosystem functioning and conservation biology (Bowlin *et al.* 2010), yet the mechanisms by which animals navigate between their seasonal migration sites are still unclear (Alerstam 2006).

Many migratory species consist of populations that utilize varying breeding sites and migratory destinations, and such variation may have important consequences for the genetic structuring of these populations (Haig *et al.* 1997). On the one hand, divergent migratory pathways and destinations may result in decreased opportunities for genetic mixing, and hence result in genetic divergence, which in some cases may result in speciation. For example, noctule bats – which migrate between hibernating and nursing sites – are genetically differentiated with respect to overwintering sites and migration direction (Petit & Mayer 2000). Similarly, beluga whales migrate between wintering sites in arctic pack ice and summering grounds in arctic and subarctic offshore waters, and genetic analysis has shown considerable levels of genetic

differentiation between belugas using different summering grounds (O'Corry-Crowe *et al.* 1997). On the other hand, the use of common breeding or overwintering grounds may result in a lack of genetic divergence, even when populations experience different selection pressures during part of their life. As one example, red-billed quelea birds in southern Africa undergo long-distance migrations in response to seasonal patterns of rain fall and grass seed production. Although different groups of birds move in north-westerly or south-easterly direction (Dallimer & Jones 2002), there is no genetic differentiation between these groups, probably because of genetic mixing that occurs when these birds re-colonize the same areas in the following season (Dallimer *et al.* 2003).

Monarch butterflies (*Danaus plexippus*) in different geographic areas have different migration strategies and thereby provide a suitable system to test the effects of divergent migration pathways on population differentiation. Monarchs are best known for their autumnal migration from eastern North America to Mexico; hundreds of millions of monarchs escape freezing temperatures and dying larval food plants in eastern North America, overwinter in the Oyamel fir forests of the Mexican Transverse Neovolcanic Range, and then re-migrate to eastern North America in the spring (Brower 1995; Urquhart 1976; Urquhart & Urquhart 1977, 1978). Allozyme analyses have indicated that these migrating monarchs form a large panmictic population, due to the genetic mixing of butterflies at the Mexican overwintering sites (Eanes & Koehn 1978). Monarch migration has captured the imagination of thousands of citizens, many of whom have helped to track the migratory routes that monarchs use on their way to their Mexican overwintering sites (Brower 1995; Urquhart & Urquhart 1977, 1978). The exact mechanisms by which eastern North American monarchs navigate to these sites remains a puzzle, although polarized light and circadian rhythms appear to be involved (Froy *et al.* 2003; Merlin *et al.* 2009; Reppert *et al.* 2004; Zhan *et al.* 2011; Zhu *et al.* 2008).

Monarch butterflies also occur in western North America (Brower 1995; Dingle *et al.* 2005; Tuskes & Brower 1978; Urquhart & Urquhart 1977). These monarchs overwinter in

Eucalyptus and Monterey Pine groves along the Californian Pacific Coast and are believed to be geographically separated from the eastern monarch butterflies by the Rocky Mountains (Fig. 1). Due to their different overwintering sites, eastern monarchs may fly up to 2500 miles to reach the Mexican Oyamel fir forests, while western monarchs generally reach the California Coast by flying less than 500 miles.

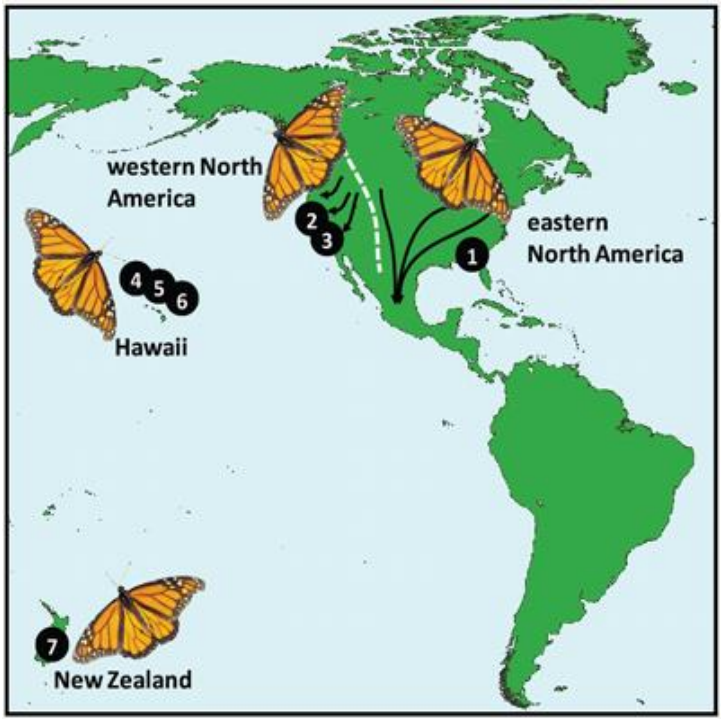

Figure 2.1. Map showing the location and migratory patterns of sampled populations. Numbers represent the sample sites as follows: 1-St. Marks, FL; 2-Pismo Beach, CA; 3-Santa Barbara, CA; 4-Kauai, Hawaii; 5-Oahu, Hawaii; 6-Maui, Hawaii; 7-Christchurch, New Zealand. The eastern North America population migrates between eastern North America and its Mexican overwintering site. The western North America population migrates along the California Pacific Coast and is believed to be geographically separated from the eastern monarchs by the Rocky Mountains (indicated in the white dashed line). The Hawaii and New Zealand populations are non-migratory.

Monarch butterfly migration has been the subject of decades' worth of research and has inspired large numbers of North American citizens. However, it is still unknown whether the different migratory pathways and destinations of eastern and western monarchs depend on genetic divergence of these butterflies, and disagreement on the occurrence and amount of gene

flow is ongoing (Brower & Pyle 2004; MonarchWatch 2011; Shephard *et al.* 2002; Urquhart & Urquhart 1977). In particular, Monarch Watch, an acclaimed educational outreach and citizen science research program proclaims that "Contact between eastern and western Monarchs is minimal suggesting that there is little exchange, or what scientists call gene flow, between these populations" (MonarchWatch 2011). This belief is widely held (e.g. Zhan *et al.* 2011), and federal regulations prohibit the shipment of monarchs across the continental divide. In contrast, limited genetic studies and flight observations of monarch butterflies in Rocky Mountain passes have led some authors to challenge the claim that eastern and western monarchs form distinct genetic populations (Brower & Pyle 2004; Shephard *et al.* 2002; Urquhart & Urquhart 1977).

Amid these conflicting views, we set out to determine whether eastern and western butterflies are genetically differentiated from each other. This is not only necessary to elucidate the genetics of monarch migration, but is also essential for monarch butterfly conservation. Monarch migration has been coined an endangered phenomenon, mostly due to the illegal deforestation of monarch overwintering sites in Mexico (Brower & Malcolm 1991). If eastern and western monarchs indeed form one genetic population, the protection of Mexican overwintering sites will not only be crucial for monarch migration in eastern North America, but also for migration in western North America (Brower & Pyle 2004).

Materials and Methods

Microsatellite development

We developed 17 polymorphic microsatellite markers to test whether eastern and western North American butterflies are genetically differentiated on the basis of these neutral genetic markers. Polymorphic microsatellite repeats were identified from a monarch expressed sequence tag (EST) database (Zhu *et al.* 2008) and Primer 3 (Rozen & Skaletsky 2000) was used to design

primers based on the contigs containing the repeats. Forward primers were fluorescently labeled on the 5' end with 6-FAM or HEX (see Table 1 for primer sequences and amplification details).

For PCR, genomic DNA was extracted from a 0.5 mm section of butterfly thorax (female butterflies) or abdomen (male butterflies) using the DNeasy Blood and Tissue Kit from Qiagen (Valencia, CA) and quantified using a Nanodrop 2000. We extracted DNA from females from the thorax rather than the abdomen to avoid the possibility of extracting male DNA in sperm transferred to the female abdomen. PCR was carried out in 15 µl multiplex reactions using the Type-It Microsatellite PCR kit (Qiagen). Each reaction contained 0.2 µM of each primer and 20-50 ng DNA template. Thermal cycling reactions for multiplex amplifications consisted of an initial 5 min at 95°C, followed by 28 cycles of 30 s at 95°C, 90 s at the primer-specific annealing temperature (see Table S1), and 30 s at 72°C. A final step of 30 min at 60°C was included to complete any partial polymerizations. Amplified DNA was genotyped on an ABI 3100 genetic analyzer (Perkin Elmer, Applied Biosystems Foster City, CA) at the Cancer Genomics Shared Resource (CGSR; Atlanta, GA) and alleles were scored using Genemarker v.4.0 (SoftGenetics LLC., State College, PA).

Table 2.1. Microsatellite loci developed in this study, showing locus name, multiplex reaction, fluorescent label, primers sequences, repeat motif and primer annealing temperature (T_A). Number of alleles and allele size range were determined by analyzing a total of 262 monarch butterflies, obtained from eastern North America (n=100), western North America (n=100), Hawaii (n=46) and New Zealand (n=16).

Locus	Multiplex reaction	Label	Primer sequence	Repeat motif	T_A	No. alleles	Allele size range
168	1	FAM	F: AGTTCAGGGTTTACGTGAGCA R: CATTATGTGAAGTGTTGCATGG	tcata	57°C	6	143-168
153	1	FAM	F: TGCGAAAAATGGTTTGAGGT R: TTATCGCCAAGTAAGTAATTTCG	ta	57°C	10	228-258
320	2	HEX	F: AATTTCTTGAGCGCTTTATCC R: CTGATCCTCGTCATCTCTCG	at	57°C	18	153-187
197	2	FAM	F: TGTCATTTCGATGTCGGCTA R: CAGAGAGAGCCTCGGGTAAA	att	57°C	4	174-183
208	3	FAM	F: TTTAGGACCCCAATCGGATTTTCG R: CGCGGACATTTTCACTTTCACGAT	at	60°C	19	178-242
203	3	HEX	F: TGACATACTTTATGTTCGTGGAAGG R: CCGCTCGCCTATATACAGGACACA	at	60°C	14	196-222
141	4	FAM	F: TCAAACCCGCATCCCTAGTGGTA R: TGGCAACGTACAGGGACGTGA	tc	60°C	13	150-178

1679	4	FAM	F: ATAGCCCTTCGACTTGTCGTTTCTC R:TCGACTGATGTTTTCGGGACTACGA	tat	60°C	4	215-224
137	5	HEX	F: AAGGTGGCGGTAAAAAGGCACAGA R: TCGCTTTCTTCCTCTTCCTCCTCA	aag	60°C	3	239-248
122	5	FAM	F: TTATAAGACCTCAACACCCACGAA R: CGCCGCTTCTAAATGAGTGGGATT	tta	60°C	6	228-252
494	6	HEX	F: CCGCGCTAGTCATTGTGTGAATGT R: CCTCGACTGATAGCCTTCGAAACG	att	60°C	7	160-181
983	6	FAM	F:AGACGCTTTGTTCAGCTTCGACCAC R: TTTACGATCACTCATACGAAACGGTA	ac	60°C	15	223-257
223	7	HEX	F: TCAAAGAATCCCGGAAACAG R: CGCTACAGTAGGAGGCAGGA	tg	52°C	21	182-248
854	8	HEX	F:AACGTCATCTGCACACGCCATACTA R:TCCAATTAAACGTGACGCCATTTTG	at	67°C	8	230-254
165	8	FAM	F:CCTCCGGAACCTGTCAAGAAAAAGA R:CACTCATCAGAACTGAAAAGTTCGAGACC	tat	67°C	8	189-213
819	8	FAM	F:GACTCGGAGACATGAGATCGACGAC R:TCGTCAGACAATTGCTCAAAATGGA	cacga	67°C	11	213-263
519	9	FAM	F:GTGGCGGGGCTTTGTGTAAATAAGA R:CAGGGTTCCATACAAACGTGTGATACAATA	att	63°C	15	221-263

Monarch butterfly collections

Our main interest was to estimate genetic differentiation between eastern and western North American monarch butterflies. However, to ensure that our microsatellite markers are able to detect population genetic differentiation we included non-migratory monarch butterflies from Hawaii and New Zealand in our analysis. Monarch populations in Hawaii and New Zealand were established within the last 170 years, and are thought to originate from North America through trans-Pacific dispersal (Vane-Wright 1993; Zalucki & Clarke 2004). Therefore, the inclusion of Hawaiian and New Zealand populations ensures that our markers are able to detect subtle and newly formed population structure. We obtained 100 monarch butterflies from St. Marks, FL (76 in October 2009; 24 in October 2010), a migration stopover of monarchs on their way to Mexico (Urquhart & Urquhart 1978). We also collected 100 monarch butterflies from the two biggest Californian overwintering sites in Pismo Beach (50 in February 2009; 34 February 2010) and Santa Barbara (12 in Nov 2009; 4 in Nov 2010). Finally, we obtained 46 butterflies from non-migratory populations in Hawaii (14, 15 and 17 from Oahu, Kauai and Maui

respectively in Nov 2009) and 16 non-migratory butterflies from New Zealand (Christchurch, Jan 2011).

Microsatellite analyses

We determined the genotype of each of the 262 butterflies at each of the 17 microsatellite loci. We then used the software Arlequin 3.5.1.2 (Excoffier & Lischer 2010) to calculate observed and expected heterozygosity at each microsatellite locus in each of 4 monarch populations: eastern North America, western North America, Hawaii and New Zealand. We also used Arlequin to calculate deviations from Hardy-Weinberg equilibrium for each locus in each population (a total of 68 statistical tests), and used a sequential Bonferroni correction (Rice 1989) to determine whether observed and expected heterozygosity levels were significantly different (α= 0.05). As shown in the results and Table 2, we discarded 6 of the 17 loci due to departure from Hardy-Weinberg equilibrium.

Table 2.2. Observed (H_o) and expected (H_e) heterozygosity at each locus within each of 4 populations, as determined by analyses in Arlequin 3.5.1.2 (Excoffier & Lischer 2010). Loci for which observed and expected hererozygosities are significantly different are indicated with asterisks; significance was determined using an α of 0.05 and a sequential Bonferroni correction (Rice 1989). Loci for which at least 3 populations were in Hardy-Weinberg equilibrium were used for subsequent analyses and are indicated in grey shading. Dashes indicate monomorphic loci.

	Eastern North America		Western North America		Hawaii		New Zealand	
Locus	**H_o**	**H_e**	**H_o**	**H_e**	**H_o**	**H_e**	**H_o**	**H_e**
168	0.74	0.64	0.64	0.60	0.72	0.59	0.19	0.18
153	0.56	0.67	0.57	0.68	0.52	0.66*	0.44	0.42
320	0.58	0.88*	0.63	0.88*	0.78	0.75	0.13	0.13
197	0.28	0.36	0.33	0.37	0.46	0.42	0.50	0.39
208	0.55	0.76*	0.36	0.76*	0.26	0.65*	0.44	0.46
203	0.67	0.78	0.78	0.82	0.48	0.65	0.63	0.53
141	0.50	0.63	0.53	0.70	0.65	0.63	0.31	0.37
1679	0.51	0.61	0.41	0.63*	0.43	0.56	0.31	0.51
137	-	-	0.05	0.05	-	-	-	-
122	0.02	0.04	0.01	0.03	0.30	0.27	-	-
494	0.29	0.28	0.32	0.32	-	-	-	-
983	0.49	0.77*	0.34	0.79*	0.33	0.50	0.31	0.61
223	0.27	0.88*	0.37	0.81*	0.14	0.66*	0.07	0.54*
854	0.30	0.59*	0.35	0.59*	0.37	0.66*	0.31	0.28
165	0.36	0.56*	0.18	0.48*	0.17	0.51*	0.38	0.51
819	0.79	0.85	0.77	0.83	0.63	0.66	0.69	0.66

519	0.83	0.80	0.84	0.80	0.65	0.67	0.81	0.68

Population genetic analyses

We used a series of analyses to test for genetic differentiation between monarchs from different populations. First, we used the Bayesian clustering analysis as implemented in the software STRUCTURE version 2.3.2.1 (Pritchard *et al.* 2000) to investigate population structure. We used an admixture model with uncorrelated allele frequencies to avoid the risk of overestimating the number of populations, K, and used the LOCPRIOR model to provide the software with location information (eastern North America, western North America, Hawaii and New Zealand) for each butterfly. We did the latter to ensure that STRUCTURE would be able to detect subtle population structure. We started simulations with K=7, to reflect the7 sample locations (Fig. 1), and then ran simulations for K values of 6, 5, 4, 3, 2 and 1. For each K, we ran multiple simulations to check for consistency between runs, using 100,000 burn-ins and 200,000 MCMC runs after burn-in. We then used log likelihood (Pritchard *et al.* 2000) and delta K (Evanno *et al.* 2005) to determine the most likely number of genetic populations present.

We also used F_{ST} and R_{ST} statistics (Holsinger & Weir 2009) to measure genetic differentiation between monarch populations. These statistics are frequently used to measure genetic differentiation, with levels of 0 indicating that individuals belong to the same panmictic population, and values higher than 0 indicating genetic differentiation. R_{ST} was developed as a more suitable statistic for microsatellite markers, based on its dependence on a stepwise mutation model (Slatkin 1995) instead of the infinite allele model that underlies F_{ST} statistics (Balloux & Lugon-Moulin 2002). However, because neither of these mutation models perfectly reflect natural mutation rates of microsatellites, studies on microsatellites often report both measures (Balloux & Lugon-Moulin 2002), and we followed this practice. We calculated pair-

wise F_{ST} and R_{ST} values between the 4 populations using the software Genepop version 4.1.0 (Rousset 2008), using the 11 microsatellite markers that were in Hardy-Weinberg in at least 3 out of 4 populations (see results and Table 2). To ensure that our estimates were not affected by the potential occurrence of null alleles, we re-calculated F_{ST} and R_{ST} values using corrected allele frequencies as determined by the software MICROCHECKER, version 2.2.3 (Van Oosterhout *et al.* 2004). Overall, statistics based on corrected and uncorrected allele frequencies were almost identical (see results), and resulted in identical conclusions.

Permutation tests (using 10,000 permutations), as implemented in the "Population comparisons" calculations in Arlequin 3.5.1.2 (Excoffier & Lischer 2010) were used to determine significance of pair-wise F_{ST} and R_{ST} values (uncorrected values were used). We also used F_{ST} and R_{ST} values to determine whether populations that are separated by greater geographic distances are genetically more differentiated (isolation by distance). We analyzed the correlation between geographic distance and measures of genetic differentiation using Mantel tests implemented in the vegan library (version 2.0-2) in the statistical package R (version 2.13.0). We ran Mantel tests on both corrected and uncorrected F_{ST} and R_{ST} measures, using 10,000 permutations.

To compare relative levels of genetic diversity between populations, we calculated genetic diversity (using the value 1-Qinter) and allelic richness in each population using Genepop version 4.1.0 (Rousset 2008). Furthermore, to understand the relative magnitude of within-and between-population genetic diversity, we carried out a locus by locus analysis of molecular variance (Excoffier *et al.* 1992) using Arlequin 3.5.1.2 (Excoffier & Lischer 2010). In this analysis, we combined eastern and western North American populations to compare genetic variation among geographic groups (North America versus Hawaii versus New Zealand), and compared this to the variation among populations within groups (i.e. variation among eastern and western North America) as well as genetic variation within populations (i.e. variation within

eastern North America, western North America, Hawaii and New Zealand). We again used 10,000 permutations.

Finally, we used the computer software POWSIM 4.1 (Ryman & Palm 2006) to determine the statistical power with which significant genetic differentiation could be determined using our microsatellite markers and observed allele frequencies. We restricted these analyses to eastern and western North America only (since these were the only populations between which we did not detect genetic differentiation; see results), and simulated the sampling of 100 individuals into two populations based on a random drawing of alleles that occurred at the observed overall frequency in eastern and western North America (as determined by MICROCHECKER-corrected allele frequencies). Simulations were carried out using a series of dictated F_{ST} values, and 60-200 runs for each value. Statistical power was then determined as the proportion of simulations for which Fisher's exact and Chi-square tests showed a significant deviation from 0 (i.e. significant genetic differentiation). Note that this software is set up for power calculations on the basis of F_{ST} values only, so we were not able to calculate power on the basis of R_{ST} values. However, because F_{ST} and R_{ST} values were similar our study (see Table 3), and because F_{ST} calculations generally provided slightly lower estimates of genetic differentiation than did R_{ST} calculations (see Table 3), the power calculations on the basis of F_{ST} values provide a conservative estimate of power to detect genetic differentiation using our genetic markers and sample sizes.

Results

A total of 6 out of 17 loci were out of Hardy-Weinberg equilibrium in at least 2 of the 4 populations, and we excluded these loci from subsequent analyses (see Table 2 for details). Of the remaining 11 loci, 9 were in Hardy-Weinberg equilibrium in all populations, and 2 were in Hardy-Weinberg equilibrium in 3 out of 4 populations.

Clustering analyses in STRUCTURE suggested that our monarch butterflies most likely form three genetically distinct populations: eastern + western North America, Hawaii, and New Zealand (Fig. 2). Thus, we found no significant genetic differentiation between eastern and western North American butterflies. Additionally, neither the Hawaiian Islands, nor the two Californian overwintering sites, are genetically distinguishable. This lack of genetic structure is unlikely to be an artifact of our microsatellite markers because these markers clearly pick up the genetic differentiation of Hawaiian and New Zealand monarchs from each other and from North American monarchs (Fig. 2). Our results are unaltered when excluding the populations for which 2 of the 11 loci are out of Hardy-Weinberg equilibrium.

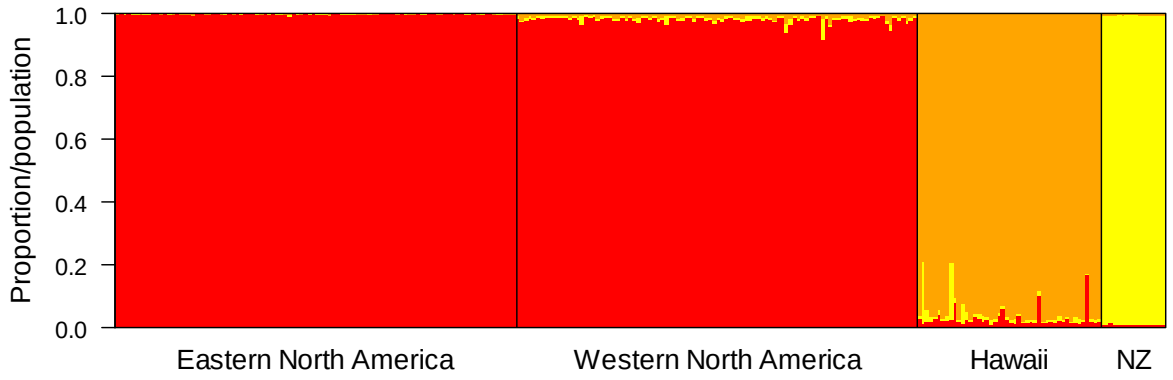

Figure 2.2. Inferred genetic proportion of individual butterflies to each of three genetic populations. Analyses were started with seven populations to reflect the seven sample locations (Fig. 1A). Bayesian analysis suggest that there are three genetic populations: eastern+western North America (indicated in red), Hawaii (orange), and New Zealand (NZ; yellow). Individual monarchs are indicated by vertical bars.

We confirmed our results by calculating genetic differentiation using F_{ST} and R_{ST} statistics. We first calculated pair-wise genetic differentiation using the same 11 loci that we used for genetic structure analysis using Genepop version 4.1.0 (Rousset 2008). We then calculated pair-wise F_{ST} and R_{ST} values based on corrected genotype and allele frequencies as obtained by the software MICROCHECKER (Van Oosterhout *et al.* 2004). For both uncorrected and

corrected allele frequencies, both F_{ST} and R_{ST} values were much lower for the comparison between eastern and western North America than for any of the other comparisons (Table 3).

Table 2.3. Pairwise F_{ST} and R_{ST} values between the four studied monarch butterfly populations, as calculated in Genepop version 4.1.0 (Rousset 2008). Values in parentheses are based on corrected allele and genotypes frequencies as determined by the software MICRO-CHECKER, version 2.2.3 (Van Oosterhout *et al.* 2004). Asterisks denote values that are significantly different from 0. The values of pairwise differentiation between eastern and western North America are not significantly different from 0 (P=0.20 for F_{ST}; P=0.43 for R_{ST}).

	Western North America	**Hawaii**	**New Zealand**
Eastern North America	F_{ST}: 0.0012 (0.0009) R_{ST}: -0.003 (-0.0006)	F_{ST}: 0.0401 (0.0388)* R_{ST}: 0.0662 (0.0607)*	F_{ST}: 0.1858 (0.1856)* R_{ST}: 0.1810 (0.1789)*
Western North America		F_{ST}: 0.0456 (0.0435)* R_{ST}: 0.0557 (0.0511)*	F_{ST}: 0.1753 (0.1745)* R_{ST}: 0.1787 (0.1790)*
Hawaii			F_{ST}: 0.1750 (0.1714)* R_{ST}: 0.0873 (0.0861)*

Indeed, population comparison calculations in Arlequin (Excoffier & Lischer 2010) suggest that the low F_{ST} and R_{ST} values calculated for the eastern-western North American comparison are not significantly different from 0, suggesting a lack of genetic differentiation between these populations (Table 3). In contrast, our analyses suggest that all other pair-wise F_{ST} and R_{ST} values are significantly different from 0, suggesting significant genetic differentiation between monarchs in North America, Hawaii and New Zealand (Table 3). Moreover, butterflies were more differentiated from each other when they were farther apart geographically (Fig. 3), suggesting that greater geographic distances reduce levels of gene flow.

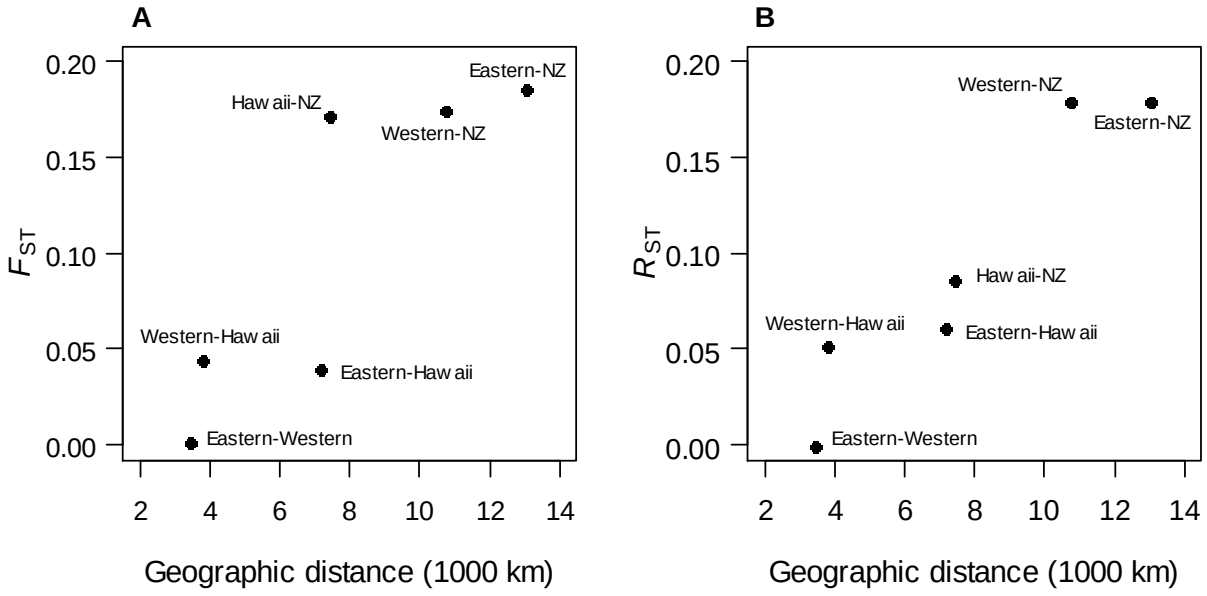

Figure 2.3. Isolation by distance. There is a strong pattern of isolation by distance demonstrated by a correlation between geographic and genetic distance, whether based on F_{ST} (panel A: r=0.85, P=0.040) or R_{ST} (panel B: r=0.95, P=0.039). Note that the graphs show corrected F_{ST} and R_{ST} values. Similar results were obtained for uncorrected values (F_{ST}: r=0.85, P=0.042; R_{ST}: r=0.95, P=0.037).

To study genetic diversity within and between populations, we first used the 11 microsatellite loci to calculate genetic diversity (using the value 1-Qinter) and allelic richness in each population using Genepop version 4.1.0 (Rousset 2008). Levels of genetic diversity were similar in eastern (0.515) and western (0.535) North America and were slightly lower in Hawaii (0.464) and New Zealand (0.339; Fig. 4A); however, although genetic diversity appeared to be lower in New Zealand than in the other populations, this was marginally non-significant ($F_{1,42}$=3.10, P=0.085). Allelic richness showed a similar and significant pattern, being highest – and similar – in eastern and western North America, and significantly lower in New Zealand (Fig. 4B; $F_{3,40}$=3.19, P=0.034). Our results thus showed similar levels of genetic diversity in eastern and North America, and a clear trend of decreasing genetic diversity and allelic richness with increasing distance from North America (Fig. 4A, B). This is consistent with the hypothesis that monarch butterflies colonized the Pacific Ocean in a stepwise fashion from an origin in North America (Vane-Wright 1993).

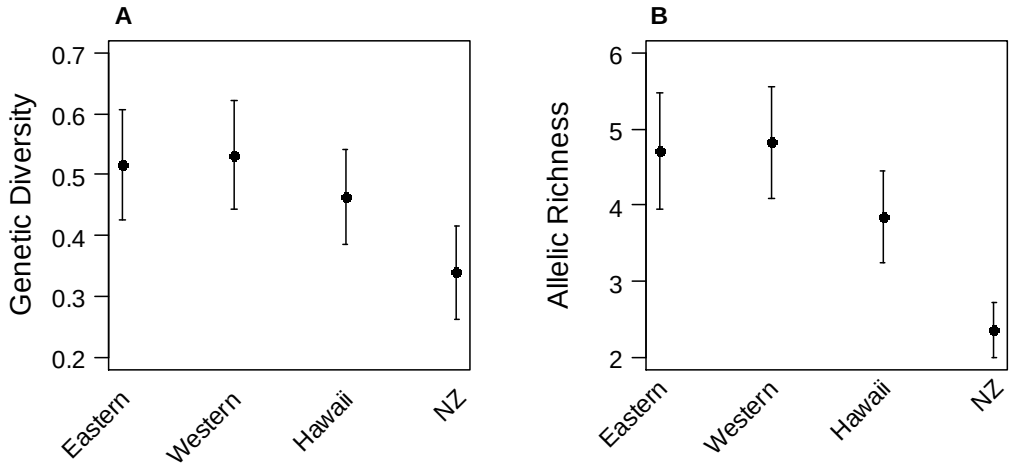

Figure 2.4. Measures of genetic diversity for eastern North America, western North America, Hawaii, and New Zealand (NZ) monarchs. (**A**) Genetic diversity (using the value 1-Qinter, the inter-individual diversity within populations) was highest in eastern and western North America and tended to decrease with increasing distance from North America. (**B**) Allelic richness was highest in eastern and western North America and significantly lower in New Zealand. Error bars show ± 1 s.e. across 11 loci.

Analysis of molecular variance (AMOVA) further confirmed a lack of genetic differentiation between eastern and western North America (Table 4). In particular, although a significant amount of genetic variation (8.52% and 9.33% for F_{ST}- and R_{ST}-based calculations respectively) was explained by different groups (i.e. genetic variation between North America, Hawaii and New Zealand), a mere and non-significant amount of variation (0.18% and 0.037% for F_{ST}- and R_{ST}-based calculations respectively) was explained by the North American east-west division (Table 4). This result was confirmed when carrying out an analysis of molecular variance on the North American populations only, which again showed that only a minimal amount of variation was explained by the North American east-west division (0.17% and 0.023% for F_{ST}- and R_{ST}-based calculations respectively).

Table 2.4. Results of Analysis of Molecular Variance (AMOVA) comparing samples from 4 populations (eastern North America, western North America, Hawaii and New Zealand). In this analysis, eastern and western North America were grouped into the same group (North America) while Hawaii and New Zealand formed their own groups. The analysis was done based on F_{ST} and R_{ST} values; results for the latter are shown in parentheses. Significant P-values, as based on permutation tests in Arlequin 3.5.1.2 (Excoffier & Lischer 2010) are indicated with asterisks.

Source of variation	d.f.	Sum of squares	Variance components	Percentage variation	P-value
Among groups	2	57.95 (5386)	0.257 (24.41)	8.52 (9.33)	0.00000 (0.00000)*
Among populations within groups	1	3.835 (257)	0.0054 (0.09767)	0.18 (0.037)	0.21017 (0.44379)
Among individuals within populations	520	1434.9 (123360)	2.759 (237.23)	91.31 (90.64)	0.00000 (0.00098)*
Total	523	1496.7 (129002)	3.022 (261.74)	100 (100)	

Finally, power calculations using the software POWSIM (Ryman & Palm 2006) suggest that our microsatellite markers and sample sizes (100 butterflies from both eastern and western North America) have sufficient statistical power to detect significant population differentiation on the basis of F_{ST} values as low as 0.0025 (Fig. 5). Thus, the lack of genetic differentiation between eastern and western North American butterflies is unlikely to be the result of inadequate molecular markers or sample sizes, and is more likely to reflect genetic mixing between these butterflies.

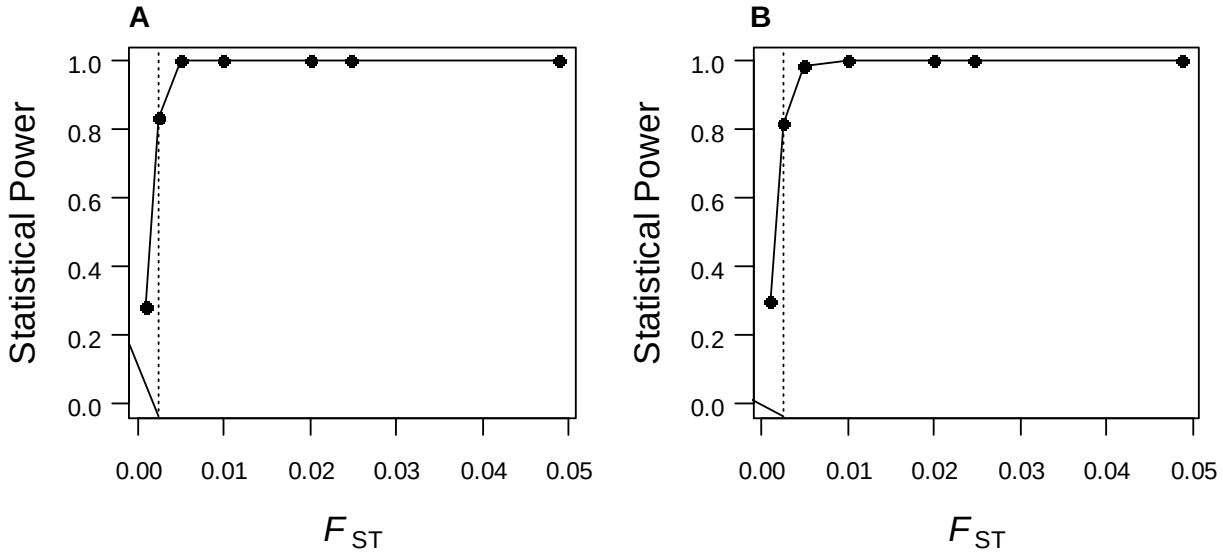

Figure 2.5. Statistical power to detect significant population genetic differentiation as a function of F_{ST}. Statistical power was determined as the proportion of POWSIM 4.1 (Ryman & Palm 2006) simulations for which Fisher's exact (A) and Chi-square tests (B) showed a

significant deviation from 0 (i.e. significant genetic differentiation). Vertical dashed lines indicate that F_{ST} values as low as 0.0025 can be detected with more than 80% probability.

Discussion

Our analyses suggest that eastern and western North American monarch butterflies form one panmictic population. These results are surprising, since these monarchs inhabit different areas of North America, migrate varying distances, and overwinter at different sites in Mexico and along the Pacific Coast, respectively. Our results suggest that North American monarchs form an admixed population, and that eastern and western migratory pathways are maintained despite extensive gene flow. These results suggest that migratory differences do not require – or result in – substantial genome-wide genetic differentiation as picked up by neutral genetic markers, and may instead be driven by two alternative mechanisms.

First, despite a lack of genetic differentiation of neutral genetic markers, eastern and western monarchs may show genetic divergence of particular genes that are involved in migration and that are under strong selection. Such a scenario was suggested for European willow warblers, which occur as European populations that either migrate from northern Scandinavia to eastern and southern Africa or from southern Scandinavia to western Africa: despite morphological divergence, these birds did not display divergence of mitochondrial and microsatellite DNA (Bensch *et al.* 1999). Second, divergent migration may not be subject to genetic differentiation, but may instead be based on differential gene expression based on varying and changing environmental conditions (Liedvogel *et al.* 2011). Such a scenario has been suggested for North American populations of Mexican free-tailed bats, which are not genetically differentiated despite their varying migration routes and overwintering sites in Mexico (Russell *et al.* 2005). The recently sequenced genome of the monarch butterfly (Zhan *et al.* 2011) will be a valuable source for testing these alternative hypotheses. In particular, re-sequencing the genomes of migratory and non-migratory butterflies will allow for the identification of genomic regions that are associated with migratory behaviors and for the

detection of differences in these regions between eastern and western North American butterflies. In addition, next-generation sequencing of the transcriptome of eastern and western migratory butterflies may reveal differential expression of genes resulting in divergent migrations; such an approach has already revealed differential expression of genes in breeding and migratory monarchs in the eastern North American population (Zhu *et al.* 2009; Zhu *et al.* 2008).

Although our results did not show genetic differentiation between eastern and western North American butterflies, we found that Hawaiian and New Zealand monarch butterflies are differentiated from North American butterflies. Although these two populations are non-migratory, these genetic differences from migratory monarchs do not necessarily mean that their lack of migration is genetically determined, as they are also differentiated from one another. Instead, their divergence is consistent with their recent dispersal across the Pacific Ocean from a source population in North America (Shephard *et al.* 2002; Vane-Wright 1993; Zalucki & Clarke 2004).

Until now, the question of genetic mixing of eastern and western North American butterflies had been unresolved, partly because of low levels of polymorphism of genetic markers in previous studies (Brower & Boyce 1991; Brower & Jeansonne 2004). However, the lack of genetic differentiation between eastern and western monarchs are consistent with a study on allozymes (Shephard *et al.* 2002) as well as a series of observational studies (Brower & Pyle 2004). In particular, western monarchs have been observed to fly in south-easterly directions, and to follow migratory pathways that are – when extrapolated – consistent with overwintering sites in Mexico (Dingle *et al.* 2005). Moreover, population sizes of overwintering western and eastern monarchs are generally correlated (Vandenbosch 2007). Finally, an overabundance of re-migrating western monarchs and a lack of eastern spring re-migrants in 1996 coincided with a westward shift of spring wind patterns and a corresponding shift of the

northward spring migration of song birds (Brower & Pyle 2004), suggesting that eastern and western monarchs do at least occasionally intermix.

Genetic mixing of eastern and western monarch populations does not invalidate the claim that eastern and western monarchs are subject to different selection pressures (Altizer *et al.* 2011; Altizer & Davis 2010; Brower *et al.* 1995). Although our results suggest that eastern and western butterflies form one panmictic population, genetic exchange is probably subject to strong seasonality, occurring during the overwintering and spring re-migration of these butterflies. This exchange is followed by a long breeding season during which there may be ample opportunity for natural selection to favor those genotypes that best suit eastern and western habitats (Altizer & Davis 2010; Dingle 1972). For example, eastern and western North American butterflies have divergent wing morphology, which is likely the result of differential selection (Altizer & Davis 2010). Moreover, strong selection may favor different genotypes at important genetic loci that go undetected with approaches based on neutral markers such as microsatellites (Bensch *et al.* 1999; Liedvogel *et al.* 2011). Until such differential selection is better understood, our results do not warrant a relaxation of the current regulations to restrict the human-facilitated movements of eastern and western monarchs (Brower *et al.* 1995). In addition, previous work has shown that western butterflies are subject to more virulent protozoan parasites than eastern butterflies (De Roode & Altizer 2010; De Roode *et al.* 2008). Thus, even if eastern and western monarch butterflies are genetically similar across their full genome, cross-continental shipments of monarchs may result in the unwanted transfer of virulent parasites (Brower *et al.* 1995).

Our findings have strong relevance to the conservation of the spectacular migration of monarch butterflies. Monarch butterfly migration is at risk (Brower & Malcolm 1991), partly due to the illegal deforestation of Mexican monarch butterfly overwintering sites. It is well known that there are many more butterflies in eastern than western North America, and it has been suggested that the western sub-population requires influxes from eastern North America for its

survival (Brower & Pyle 2004). Our results support this hypothesis, by suggesting that eastern and western monarchs form one genetic population. As such, the conservation of Mexican overwintering sites may be essential not only to protect eastern monarch migration, but also to conserve monarchs and their migration in western North America.

Chapter 3

Unraveling the mysteries of monarch migration and global dispersal through molecular genetic techniques

Reprinted material from: A.A. Pierce, N. Chamberlain, M. Kronforst, J.C. de Roode. In: *Monarchs in a changing world: biology and conservation of an iconic insect* (eds. Oberhauser K, Nail KR, Altizer S).

Introduction

Monarchs are iconic insects best known for their spectacular annual fall migration from Canada and the United States to Mexico (Urquhart & Urquhart 1978). Despite decades of study, many aspects of this migration remain clouded in mystery (Brower 1995, 1996). Unsolved questions regarding the monarchs' southward fall migration focus on the mechanisms by which monarchs orient towards and locate their overwintering sites (Guerra *et al.* 2012; Merlin *et al.* 2009; Reppert *et al.* 2004; Zhan *et al.* 2011; Zhu *et al.* 2009; Zhu *et al.* 2008) and the relative role of active navigation and passive wind-based movement (Wenner & Harris 1993). Importantly, monarchs occupy locations ranging from the New World tropics to more recently-colonized Pacific islands (Ackery & Vane-Wright 1984) to Europe, and most of these populations are non-migratory (Altizer *et al.* 2000; James 1993). A population is deemed non-migratory if their habitat range does not significantly differ throughout the year. The wide distribution of monarchs raises important questions regarding the genetic differences between and interconnectedness among existing populations (Lyons *et al.* 2012). For example, is genetic separation a prerequisite for monarch populations to maintain different migration strategies and destinations; and have non-migratory populations repeatedly arisen from migratory ancestors? The occurrence of monarchs around the world raises the additional question of where all these monarchs came from. Which ancestral populations served as sources for more

recent colonization events, and in what patterns did monarchs spread across the Atlantic and Pacific Oceans from the New World (Vane-Wright 1993; Zalucki & Clarke 2004)?

Given the popularity of monarchs and the long history of scientific study focused on them, it is perhaps surprising that many questions regarding their migration and evolutionary history remain unanswered. However, modern molecular genetic approaches are required to investigate many of these issues, and these techniques have not been widely available for monarchs until recently. The recent development of microsatellite markers (Lyons *et al.* 2012), the publication of the monarch's genome sequence (Zhan *et al.* 2011), and the genomic work that preceded it (Zhu *et al.* 2008) now offer the potential to explore many aspects of monarch biology from a genetic perspective. In this chapter, we address some of these questions and discuss answers offered by recent analyses. We begin by summarizing early genetic work on monarchs based on studies of allozyme variation and mitochondrial DNA. We then describe how the use of microsatellite markers provides insight into the genetic connectedness between monarch populations and worldwide monarch dispersal, and how these markers may change our thinking on monarch migration. Finally, we briefly describe how genomic work has provided insights into monarch navigation and how it is likely to improve our understanding of monarch migration.

Allozyme markers show seasonal mixing and shed light on the origins of Pacific monarchs

Migration can affect levels of gene flow within and genetic differentiation among populations, especially when species occupy multiple breeding grounds and migration destinations (Haig *et al.* 1997). Divergent migratory pathways and destinations could lower opportunities for genetic mixing, and hence result in genetic divergence. For example, the Old World noctule bats, *Nyctalus noctula*, which migrate between hibernating and summer nursing sites, are genetically differentiated with respect to overwintering sites and migration flyways (Petit & Mayer 2000).

Similarly, beluga whales, *Delphinapterus leucas*, migrate between wintering sites in arctic pack ice and summering grounds in arctic and subarctic offshore waters, and there are considerable levels of differentiation between belugas using different summering grounds (O'Corry-Crowe *et al.* 1997). In contrast, the use of common migratory flyways, breeding grounds, or overwintering areas can lead to high levels of genetic mixing, even when populations experience different selection pressures or population sub-structuring at other points in their migratory cycle. For example, red-billed quelea birds, *Quelea quelea*, in southern Africa undergo long-distance migrations in response to seasonal patterns of rain fall and grass seed production. Although different groups of birds move in north-westerly or south-easterly directions (Dallimer & Jones 2002), genetic analysis shows high levels of mixing between these groups, which probably occurs when birds re-colonize the same areas in the following season (Dallimer *et al.* 2003).

As with other species, the use of widely dispersed breeding grounds, distinct wintering sites, and different migratory flyways could cause local genetic differences in migratory monarch populations. The earliest published population genetic study on monarchs examined this issue using allozymes. In a seminal paper, Eanes and Koehn (1978) examined the population structure of eastern North American monarchs by collecting 30 geographic samples throughout the monarchs' summer breeding grounds and along their fall migration routes in the eastern United States. Using six allozyme loci, Eanes and Koehn found differentiation between monarch groups during the summer breeding season. They hypothesized that this differentiation may result from genetic drift, which involves random processes that alter allelic frequencies such as founder effects. Because monarchs are more regionally contained during the summer than during the migration, such random effects could result in differences in allele frequencies between sub-populations. It is also possible that differential selection could cause allele frequencies to vary among sampling locations. Such differential selection may act on allozyme markers, some of which have been linked with flight metabolism (Hughes & Zalucki 1993; Solensky & Oberhauser 2009; Zalucki *et al.* 1993). Eanes & Koehn also found that the annual

migration erased the genetic differentiation detected across summer breeding sites, by mixing monarchs from different breeding regions. Thus, as with red-billed quelea birds (Dallimer *et al.* 2003), monarchs originating from a range of breeding sites appear to mix randomly during the migration season.

Allozymes have also helped elucidate the genetic origin of monarchs in Australia and showed that the lack of migration of Australian monarchs affects their genetic structure (Shephard *et al.* 2002). Monarchs likely spread beyond the New World within the last 200 years, across the Pacific and Atlantic Oceans to destinations as distant as Australia and Spain (Fig. 1). But the exact routes by which they have done so remain unclear (Vane-Wright 1993; Zalucki & Clarke 2004); in particular, did monarchs spread in a stepwise fashion from North America to far-flung locations across the Pacific and Atlantic oceans (Vane-Wright 1993), or did multiple independent dispersal events occur (Zalucki & Clarke 2004)? To address this question, Shephard and colleagues (2002) collected 1194 butterflies from 15 sites in Australia, North America, and Hawaii. They found that the North American monarchs had more allelic diversity than monarchs from Hawaii and Australia, and that Australia and Hawaii had different subsets of alleles. These results suggest that both Hawaiian and Australian monarchs are derived from North America, but that the colonization of each location resulted from an independent dispersal event. Had the Australian population derived from the Hawaiian population, the alleles found in Australia would have been a subset of those found in Hawaii. Shephard and colleagues also found that monarchs obtained from different regions in Australia were more similar genetically than they were to monarchs from either Hawaii or North America, suggesting that monarchs colonized Australia in a single event. The lack of genetic differentiation among Australian sites also suggests seasonal mixing; this supports the hypothesis that, despite their lack of a two-way migration, Australian monarchs undergo alternating bouts of seasonal dispersal and range contraction through which monarchs from different regions mix (James 1993).

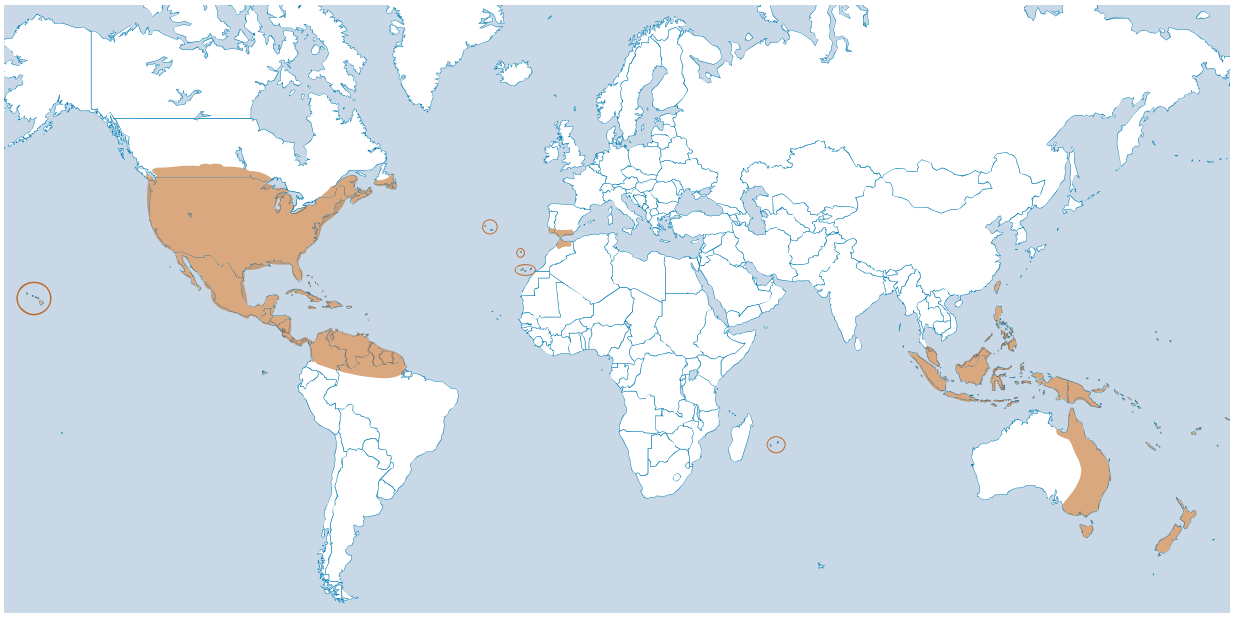

Figure 3.1. Worldwide distribution of monarchs. Orange shading and circles indicate known monarch range, following Ackery and Vane-Wright (1984) and updated following Neves et al. (2001) and personal communications.

East meets west: on the origins of and mixing between North American monarchs

One long-standing question in monarch biology has been whether monarchs in the eastern and western regions of the United States and Canada are genetically similar or distinct. Although monarchs are best known for their migration from eastern Canada and the United States to the Oyamel fir forests in central Mexico, monarchs in the western states embark on a shorter-distance migration to overwinter in eucalyptus and Monterey pine groves along coastal California (Brower 1995; Dingle *et al.* 2005; Tuskes & Brower 1978; Urquhart & Urquhart 1977). Conventional wisdom suggested that these monarchs were geographically separated by the Rocky Mountains and continental divide (Fig. 2A). Moreover, due to their different overwintering sites, eastern monarchs can travel up to 3500 km during the fall migration (Urquhart & Urquhart 1979), while western monarchs generally fly shorter distances, usually less than 500 km. Until recently, it was unknown whether the different migratory pathways and destinations of eastern and western monarchs depend on, or have resulted in, genetic

divergence of these butterflies, and disagreement on the amount of gene flow is ongoing (Brower & Pyle 2004; MonarchWatch 2011; Shephard *et al.* 2002; Urquhart & Urquhart 1977).

Brower and Boyce (1991) used mitochondrial DNA markers to determine whether eastern and western migratory monarchs were genetically differentiated, and to compare these monarchs with those from ancestral populations in the neotropics. Mitochondrial DNA (mtDNA) is a useful genetic marker for several reasons. First, it is relatively easy to replicate and thus generates high numbers of copies to study in the laboratory. Second, it contains highly conserved regions (found across a wide range of species) that make it possible to use the same primers for replication across species and populations. These conserved regions surround regions with an elevated mutation rate that can cause differences between isolated populations of the same species. Brower and Boyce used 12 butterflies from each of the eastern and western migratory groups, and also included monarchs from the West Indies islands of Trinidad and Tobago. Surprisingly, in the fragments of mtDNA that they examined, they found virtually identical patterns in all of the populations. The only unique variants discovered were in one individual from the eastern population and one individual from the western population. This indicates that based on mtDNA fragments alone, eastern and western migratory monarchs cannot be distinguished from each other, nor can they be distinguished from monarchs in the neotropics, despite the large distances and geographic barriers separating them (Brower & Boyce 1991).

Brower and Boyce hypothesized three non-exclusive explanations for such low polymorphism and variability: low mutation rates, stabilizing natural selection, and random processes. However, they believed the most plausible of these explanations lay in random processes driven by a recent genetic bottleneck. A bottleneck would reduce overall levels of genetic diversity, especially because mtDNA is maternally inherited and genetic recombination does not occur. Thus, their data indicate that monarchs may have experienced a significant population reduction sometime in the recent past, followed by a rapid radiation into the

temperate zone from the tropics. This was later confirmed by Brower and Jeansonne (2004), when mtDNA markers indicated a lack of genetic divergence between North American and South American monarchs, despite clear differences in morphology and behavior.

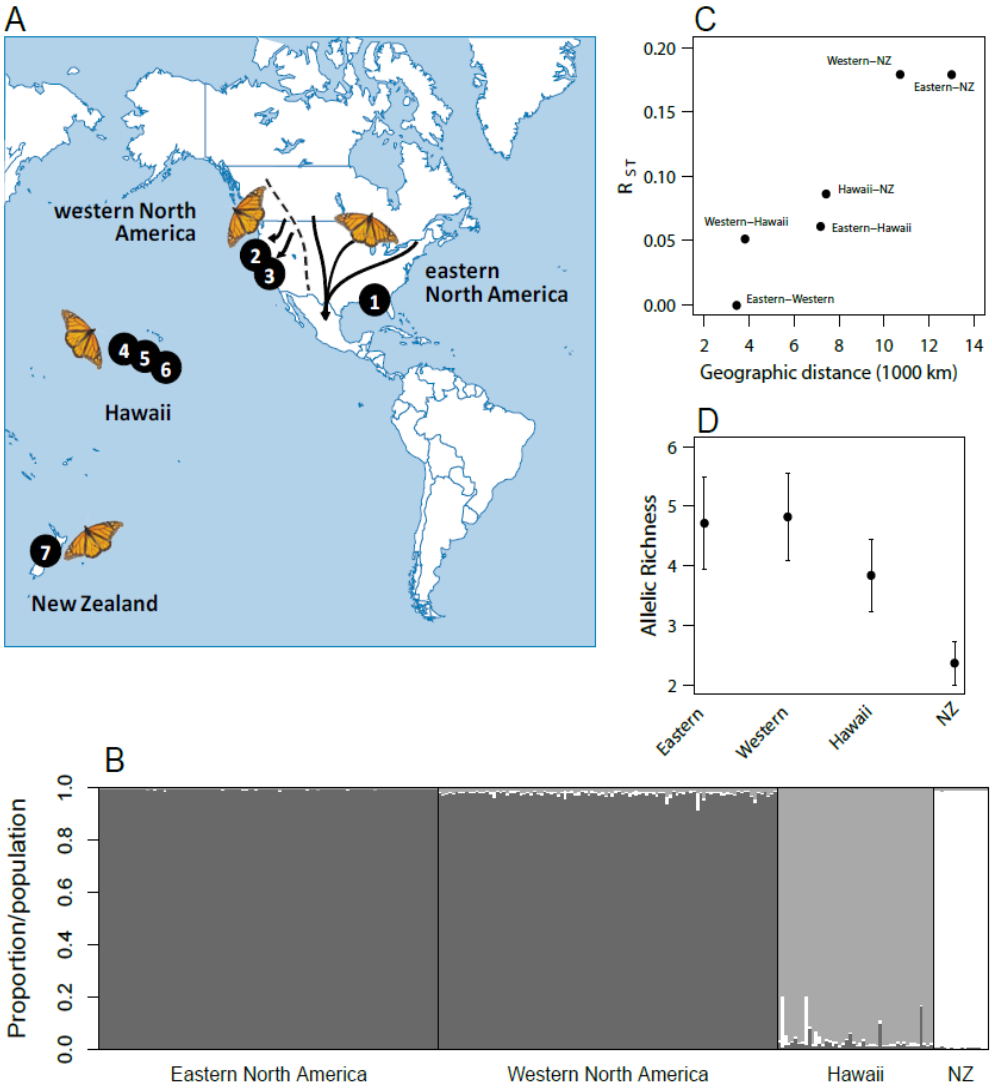

Figure 3.2. East meets west (adapted from Lyons et al. 2012). A) Location and migratory patterns of sampled populations (Lyons *et al.* 2012): 1-St. Marks FL (n=100); 2-Pismo Beach CA (n=84), 3-Santa Barbara CA (n=16); 4-Kauai HI (n=15); 5-Oahu HI (n=14); 6-Maui HI (n=17); 7-Christchurch, New Zealand (n=16). Eastern and western North America populations are at least somewhat separated by the Rocky Mountains (black dashed line). Hawaii and New Zealand populations are non-migratory. B) Inferred genetic proportion of individual butterflies from each of three populations. Genetic clustering analysis was used to determine the likely proportion of alleles of each butterfly that originates from each of three genetic populations: eastern+western North America (dark grey), Hawaii (light grey), and New Zealand (NZ; white). Individual monarchs are indicated by vertical bars; bars for butterflies with

alleles from different populations are divided into different portions accordingly. Genetic asignments were determined on the basis of 11 microsatellite loci using the software STRUCTURE (Pritchard *et al.* 2000). C) There is a strong pattern of isolation by distance demonstrated by a correlation between geographic and genetic distance (r=0.95, P=0.037). R_{ST} values closer to zero indicate a lack of genetic differentiation. D) Allelic richness was highest in North America and significantly lower in New Zealand. Error bars show ± 1 s.e. across 11 loci.

To better resolve genetic variation within and among contemporary monarch populations, Lyons et al. (2012) recently developed microsatellite markers. Microsatellites are selectively neutral markers comprised of repeats of nucleotide sequences that are scattered throughout the genome and tend to show extreme variability, which makes them ideal for studying genetic variation within and between populations. To compare eastern and western migratory monarchs, Lyons and colleagues collected 100 butterflies from St. Marks Florida, a stopover location along the eastern autumn migration to Mexico (Urquhart & Urquhart 1978), and another 100 butterflies from Pismo Beach and Ellwood California, two western North American overwintering sites. To ensure that their microsatellite markers were able to detect subtle genetic differentiation, they also included monarchs from Hawaii and New Zealand in their analysis (Fig. 2A).

Using a set of 11 polymorphic microsatellite markers and a series of population genetic analysis tools, Lyons et al. (2012) found that, despite differences in migration destination and the Rocky Mountains to serve as a potential barrier, eastern and western North American monarchs are genetically indistinguishable on the basis of their microsatellite genetic make-up. Using a genetic clustering analysis to determine the most likely number of genetic populations from which these 262 butterflies were derived (Fig. 2B), this work offered support for three, rather than four, genetically distinct populations. Briefly, this clustering analysis assigned individual butterflies to a source population based on which alleles are present in each butterfly, and whether those alleles appear to be shared across all populations, or restricted to a single population. In some cases, an individual butterfly will show up as a single-color bar in Fig. 2B; if a butterfly has alleles representative of two or more populations, it will show up as a bar with

multiple colors. The results of this clustering analysis showed that eastern and western monarchs belong to a single genetic population (all North American monarchs are all represented by dark grey bars in Fig. 2B). Thus, sufficient gene flow exists between eastern and western migratory monarchs to homogenize the selectively neutral molecular variation examined in this analysis.

The results of Lyons et al. (2012) indicate that eastern and western migratory monarchs regularly exchange genes, despite the Rocky Mountains separating their breeding ranges. Such a conclusion was suggested by Shephard and colleagues, who found high levels of gene flow between monarchs from California and Michigan based on allozyme analysis (Shephard *et al.* 2002). How these genetic exchanges occur is not clear, although it has been suggested that monarchs dispersing from Mexico in the spring can populate areas in the western United States in high numbers (Brower & Pyle 2004; Vandenbosch 2007), and some monarchs tagged in the west are retrieved at overwintering sites in Mexico (Southwest Monarch Study 2012).

Results of Lyons et al. (2012) further suggest that large-scale genetic differentiation is neither a prerequisite for, nor a result of, differential migration of eastern and western monarchs. This does not necessarily mean that there is no differential selection operating on eastern and western monarchs; despite a lack of genetic differentiation of neutral genetic markers, eastern and western monarchs could still show divergence of particular genes that are involved in migration and that are under strong selection. Alternatively, divergent migration pathways might arise from differential gene expression (based on varying and seasonally changing environmental conditions) rather than genetic differences *per se* (Liedvogel *et al.* 2011). Such a scenario has been suggested for North American populations of Mexican free-tailed bats, *Tadarida brasiliensis*, which are not genetically differentiated despite their varying migration routes and overwintering sites (Russell *et al.* 2005). Next-generation sequencing of the transcriptome (which provides information on the genes that are being transcribed at a given time) of eastern and western migratory butterflies might reveal differential expression of

genes resulting in divergent migrations; such an approach has already revealed differential expression of genes in breeding versus migratory monarchs in the eastern United States and Canada (Zhu *et al.* 2009; Zhu *et al.* 2008).

Although Lyons and colleagues did not find genetic differentiation between eastern and western migratory butterflies, they did find that Hawaii and New Zealand monarchs are differentiated from North American monarchs as well as from each other (Fig. 2B). They also found that monarchs were more differentiated from one another when they were farther apart geographically (Fig. 2C), suggesting that greater geographic distances reduce levels of gene flow. Levels of genetic diversity, as measured by allelic richness, appeared similar in the eastern and western United States, decreased in Hawaii and decreased further in New Zealand (Fig. 2D). This trend of decreasing genetic diversity with increasing distance from North America is consistent with the hypothesis that monarchs dispersed across the Pacific Ocean from an origin in North America (Clarke & Zalucki 2004; Vane-Wright 1993; Zalucki & Clarke 2004). Moreover, the lower genetic diversity in New Zealand than Hawaii is consistent with serial dispersal events, each leading to an additional loss in genetic diversity. However, an alternative hypothesis for this pattern is that monarchs dispersed to Hawaii on more occasions than to New Zealand. Further studies, including a greater number of Pacific islands, are necessary to distinguish between these hypotheses.

Does open water impede gene flow?

The long-distance migration and dispersal of monarchs across the globe demonstrate that monarchs have strong flight ability and therefore the potential to exchange genes between distant geographic regions. At the same time, the study by Lyons et al. (2012) suggests that although monarchs can traverse seas and oceans to colonize distant islands, such large water bodies also present barriers (albeit imperfect) to high and recurrent gene flow. We tested this idea in a separate analysis by estimating the amounts of genetic exchange between several North

American monarch populations that are varying distances from each other and that are either connected by land or separated by sea.

Between 2007-2012, we obtained 144 monarchs from Mexico overwintering sites, Bermuda, Belize, Costa Rica, Puerto Rico, and Ecuador (see Fig. 3 for locations and sample sizes), and compared them to the 200 butterflies collected from St Marks, FL and coastal California for the study by Lyons et al. (2012). We refer to the original 200 butterflies as the United States sample from here on. Monarchs collected from Ecuador represent the subspecies *Danaus plexippus megalippe*, while all other butterflies represent *Danaus plexippus plexippus*. Like Lyons et al. (2012), we used 11 polymorphic microsatellite markers to determine whether the butterflies are genetically differentiated by location, or whether extensive gene flow occurs. DNA extractions and PCR protocols followed those published by Lyons et al. (2012). To investigate population structure, we used the same genetic clustering analysis and software settings in STRUCTURE version 2.3.2.1 (Pritchard *et al.* 2000) as described in Lyons et al. (2012). The parameters were sensitive enough to detect subtle or newly formed population structure, and thus to determine the most likely number of genetic populations present.

Figure 3.3. Location and hypothesized historical dispersal of sampled monarch populations. Letters represent the sample sites as follows: A-Mexico (n=27), B-United States (n=200), C- Puerto Rico (n=29), D-Bermuda (n=13), E-Belize (n=31), F-Costa Rica (n=30), G-Ecuador (n=14).

Our analysis showed that monarchs from the United States, Mexico, Belize, and Costa Rica are genetically indistinguishable (*i.e.,* they are all derived from a single genetic population, indicated in dark grey in Fig. 4), suggesting significant genetic mixing between the monarchs from these locations. In addition, although butterflies from Bermuda and Puerto Rico are genetically similar to mainland North-American monarchs, we found moderate genetic differentiation between island and mainland populations (island butterflies carry high proportions of alleles from both the dark and light grey populations in Fig. 4). Furthermore, monarchs from Ecuador form a very distinct population (indicated in black in Fig. 4), different from all other mainland and island groups, a finding that is not surprising given that these monarchs have been previously characterized as a different subspecies.

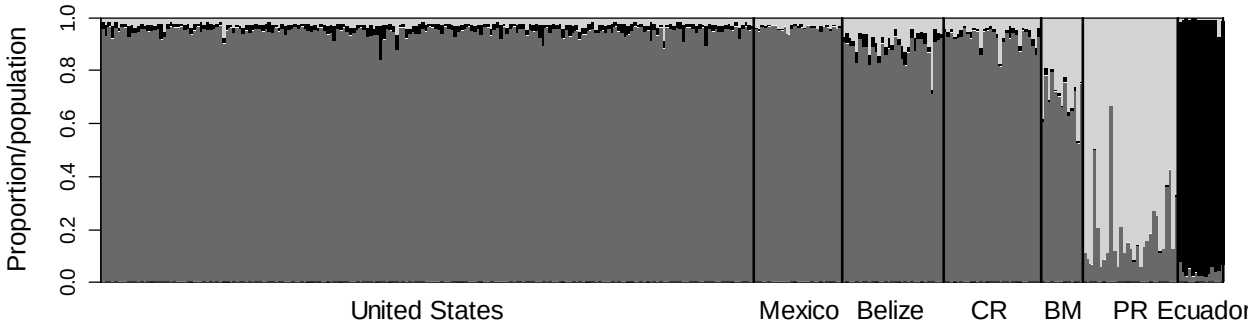

Figure 3.4. Inferred genetic proportion of individual butterflies to each of seven populations. Genetic clustering analysis suggests that there are three genetic populations: North America (United States, Mexico, Belize, Costa Rica (CR); indicated in dark grey), island populations including Bermuda (BM) and Puerto Rico (PR; light grey), and Ecuador (EC; black). Individual monarchs are indicated by vertical bars. Genetic asignments were determined on the basis of 11 microsatellite loci using the software STRUCTURE (Pritchard *et al.* 2000).

To confirm our STRUCTURE results, we used F_{ST} and R_{ST} statistics (Holsinger & Weir 2009) to measure genetic population sub-structuring among the six *D. plexippus plexippus* populations. These statistics are frequently used to measure genetic differentiation, with levels of 0 indicating that individuals belong to the same panmictic population, and values higher than 0 indicating genetic differentiation. Typically, population geneticists consider a value of 0-0.05 to indicate little differentiation; 0.05-0.15 as moderate; 0.15-0.25 as great and >0.25 as very great; and permutation tests are often used to determine if a value is considered significantly different than 0. R_{ST} was developed as a more suitable statistic for microsatellite markers, based on its dependence on a stepwise mutation model (Slatkin 1995) instead of the infinite alleles model that underlies F_{ST} statistics (Balloux & Lugon-Moulin 2002). However, because neither of these mutation models perfectly reflect natural mutation rates of microsatellites, studies on microsatellites often report both measures (Balloux & Lugon-Moulin 2002), and we followed this practice. We calculated pairwise F_{ST} and R_{ST} values between the six *D. plexippus plexippus* populations using the software Genepop version 4.1.0 (Rousset 2008). Permutation tests (using 10,000 permutations), as implemented in the 'Population comparisons' calculations in Arlequin 3.5.1.2 (Excoffier & Lischer 2010) were used to determine whether pairwise F_{ST} and R_{ST} values differed significantly from 0 (*i.e.* to determine if populations are differentiated from each other).

Our analyses of F_{ST} and R_{ST} confirm the results from our STRUCTURE analysis; there was no significant genetic differentiation between monarchs obtained from the United States, Mexico, Belize, and Costa Rica, but Bermuda and Puerto Rico were differentiated from each other and from the mainland populations (Table 23.1). Thus, populations separated by land masses did not show genetic differentiation, whereas populations separated by water bodies did. This suggests that monarchs are able to travel across land freely enough for gene flow to occur, even across large geographic distances. However, large expanses of open water appear to limit the amount of genetic mixing, resulting in population differentiation of island monarchs.

Table 3.1 Pairwise R_{ST} and F_{ST} values between monarchs from 7 geographic regions.

	United States	Mexico	Belize	Costa Rica	Bermuda	Puerto Rico
Mexico	Rst: -0.01046					
	Fst: -0.00044					
Belize	Rst: -0.02144	Rst: -0.01301				
	Fst: -0.00062	Fst: 0.00044				
Costa Rica	Rst: -0.00964	Rst: 0.00658	Rst: 0.00074			
	Fst: -0.00153	Fst: 0.00565	Fst: -0.00042			
Bermuda	Rst: 0.00458	Rst: 0.01936	Rst: 0.04018*	Rst: 0.00521		
	Fst: 0.07429*	Fst: 0.06768*	Fst: 0.08379*	Fst: 0.08776*		
Puerto Rico	Rst: 0.01794*	Rst: 0.09878*	Rst: 0.08454*	Rst: 0.03750*	Rst: 0.06209*	
	Fst: 0.08264*	Fst: 0.11018*	Fst: 0.08456*	Fst: 0.09715*	Fst: 0.16965*	
Ecuador	Rst: 0.13180*	Rst: 0.25540*	Rst: 0.20782*	Rst: 0.20677*	Rst: 0.28394*	Rst: 0.27552*
	Fst: 0.18721*	Fst: 0.24336*	Fst: 0.21644*	Fst: 0.20134*	Fst: 0.30676*	Fst: 0.34122*

Values near zero indicate that populations are not genetically distinct, whereas higher values indicate differentiation. Asterisks and shading denote values that are significantly different from zero (indicating that those populations are significantly differentiated from each other).

Apart from generating insights into monarch dispersal ability, our results might also change the way scientists view monarch migration. The migration of monarchs from Canada and the United States to Mexico and back again has generally been viewed as a directed two-way migration. A major reason for this view is that monarch migration has been best studied by United States and Canadian citizens and scientists who observe monarchs flying south/southwest towards Mexico in the fall and returning in the spring. However, an alternative scenario is that North-American monarchs aggregate in Mexico during the winter, and then disperse in all directions, not just northwards to the United States and Canada, during the spring. Such undirected dispersal has been suggested by Wenner and Harris (1993) with regards to spring migrants in California, although the number of monarchs returning to the south central U.S. in the spring is evidence that at least a large number of monarchs from Mexico do return to the breeding range of the eastern migratory population (Miller *et al.* 2012).

Importantly, either of these scenarios would result in the genetic mixing of monarch populations north and south of the Mexican overwintering sites that our analyses showed. Further study into dispersal from the overwintering sites is required to understand the full implications of these results as well as their relevance to understanding the navigational and homing mechanisms involved in monarch migration.

A new view of monarch migration and evolution in the genomic era

Past studies of gene flow within and among monarch populations were aided by genetic markers that are selectively neutral, including allozymes, mitochondrial DNA markers, and microsatellites. But different genetic approaches are required to examine evolutionary changes in monarchs that might have arisen from selection pressures, including those driven by long-distance migration. For example, even though selectively neutral markers have suggested a lack of genetic differentiation between monarchs with different migration destinations in eastern and western North America (Lyons *et al.* 2012), monarchs from these different locations could still be differentiated at loci that are under differential selection in their respective geographic areas. Thus, it is possible that locally adapted variants of specific genes are selected for during the breeding and migration seasons despite a high influx of neutral genes via gene flow. Likewise, despite the lack of differentiation of neutral markers across migratory and non-migratory monarch populations in the New World, these same populations could be differentiated at loci that influence traits such as flight ability or metabolism, being favored in populations that migrate annually, but not in populations that do not migrate.

Work on phenotypic variation across wild monarch populations already supports the idea that populations experiencing gene flow can continue to diverge at traits under selection. For example, a recent study focused on wing morphology across multiple wild monarch populations and showed that forewings were larger and more angular in shape (higher aspect ratio) in North American migratory monarchs relative to non-migratory monarchs in Hawaii,

South Florida, and Costa Rica (Altizer & Davis 2010); this work further showed that wing traits were heritable and among-population differences were maintained even when monarchs were reared in common garden experiments. A similar study on monarchs in Cuba showed that resident individuals (identified based on stable isotopes and cardenolide fingerprints) had shorter wings than migrants from eastern North America (Dockx 2007, 2012).

Molecular genetics approaches allow us to examine variation across multiple traits, including those that might be difficult to measure phenotypically, such as flight metabolism or navigational systems. One way to identify such "migration genes" is to carry out a candidate gene approach, in which specific genes are sequenced in migratory and non-migratory monarchs. One likely candidate is the phosphoglucose isomerase (*Pgi*) gene; *Pgi* is a central enzyme in glycolysis and affects the flight and dispersal ability of monarchs and other butterflies (Hughes & Zalucki 1993; Niitepõld *et al.* 2009). It is likely that butterflies that migrate seasonally will carry different alleles of this gene than non-migratory butterflies. Indeed, preliminary analyses have found that migratory and non-migratory monarch populations from North and South America are nearly fixed for different *Pgi* haplotypes (NLC & MRK, unpublished data).

The downside of a candidate gene approach is that most candidate sites that could be responsible for differences in migration are currently unknown. With this in mind, a genome-wide analysis of genetic differences between migratory and non-migratory monarchs might offer more promise, and the recently published sequence of the full monarch genome (Zhan *et al.* 2011) will make this approach feasible. In particular, by sequencing the genomes of monarchs from migratory and non-migratory populations, as well as related *Danaus* species that do not migrate, we can investigate associations between particular genes and alleles and the extent to which populations undergo long-distance migration. Such sequencing, currently underway, will provide a powerful way to uncover novel migration genes, and could confirm whether migratory monarchs have variations in genes that are involved in flight ability, such as *Pgi*.

Recent work in developing an expressed sequence tag library and the sequencing of the monarch genome has opened the door to almost unlimited possibilities in monarch genetics (Zhan *et al.* 2011; Zhu *et al.* 2008). This work has identified multiple genes that may be involved in circadian rhythm and other aspects involved in migration (Zhu *et al.* 2008). For example, by comparing expressed sequences in the brains of summer breeding versus fall migratory monarchs, Zhu and colleagues showed that genes like *turtle*, which affects fruit fly locomotion, are up-regulated during migration and could be involved in migratory locomotor behavior. Another gene, *rosy*, related to increased longevity in fruit flies, was also found to be up-regulated in migratory monarch adults, which in North America live for up to 8 months, compared to only about a month for their summer counterparts. Thus, along with 70 other genes identified in the genome, *rosy* could account for the increased lifespan of migratory monarchs (Zhu *et al.* 2008). Genes relating to eye development and neural processing have also been identified, which may provide insights into the monarchs' possible use of a sun compass (Zhan *et al.* 2011). These genomic tools have just begun to allow researchers to discover genes related to the monarch's fascinating migration. They will surely prove themselves invaluable in the future understanding of monarch biology in general.

Outlook

The field of monarch genetics has rapidly expanded and this growth will continue with improving technologies. Recent years have seen the transition from the use of traditional population genetics approaches to the use of genomics and the investigation of genome-wide gene expression, and the recently sequenced monarch genome has much to offer in determining the underlying mechanisms of monarch migration. Efforts to resequence, where part of an individual's genome is sequenced and compared to the standard genome to detect differences, will help identify genes that underlie migration, and transcriptomics will continue to offer insights in the role of differential gene expression in monarch migration. Moreover, the

monarch genome can now be mined for thousands of single nucleotide polymorphisms, which will help scientists more accurately quantify the evolutionary history of monarch populations worldwide. The enticing aspect of genetic studies is that they can show us what is currently happening with monarchs, and can also give insights into their past. Differing allele frequencies and genetic diversity levels provide glimpses into monarchs' historical colonization pathways and evolutionary history. We can use the monarch genome to better understand how and when monarchs colonized the world, and how the exposure of monarchs to novel habitats and environments has affected their genetic adaptation.

Although most previous genetic studies on monarchs have focused on some aspect of migration, scientists have now reached an era where genetics can be used to investigate a wide variety of monarch features and behaviors. For example, recent genomic studies have identified the genetics underlying the chemical defense mechanism of monarchs, which includes a variant of the sodium/potassium pump that makes them more resistant to the toxic effects of milkweed chemicals (Dobler *et al.* 2012; Zhan *et al.* 2011; Zhen *et al.* 2012; Zhu *et al.* 2008). We expect that the ongoing development of genetic tools will help scientists better understand the evolution of monarch metamorphosis, host plant specialization, warning coloration, and resistance to disease, to name just a few topics. Undoubtedly, future work in genetics, coupled with traditional studies and observation, will lead to exciting new breakthroughs in the field of monarch biology.

Chapter 4

Serial founder effects and genetic differentiation during worldwide range expansion of monarch butterflies

Reprinted material from: A.A. Pierce, M.P. Zalucki, M. Bangura, M. Udawatta, M. Kronforst, S. Altizer, J. Fernandez Haeger, J.C. de Roode. *Proc. R. Soc. B.* 281 (Dec, 2014). Used by permission of the publisher, Elsevier.

Introduction

Range expansions often result in decreased genetic diversity and increased between-population differentiation with increasing geographic distance from the source population (Eckert *et al.* 2008; Peter & Slatkin 2013; Schulte *et al.* 2013). Theoretical and empirical studies show that genetic drift and serial founder events are important mechanisms producing such patterns (Austerlitz *et al.* 1997; Francois *et al.* 2008; Slatkin & Excoffier 2012; Taberlet *et al.* 1998). In *Homo sapiens*, migration out of Africa resulted in serial founder effects, including a decrease in heterozygosity with increasing distance from the African origin (Deshpande *et al.* 2009; Li *et al.* 2008; Ramachandran *et al.* 2005), as well as a loss of genetic diversity in their accompanying human malaria parasite, *Plasmodium falciparum* (Tanabe *et al.* 2010). Importantly, genetic drift in edge populations can override selection during range expansion (Muller *et al.* 2014), sometimes causing fixation of harmful alleles and extinction of beneficial alleles (Hallatschek *et al.* 2007).

Serial founder events can result in gene surfing: this spatial analog of genetic drift occurs when neutral or even deleterious alleles reach higher than expected frequencies along the front of an expansion (Edmonds *et al.* 2004; Flaxman 2013; Hallatschek & Nelson 2008; Klopfstein *et al.* 2006; Slatkin & Excoffier 2012). Gene surfing has been observed in multiple species, many with relatively low dispersal rates (Gracia *et al.* 2013; White *et al.* 2013). While these studies demonstrated the effects of genetic drift in shaping population genetics associated with range expansions, few studies have been carried out on organisms that undertake frequent and long-

distance migrations. Simulations predicted that while serial founder events can cause genetic diversity loss and greater differentiation, these effects can be reduced by high dispersal (Bialozyt *et al.* 2006; Fayard *et al.* 2009; Ray & Excoffier 2010). Indeed, the disappearance or reduction of an isolation-by-distance pattern has been shown in a handful of species cable of long-range dispersal (Berthouly-Salazar *et al.* 2013; Szovenyi *et al.* 2012).

Here we study how the range expansion of a migratory species has shaped its population genetics. We focus on the monarch butterfly, *Danaus plexippus*, which has expanded its ancestral range to colonize locations around the globe (Ackery & Vane-Wright 1984). Monarchs are famous for their annual migration from the eastern parts of Canada and the United States to overwintering sites in central Mexico (Urquhart & Urquhart 1978); however, monarchs also occur in locations around the world, ranging from the New World tropics to the Pacific islands (Ackery & Vane-Wright 1984; Zalucki & Clarke 2004) and southern Europe (Fernandez Haeger & Jordano Barbudo 2009). Most monarch populations outside of North America are, in fact, non-migratory (Altizer *et al.* 2000; James 1993), or locally travel only short distances in the form of modest range shifts (Dingle *et al.* 1999). Historical records and recent genomic analyses suggest that monarchs colonized other locations throughout the world from North American origins (Zhan *et al.* 2014).

Records of monarchs outside of the New World are limited to the last 200 years, after suitable milkweed host plants (mostly in the genus *Asclepias*) were introduced (Zalucki & Clarke 2004); however, recent genome-wide analyses suggest that these colonizations may have occurred earlier (Zhan *et al.* 2014). Monarchs likely colonized locations around the world through serial founder effects (stepwise dispersal) or multiple independent colonization events (Shephard *et al.* 2002; Zalucki & Clarke 2004), and were aided in reaching new locations through human transportation of both milkweeds and monarchs (Vane-Wright 1993), and extreme weather events (Clarke & Zalucki 2004). For example, monarchs were recorded in Australia in 1870, and were most likely carried there on cyclonic winds from a source population

in New Caledonia (Clarke & Zalucki 2004). Regarding eastward dispersal across the Atlantic Ocean, monarchs have been spotted along the southern Iberian peninsula and northern Africa, where their population sizes vary seasonally (Fernandez Haeger *et al.* 2011), pointing to a potential role for multiple introductions in facilitating population persistence (Fernandez Haeger *et al.* 2013).

We sampled monarchs from 18 locations worldwide (Fig. 1) and used microsatellite analyses to investigate patterns of isolation by distance and genetic diversity. We found that monarch populations connected with the North-American source population by land maintained high levels of gene flow. We also detected strong genetic differentiation associated with range expansion out of North America. Across the Pacific Ocean, we found evidence of gene surfing and colonization through a pattern of serial stepwise dispersal. Conversely, across the Atlantic, monarchs showed evidence for multiple colonization events and dramatic founder effects within small, isolated populations.

Figure 4.1. Map showing sampling locations of monarchs used for this study. Numbers represent sample sizes for each location. The inset shows a cluster of North American migratory monarchs at an overwintering site in Central Mexico (left) and a female monarch laying eggs on the tropical milkweed *Asclepias curassavica* (right). (Photos by J.C. de Roode).

Methods

***Monarch field collections*:** We obtained 746 monarchs from 18 locations around the world between 2007-2011(Fig. 1; Table 1). From North American migratory monarchs we included samples from one fall migratory stopover site in north Florida (St Marks: 100 monarchs), two California overwintering sites (Pismo Beach and Santa Barbara: 100 monarchs) and two Mexico overwintering sites (Cerro Pelon and Sierra Chincua: 27 monarchs). We refer to the monarchs sampled from the eastern and western United States as the USA sample from hereon because despite overwintering in different locations, eastern and western North American monarchs are genetically indistinguishable on the basis of microsatellites (Lyons *et al.* 2012). We also included samples from non-migratory New World monarchs inhabiting South Florida, southern North America (Costa Rica and Belize), Caribbean and Atlantic Islands (Puerto Rico and Bermuda), and South America (Ecuador and Aruba). Additionally, we included non-migratory monarchs from islands across the Pacific Ocean (Hawaii, Samoa, Fiji, New Caledonia, New Zealand and Australia) and monarchs from the Iberian Peninsula and northern Africa (Spain, Portugal, and Morocco).

Table 4.1. Sampling sites and numbers of monarch butterflies

Region	Sampling location	Number of monarchs	Year collected
North America	United States	200	2009, 2010
	Mexico	27	2008
	Belize	31	2011
	Costa Rica	30	2009, 2010
	South Florida	36	2008, 2009
	Bermuda	13	2012
	Puerto Rico	29	2010
South America	Ecuador	14	2008
	Aruba	29	2012
Hawaiian Islands	Hawaii	114	2007, 2009, 2010
Pacific Islands	Samoa	32	2006, 2007
	Fiji	10	2009, 2010
	New Caledonia	40	1991, 2008, 2010

	New Zealand	22	2007, 2011
	Australia	27	2009
Iberian Peninsula	Spain	31	2012
	Portugal	32	2012
northern Africa	Morocco	29	2012

***Genetic work*:** We used 16 microsatellite markers [48] to genotype each of our samples. For PCR, genomic DNA was extracted from a 0.5 mm section of butterfly thorax (females) or abdomen (males) using the Ultraclean DNA Isolation Kit from Mo-Bio (Carlsbad, CA, USA) and quantified using a Nanodrop 2000. We extracted DNA from females from the thorax rather than the abdomen to avoid possible contamination from male sperm. We followed PCR protocols as described [ref 48]. Amplified DNA was genotyped on an ABI 3100 genetic analyzer (Perkin Elmer, Applied Biosystems, Foster City, CA, USA) at the Emory Integrated Genomics Core (EIGC; Atlanta, GA, USA) and alleles were scored using Genemarker v. 4.0 (SoftGenetics LLC, State College, PA, USA).

Microsatellite analyses: We used the software Arlequin 3.5.1.2 (Excoffier & Lischer 2010) to calculate observed and expected heterozygosity at each locus in each population and calculated deviations from Hardy-Weinberg equilibrium using sequential Bonferroni correction at $\alpha = 0.05$ (Rice 1989). For subsequent analyses, we used 11 microsatellite loci that were in Hardy-Weinberg equilibrium in at least 11 out of 18 sampling locations (Table S1).

Worldwide population genetic analyses: To investigate worldwide population structure, we used the software STRUCTURE version 2.3.3 (Pritchard *et al.* 2000) and implemented 100,000 burn-ins and 200,000 MCMC runs after the burn-in. We used an admixture model with uncorrelated allele frequencies to avoid the risk of overestimating the number of populations and used the LOCPRIOR model to provide the software with location information for each butterfly to ensure the detection of subtle population structure. We started simulations with K=18, to reflect the 18 sampling locations, and then ran simulations for K values of 18 through 1. For each K we ran ten simulations to check for consistency between runs,

and used the log likelihood (Pritchard *et al.* 2000) and delta *K* method (Evanno *et al.* 2005) to determine the most likely number of genetic populations.

We calculated Nei's standard genetic distance (D_{ST}) (Nei 1973) using the GenAlEx Excel plugin (Peakall & Smouse 2012). With this measurement, we created a distance matrix and used it to build a neighbor-joining distance tree in MEGA5 (Tamura *et al.* 2011) to visualize the genetic relationships between monarchs from our sampling sites. With the same distance matrix, we performed a Principal Coordinate Analysis using the cmdscale function in R (version 3.0.1).

To confirm our results, we used F_{ST} and R_{ST} statistics (Holsinger & Weir 2009; Slatkin 1995) to measure genetic differentiation between monarch populations. Permutation tests (using 10,000 permutations), implemented in Arlequin 3.5.1.2 (Excoffier & Lischer 2010) were used to determine whether pairwise F_{ST} and R_{ST} values were significantly different from 0. To account for the potential occurrence of null alleles, we recalculated F_{ST} and R_{ST} values using corrected allele frequencies as determined by the software MICRO-CHECKER, version 2.2.3 (Van Oosterhout *et al.* 2004). Overall, statistics based on corrected and uncorrected allele frequencies were highly similar (Table S2 and Table S3), and resulted in the same conclusions.

***Population genetic analyses of monarchs across the Pacific and Atlantic*:** We used STRUCTURE (settings and determination of K as above) to investigate population structure at a smaller scale among locations across the Pacific versus the Atlantic. For six locations across the Pacific, we began with K=6 and ran simulations for K values 6 through 1. For eight locations across the Atlantic, we started simulations with K=8 and ran simulations for K values of 8 through 1. To analyze isolation by distance, we analyzed the correlation between geographic distance and genetic differentiation (F_{ST} and R_{ST}) for locations across the Pacific and across the Atlantic using Mantel tests implemented in the vegan library version 2.0-0 (Oksanen *et al.* 2011) in the statistical package R (version 3.0.1).

To determine likely dispersal routes, we resampled monarchs from each location with replacement using Poptools (Hood 2010) to standardize sample size across sites (n=22) prior to comparison of relative levels of genetic diversity (using the value 1-Qinter, the inter-individual diversity within populations), which were measured using Genepop version 4.1.0 (Rousset 2008). We also calculated allelic richness, the number of alleles per locus, using ADZE-1.0 (Szpiech *et al.* 2008), which utilizes a rarefaction approach to account for differences in sample size. Locations with fewer than 20 samples were excluded from these analyses. We then grouped locations based on proximity to either the Pacific or Atlantic Ocean to investigate the possibility of allelic clinal patterns.

We calculated levels of gene flow, measured as the number of migrants reaching one population from another per generation, using likelihood ratio tests implemented in the coalescent-based software package MIGRATE-N 3.2 (Beerli 2009). We used a Brownian motion approximation to the ladder model and Bayesian inference with multiple heating chains to jointly estimate parameters (Beerli 2006; Beerli & Felsenstein 2001). We arbitrarily limited four locations for analysis of gene flow across the Pacific and four locations for determining gene flow across the Atlantic to maintain program accuracy and efficiency. Each analysis was run independently with five replicates.

Results

Worldwide patterns: Monarchs showed substantial genetic differentiation among the sampled locations (Fig. 2). STRUCTURE analyses supported a total of seven worldwide monarch populations composed of (i) North America (including USA, Mexico, Central America, and neighboring Caribbean/Atlantic islands), (ii) South America (Ecuador), (iii) Aruba, (iv) Spain, (v) Portugal/Morocco, (vi) the Hawaiian Islands, and (vii) a series of Pacific Islands (including Australia and New Zealand; Fig. 2).

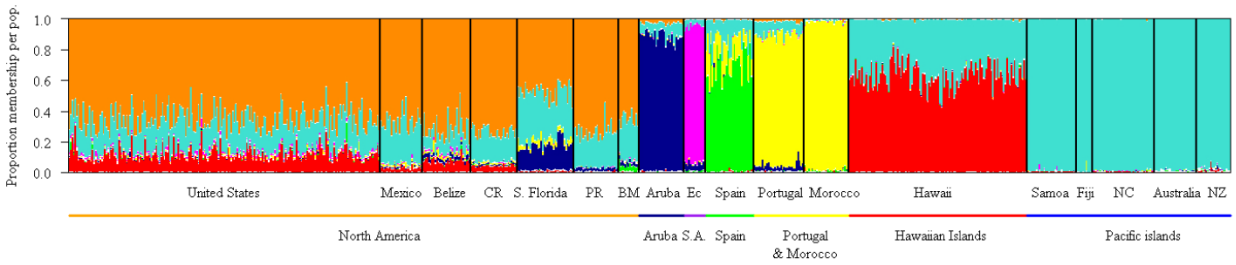

Figure 4.2. Worldwide structure plot showing that K (number of distinct populations) = 7. Microsatellite loci show that monarchs across the globe form seven genetic populations (orange bar=North America, dark blue bar= Aruba, purple bar= South America (S.A.), green bar= Spain, yellow bar= Portugal & Morocco, red bar= Hawaiian Islands, blue bar= Pacific Islands). The North American population is composed of the United States, Mexico, Belize, and Costa Rica (CR) as well as slightly more genetically differentiated locations such as South Florida (S. Florida), Puerto Rico (PR), and Bermuda (BM). South America is composed of Ecuador (Ec) and the Pacific Islands are composed of Samoa, Fiji, New Caledonia (NC), Australia, and New Zealand (NZ). Individual monarchs are indicated by vertical bars and color denotes population membership

We inferred similar relationships from a distance tree based on Nei's standard genetic distance (Fig. 3) with the locations grouping intuitively based on geographic proximity. In addition, the relatively long branch lengths between USA/Mexico monarchs and those from two neighbouring locations (Puerto Rico and South Florida/Bermuda) support the intermediate levels of differentiation identified by STRUCTURE. The nested structure of the distance tree (Fig. 3) also offers insight into the colonization route. For example, the branch capturing the Pacific Islands shows Hawaii having intermediate differentiation, with other Pacific sites showing increased differentiation from the USA; this pattern supports the serial stepwise dispersal hypothesis. Similarly, the clustered nature of the Iberian Peninsula and northern Africa locations suggest common colonization events. Similar patterns were observed using Principal Coordinate Analysis (Fig. S1). These relationships are further supported by F_{ST} and R_{ST} values, which are much lower between populations within the same geographic group and in closer geographic proximity than between populations in different geographic groups (Tables S2, S3).

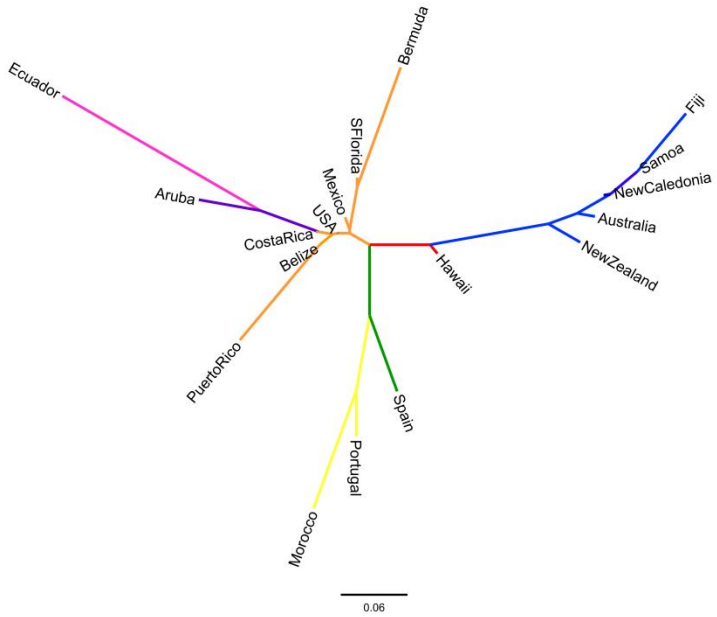

Figure 4.3. Distance tree based on Nei's standard genetic distance (D_{ST}). Sampling locations group intuitively based on geographic proximity and indicate groupings similar to the STRUCTURE plot in Fig 1. Branches are colored to demonstrate group membership (orange=North America, dark blue= Aruba, purple= South America (S.A.), green = Spain, yellow= Portugal & Morocco, red = Hawaiian Islands, blue = Pacific Islands).

We found evidence for a panmictic population spanning the continent of North America, composed of the USA, Mexico, Belize, and Costa Rica. South Florida, Puerto Rico, and Bermuda were clearly differentiated from these North American mainland populations, as indicated by F_{ST} and R_{ST} values (Tables S2, S3), and STRUCTURE analysis (Fig. 2).

Monarchs across the Pacific: In our worldwide analyses, when compared with other populations, monarchs from the Pacific Islands appear very homogeneous. To examine more subtle patterns of variation, we ran STRUCTURE for the Hawaiian and Pacific Islands only (Fig. 4E), and were able to detect more population differentiation among the Pacific Island locations. To further explore this pattern, we examined allele frequencies for each of the 11 loci and found a distinct clinal pattern (Fig. S2), consistent with stepwise dispersal and gene surfing. Allelic richness was highest in the USA, with subsequent loss of some alleles and enrichment of others

with increasing distance westward towards Australia and New Zealand. This pattern was observed across all 11 loci (Fig. S2), which is highly indicative of a stepwise dispersal across the Pacific, and the presence of gene surfing in these colonization events.

In addition to this clinal pattern, we found a decrease in genetic diversity and allelic richness as monarchs moved farther west from North America (Fig. 4A, 4B), further supporting stepwise serial dispersal. Conversely, the observation of relatively high frequencies of rare alleles in more distant locations is also consistent with the possibility that some sites were founded by independent colonization events. Additionally, there was only a weak pattern of isolation by distance (Fig. 4C; $p=0.06$), as has been seen in other highly mobile species, such as the European Starling in South Africa (Berthouly-Salazar *et al.* 2013), which might arise from ongoing gene flow or intermittent dispersal events among colonized locations. To determine levels of gene flow, we ran Migrate-N using samples across the Pacific. Estimated number of migrants per generation between locations ranged from only two individuals from New Zealand to Australia to 15 individuals from Hawaii to the USA (Fig. 4D). This confirms the presence of ongoing gene flow even among distantly separated populations.

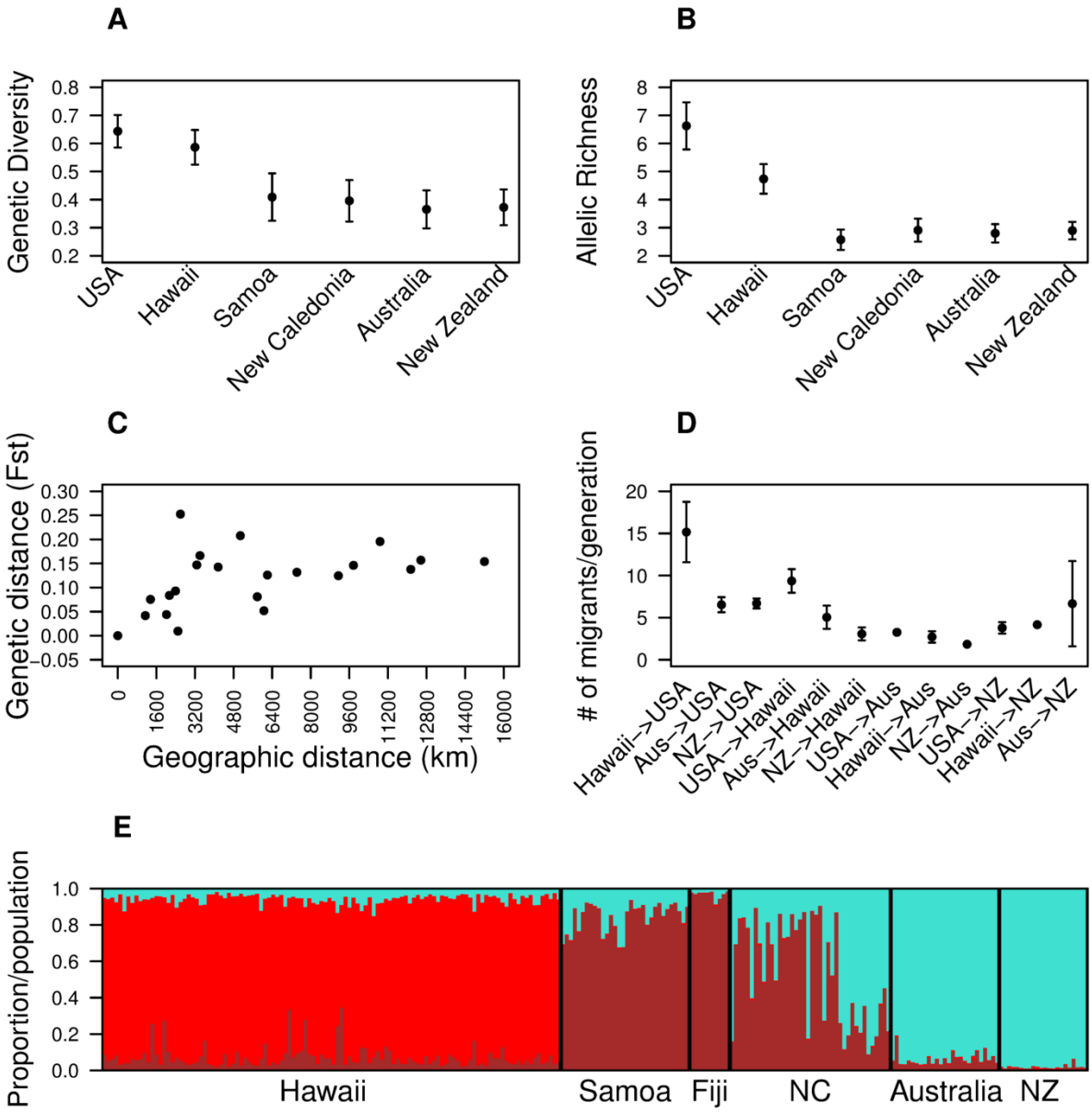

Figure 4.4. Genetic trends across the Pacific. A) Genetic diversity (using the value 1-Qinter, the inter-individual diversity within populations) was highest in the USA and tended to decrease with increasing distance from the source population. B) Allelic richness was higher in the United States and significantly decreased across the Pacific. C) No pattern of isolation by distance was found (r=0.3833, p=0.062). D) Gene flow ranged between 2 and 15 individuals per generation and was highest from Hawaii to the United States. E) Pacific STRUCTURE plot indicating K=3. Increased population differentiation and a clinal pattern appear when examined at a finer spatial scale.

***Monarchs across the Atlantic*:** We ran STRUCTURE using samples from the USA and locations across the Atlantic Ocean, including South Florida, Puerto Rico, Bermuda, Spain, Portugal, and Morocco. Results supported the same structure shown in Fig. 2, which is unsurprising as we were already able to detect ample population structure. We observed a decrease in allelic richness and genetic diversity among monarchs sampled in these Old World locations relative to the USA, although the pattern was not as strong as that observed across the Pacific (Fig. 5A, 5B). To investigate potential allelic clinal patterns, we examined allele frequencies for the USA, Spain, Portugal and Morocco. Many of the alleles found in the Iberian Peninsula and northern Africa represent a subset of alleles found in the USA (Fig. S3), indicative of founder effects.

We found a weak pattern of isolation by distance (Fig. 5C, $p=0.048$), with increasing distance from the USA associated with increasing genetic differentiation. To determine gene flow across the Atlantic and among Old World locations, we ran Migrate-N and found relatively low levels of gene flow in comparison to the sites across the Pacific (Fig. 5D). This suggests that while there may have been multiple colonization events from North America, they occur infrequently or involve few individuals.

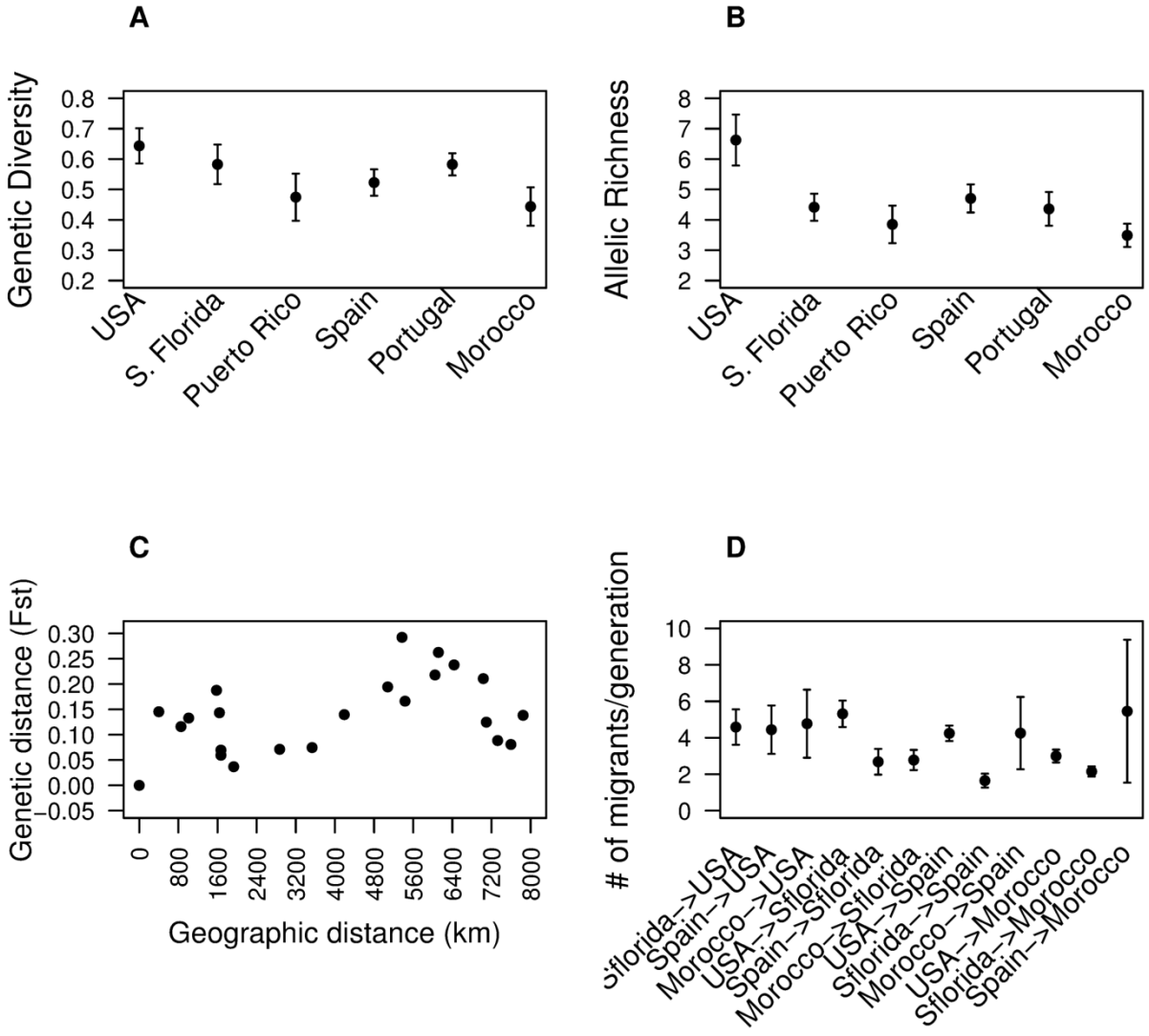

Figure 4.5. Genetic trends across the Atlantic. A) Genetic diversity (using the value 1-Qinter, the inter-individual diversity within populations) was highest in the United States and tended to decrease with increasing distance from the source population. B) Allelic richness was higher in the United States and significantly decreased across the Atlantic. C) A weak pattern of isolation by distance was found (r=0.3181, p=0.048), with increasing geographic distance correlating with increasing genetic distance. D) Gene flow ranged between 2 and 6 individuals per generation.

Discussion

A strong signal of worldwide genetic population structure was detected in monarch butterflies, despite their high propensity for flight and their long-distance migration. With

increasing distance from the likely source population in North America, we found a trend of decreasing genetic diversity and clines in allelic richness, with some alleles disappearing and others becoming enriched. It thus appears that genetic drift has been the main driver in determining allele frequencies as monarchs colonized new areas. Importantly, spatial expansions can result in rare alleles reaching high frequencies through gene surfing, a process associated with genetic drift (Slatkin & Excoffier 2012), and this process has likely been a strong force shaping allele frequencies of newly colonized monarch populations across both the Atlantic and the Pacific Oceans.

Pacific versus Atlantic

Our results support a serial founder effect in monarchs from locations across the Pacific. The robust clinal pattern in allele frequencies (Fig. S2) and decreased allelic richness and genetic diversity (Fig. 4) strongly suggest that stepwise dispersal was the primary mechanism of colonization (see also ref (Shephard *et al.* 2002)), as has been seen in other range expansions like that of the spur-thighed tortoise in Spain (Gracia *et al.* 2013). Conversely, across the Atlantic, we found relatively higher genetic diversity as well as evidence for multiple colonization events and dramatic founder effects with rare alleles in the donor population found at relatively high frequencies. This pattern is more similar to that found in the speckled wood butterfly, in which range expansion was characterized by multiple colonization events resulting from long distance dispersal. There, it was found that while dispersal events resulted in reduced genetic diversity, this effect was lessened by subsequent gene flow (Vandewoestijne & Van Dyck 2010). Indeed, in monarchs, historical records provide evidence of infrequent “migrants” or vagrants that strayed across the Atlantic on winds from North America with only a few individuals arriving in each dispersal event. In the Iberian Peninsula and northern Africa, monarchs are faced with hot and dry summers, resulting in host plants being scattered and restricted to wet soils (river beds, ponds, etc.) (Fernandez Haeger & Jordano Barbudo 2009).

Although this fragmented distribution could contribute to reduced gene flow between different patches, the most likely explanation for rare alleles is the high fluctuations in butterfly density in those patches. It is common for Iberian monarchs to suffer periodic bottlenecks as a consequence of excessive larval density or the destruction of plants by flooding rivers (Fernandez Haeger *et al.* 2013).

Like many introduced or invasive species, monarch colonization across both the Pacific and the Atlantic was made possible by anthropogenic interactions, specifically human-driven colonization by milkweeds (which are non-native in most locations where monarchs are now found). Additionally, colonization was likely affected by between-location distances and wind patterns (Zalucki & Clarke 2004). The distances between locations inhabited by monarchs in the Iberian Peninsula and northern Africa are much shorter (and often connected by land) relative to the distances between oceanic islands in the Pacific. Importantly, both normal wind patterns and severe weather events such as hurricanes or cyclones have been known to transport monarchs across oceans (Clarke & Zalucki 2004), thus facilitating colonization and gene flow. Winds in both directions moving east and west away from of North America might have aided monarchs in trans-oceanic movements; additionally, the presence of the Gulf Stream and strong winds in middle latitudes from the USA to Europe may have resulted in more regular introductions across the Atlantic. Our analyses suggest that gene flow was generally higher among the Pacific locations than among those across the Atlantic, suggesting that monarchs move frequently enough between Pacific islands to limit population divergence.

Implications

Our analyses show that monarch migration provides sufficient gene flow to create a panmictic population of butterflies that spans the mainland North-American continent. These results indicate that the long-distance migration of monarchs from the USA and Canada to Mexico each year likely results in genetic mixing which affects monarchs in other North

American locations. One hypothesis for how this occurs is that rather than solely re-migrating north each spring, a portion of monarchs radiate outwards from their overwintering sites in search of larval host plants (Pierce *et al.* 2015; Wenner & Harris 1993). Due to the panmictic nature of the population in North America, our results provide a new outlook on monarch butterfly migration pathways.

Results here also indicate that different migration strategies, such as the use of different overwintering sites or the presence of non-migratory behavior, can be maintained between populations despite high levels of gene flow (Lyons *et al.* 2012) as also seen in the admixed populations of migratory and sedentary groups of broad-tailed hummingbirds in North America (Malpica & Ornelas 2014). A similar lack of neutral genetic differentiation has been observed in other highly mobile, migrating species, such as hoverflies with different overwintering strategies in France (Raymond *et al.* 2013) and Blackcaps with different migratory behaviors and destinations in Europe (Mettler *et al.* 2013).

To investigate evolutionary changes in monarchs and other species that might have arisen from the selective pressures of long-distance migration, different genetic approaches are required. For example, despite a lack of genetic differentiation based on selectively neutral markers, monarchs from populations with different migratory destinations or strategies could be differentiated at loci under differential selection in their respective geographic areas (Altizer & Davis 2010; Zhan *et al.* 2014). Thus, it is possible that locally adapted alleles are selected for during the breeding and migration seasons despite a high influx of neutral genes via gene flow. To better understand the role of genetics in migration as well as the potential role of local adaptation, comparative studies investigating non-neutral alleles and traits should be addressed in the future.

With respect to other systems, our work clearly demonstrates that range expansion crucially affects levels of genetic diversity and population differentiation. Although these results are generally expected for species with low levels of dispersal, theoretical models have suggested

that high levels of dispersal could limit the effects of range expansion on reducing genetic diversity and creating genetic differentiation (Bialozyt *et al.* 2006; Fayard *et al.* 2009; Ray & Excoffier 2010). Some empirical studies have confirmed these predictions. In European Starlings inhabiting South Africa, for example, and peat mosses in the Stockholm archipelago, long-distance dispersal events have maintained genetic diversity (Berthouly-Salazar *et al.* 2013; Szovenyi *et al.* 2012) and multiple colonization events are hypothesized to have maintained relatively high genetic diversity in expanding populations of the Mediterranean damselfly (Swaegers *et al.* 2013). Additionally, when comparing bats with different dispersal abilities suffering from habitat fragmentation, highly mobile bats presented less genetic diversity loss (Meyer *et al.* 2009). In contrast to these findings, our results suggest that the genetic signature of range expansion can be maintained even in species with great capacity for dispersal. Beyond furthering our understanding of the population genetic effects of range expansion, these results are relevant for conservation biology by showing that even highly mobile species can suffer from reduced genetic diversity in widely separated geographic fragments.

In summary, our work demonstrates that genetic drift has played a key role in shaping allele frequencies in colonized monarch butterfly populations throughout the world. In addition to demonstrating the impact of migration on gene flow and its role in creating a panmictic population spanning the North American continent, this work is also relevant for conservation, as connectivity impacts population size and dynamics. Especially with global climate change becoming increasingly important and resulting in habitat and range shifts for numerous species (Thomas *et al.* 2001), a deeper understanding of the mechanisms and effects of range expansions are critical in conservation efforts and the prevention and control of invasive species. Finally, our results confirm theoretical predictions that range expansions can result in allele clines (Excoffier *et al.* 2009; Slatkin & Excoffier 2012) and gene surfing (Edmonds *et al.* 2004; Excoffier & Ray 2008; Flaxman 2013; Hallatschek & Nelson 2008; Klopfstein *et al.* 2006), and suggest that these patterns may occur even for species with great migratory ability.

Supplementary Figures/Tables

Table S4.1 Observed (H_O) and expected (H_E) heterozygosity for each population at each locus as determined by Arlequin 3.5.1.2 (Excoffier and Lischer 2010).

	Locus	168	153	320	197	208	203	141	1679	137	122	494	983	854	165	819	519
United States	H_o	**0.690**	**0.565**	**0.605**	**0.305**	**0.455**	**0.725**	**0.515**	**0.460**	0.025	**0.015**	0.305	**0.415**	**0.372**	**0.270**	0.780	0.835
	H_E	**0.619**	**0.676**	**0.884**	**0.373**	**0.771**	**0.802**	**0.662**	**0.622**	0.025	**0.035**	0.300	**0.782**	**0.589**	**0.526**	0.844	0.800
Mexico	H_o	0.629	0.703	0.778	0.444	0.519	0.519	0.481	0.519	-	-	0.111	**0.407**	0.462	**0.192**	0.852	0.667
	H_E	0.561	0.710	0.860	0.488	0.784	0.789	0.613	0.638	-	-	0.108	**0.806**	0.609	**0.438**	0.832	0.684
Belize	H_o	0.742	0.645	0.677	0.129	**0.323**	0.710	0.548	0.581	-	-	0.167	**0.367**	**0.300**	**0.200**	0.733	0.556
	H_E	0.637	0.670	0.888	0.182	**0.323**	0.710	0.548	0.668	-	-	0.160	**0.794**	**0.524**	**0.454**	0.805	0.741
Coasta Rica	H_o	0.600	0.500	**0.600**	0.267	**0.345**	0.933	0.533	0.533	**0.000**	**0.000**	0.233	**0.467**	**0.300**	**0.345**	0.733	0.800
	H_E	0.552	0.705	**0.868**	0.282	**0.683**	0.802	0.569	0.602	**0.235**	**0.131**	0.268	**0.814**	**0.581**	**0.649**	0.799	0.786
South Florida	H_o	**0.639**	**0.833**	0.556	**0.722**	0.417	0.667	**0.361**	**0.722**	-	-	0.566	**0.444**	0.306	**0.194**	0.619	0.806
	H_E	**0.644**	**0.644**	0.734	**0.520**	0.720	0.690	**0.513**	**0.640**	-	-	0.055	**0.769**	0.372	**0.440**	0.694	0.682
Puerto Rico	H_o	0.483	0.551	**0.621**	0.069	0.207	0.517	0.276	0.517	-	-	-	**0.345**	0.379	**0.069**	0.690	0.517
	H_E	0.494	0.578	**0.782**	0.131	0.324	0.586	0.319	0.602	-	-	-	**0.664**	0.422	**0.463**	0.729	0.667
Aruba	H_o	0.517	**0.214**	0.448	0.034	0.370	0.607	**0.414**	**0.571**	-	-	-	**0.621**	0.241	0.069	0.620	0.690
	H_E	0.668	**0.560**	0.661	0.034	0.673	0.674	**0.515**	**0.689**	-	-	-	**0.843**	0.313	0.346	0.713	0.766
Ecuador	H_o	0.286	**0.143**	0.214	0.286	0.077	0.214	**0.357**	0.000	-	-	0.071	0.071	0.143	**0.143**	0.571	0.500
	H_E	0.423	**0.518**	0.532	0.254	0.151	0.206	**0.614**	0.138	-	-	0.071	0.071	0.349	**0.519**	0.582	0.598
Bermuda	H_o	0.615	0.462	0.692	0.462	0.462	0.692	0.538	0.615	-	-	-	0.538	0.154	0.231	0.769	0.500
	H_E	0.624	0.723	0.806	0.369	0.806	0.766	0.710	0.732	-	-	-	0.732	0.148	0.532	0.649	0.797
Spain	H_o	0.419	0.452	**0.097**	0.258	**0.710**	0.516	0.839	0.387	-	-	0.516	**0.226**	0.542	**0.097**	0.581	0.709

	H_E	0.424	0.552	**0.531**	0.275	**0.774**	0.658	0.643	0.379	-	-	0.439	**0.762**	0.597	**0.288**	0.633	0.664
Portugal	H_O	0.219	**0.375**	0.167	0.226	0.313	0.686	**0.313**	**0.563**	-	-	0.438	**0.594**	0.469	0.250	0.656	0.625
	H_E	0.299	**0.651**	0.658	0.252	0.605	0.792	**0.425**	**0.616**	-	-	0.563	**0.776**	0.565	0.573	0.803	0.607
Morocco	H_O	0.620	**0.172**	0.000	-	0.172	0.483	**0.103**	**0.620**	-	-	0.517	**0.172**	0.655	0.069	0.620	0.655
	H_E	0.460	**0.588**	0.361	-	0.477	0.477	**0.650**	**0.608**	-	-	0.527	**0.379**	0.563	0.131	0.660	0.641
Hawaii	H_O	0.684	**0.491**	**0.614**	0.412	**0.245**	**0.558**	0.675	**0.351**	-	-	0.070	**0.286**	**0.315**	**0.228**	0.640	0.640
	H_E	0.593	**0.588**	**0.718**	0.404	**0.676**	**0.665**	0.648	**0.489**	-	-	0.069	**0.502**	**0.562**	**0.486**	0.645	0.634
Samoa	H_O	0.531	0.031	0.419	0.250	0.452	0.133	0.031	0.094	-	-	-	0.594	**0.387**	0.186	0.844	0.613
	H_E	0.448	0.031	0.601	0.268	0.508	0.325	0.031	0.091	-	-	-	0.726	**0.542**	0.376	0.788	0.707
Fiji	H_O	0.800	-	0.500	0.000	0.600	-	0.100	-	-	-	-	0.900	0.500	-	0.700	0.900
	H_E	0.611	-	0.395	0.189	0.526	-	0.100	-	-	-	-	0.574	0.395	-	0.700	0.595
New Caledonia	H_O	0.462	0.256	0.385	0.179	0.462	0.243	0.179	0.053	-	-	-	0.667	0.459	0.237	0.821	0.676
	H_E	0.533	0.295	0.569	0.165	0.499	0.434	0.213	0.052	-	-	-	0.774	0.507	0.235	0.765	0.700
Austalia	H_O	0.593	0.037	0.556	0.370	0.370	0.629	0.222	0.037	-	-	0.037	0.556	0.222	0.481	0.444	0.593
	H_E	0.498	0.037	0.531	0.307	0.509	0.498	0.208	0.037	-	-	0.037	0.711	0.391	0.498	0.532	0.623
New Zealand	H_O	0.136	**0.350**	0.136	0.409	0.409	0.619	0.227	0.227	-	-	-	0.455	0.273	0.500	0.727	0.714
	H_E	0.132	**0.594**	0.132	0.332	0.467	0.535	0.284	0.485	-	-	-	0.614	0.383	0.502	0.659	0.670

Significant differences between observed and expected heterozygosities are indicated in bold and were determined using an α of 0.05 and a sequential Bonferroni correction (Rice 1989). Loci for which at least eleven populations were in Hardy–Weinberg equilibrium were used for subsequent analyses and are indicated in grey shading. Dashes indicate monomorphic loci.

Table S4.2 Pairwise F_{ST} and R_{ST} values between 18 sampling locations, as calculated by Genepop version 4.1.0 (Rousset 2008). Shading indicates values significantly different from 0.

	USA	Mexico	Belize	Costa Rica	S. Florida	Puerto Rico	Aruba	Ecuador	Bermuda	Spain	Portugal	Morocco	Hawaii	Samoa	Fiji	New Caledonia	Australia
Mexico	R_{ST}: 0.008 F_{ST}: 0.004																
Belize	R_{ST}: -0.005 F_{ST}: 0.005	R_{ST}: -0.010 F_{ST}: 0.004															
Costa Rica	R_{ST}: 0.002 F_{ST}: 0.002	R_{ST}: 0.034 F_{ST}: 0.009	R_{ST}: 0.032 F_{ST}: 0.006														
S. Florida	R_{ST}: 0.002 F_{ST}: 0.037	R_{ST}: -0.003 F_{ST}: 0.029	R_{ST}: 0.008 F_{ST}: 0.032	R_{ST}: 0.034 F_{ST}: 0.039													
Puerto Rico	R_{ST}: 0.041 F_{ST}: 0.075	R_{ST}: 0.162 F_{ST}: 0.096	R_{ST}: 0.137 F_{ST}: 0.078	R_{ST}: 0.050 F_{ST}: 0.090	R_{ST}: 0.152 F_{ST}: 0.143												
Aruba	R_{ST}: 0.060 F_{ST}: 0.079	R_{ST}: 0.076 F_{ST}: 0.090	R_{ST}: 0.102 F_{ST}: 0.065	R_{ST}: 0.113 F_{ST}: 0.072	R_{ST}: 0.101 F_{ST}: 0.096	R_{ST}: 0.265 F_{ST}: 0.172											
Ecuador	R_{ST}: 0.07 F_{ST}: 0.164	R_{ST}: 0.192 F_{ST}: 0.201	R_{ST}: 0.170 F_{ST}: 0.184	R_{ST}: 0.144 F_{ST}: 0.177	R_{ST}: 0.200 F_{ST}: 0.232	R_{ST}: 0.186 F_{ST}: 0.333	R_{ST}: 0.233 F_{ST}: 0.213										
Bermuda	R_{ST}: 0.004 F_{ST}: 0.071	R_{ST}: 0.040 F_{ST}: 0.067	R_{ST}: 0.053 F_{ST}: 0.088	R_{ST}: 0.017 F_{ST}: 0.086	R_{ST}: 0.027 F_{ST}: 0.060	R_{ST}: 0.117 F_{ST}: 0.188	R_{ST}: 0.159 F_{ST}: 0.151	R_{ST}: 0.162 F_{ST}: 0.272									
Spain	R_{ST}: 0.140 F_{ST}: 0.081	R_{ST}: 0.139 F_{ST}: 0.094	R_{ST}: 0.117 F_{ST}: 0.090	R_{ST}: 0.167 F_{ST}: 0.099	R_{ST}: 0.146 F_{ST}: 0.125	R_{ST}: 0.241 F_{ST}: 0.238	R_{ST}: 0.212 F_{ST}: 0.189	R_{ST}: 0.171 F_{ST}: 0.221	R_{ST}: 0.089 F_{ST}: 0.166								
Portugal	R_{ST}: 0.066 F_{ST}: 0.089	R_{ST}: 0.052 Fst: 0.112	R_{ST}: 0.074 F_{ST}: 0.114	R_{ST}: 0.088 F_{ST}: 0.085	R_{ST}: 0.088 F_{ST}: 0.140	R_{ST}: 0.224 F_{ST}: 0.218	R_{ST}: 0.155 F_{ST}: 0.176	R_{ST}: 0.200 F_{ST}: 0.276	R_{ST}: 0.074 F_{ST}: 0.194	R_{ST}: 0.150 F_{ST}: 0.146							

Morocco	R_{ST}: 0.040	R_{ST}: 0.029	R_{ST}: 0.031	R_{ST}: 0.047	R_{ST}: 0.047	R_{ST}: 0.138	R_{ST}: 0.142	R_{ST}: 0.131	R_{ST}: 0.026	R_{ST}: 0.090	R_{ST}: 0.013						
	F_{ST}: 0.138	F_{ST}: 0.174	F_{ST}: 0.170	F_{ST}: 0.174	F_{ST}: 0.211	F_{ST}: 0.263	F_{ST}: 0.282	F_{ST}: 0.380	F_{ST}: 0.292	F_{ST}: 0.133	F_{ST}: 0.116						
Hawaii	R_{ST}: 0.010	R_{ST}: 0.035	R_{ST}: 0.013	R_{ST}: 0.011	R_{ST}: 0.025	R_{ST}: 0.056	R_{ST}: 0.112	R_{ST}: 0.124	R_{ST}: 0.024	R_{ST}: 0.137	R_{ST}: 0.103	R_{ST}: 0.059					
	F_{ST}: 0.052	F_{ST}: 0.050	F_{ST}: 0.059	F_{ST}: 0.066	F_{ST}: 0.067	F_{ST}: 0.150	F_{ST}: 0.142	F_{ST}: 0.251	F_{ST}: 0.140	F_{ST}: 0.136	F_{ST}: 0.139	F_{ST}: 0.192					
Samoa	R_{ST}: 0.026	R_{ST}: 0.008	R_{ST}: 0.021	R_{ST}: 0.070	R_{ST}: 0.024	R_{ST}: 0.189	R_{ST}: 0.055	R_{ST}: 0.177	R_{ST}: 0.055	R_{ST}: 0.151	R_{ST}: 0.059	R_{ST}: 0.065	R_{ST}: 0.067				
	F_{ST}: 0.146	F_{ST}: 0.166	F_{ST}: 0.226	F_{ST}: 0.203	F_{ST}: 0.205	F_{ST}: 0.266	F_{ST}: 0.323	F_{ST}: 0.427	F_{ST}: 0.233	F_{ST}: 0.267	F_{ST}: 0.252	F_{ST}: 0.299	F_{ST}: 0.143				
Fiji	R_{ST}: 0.170	R_{ST}: 0.135	R_{ST}: 0.174	R_{ST}: 0.283	R_{ST}: 0.206	R_{ST}: 0.461	R_{ST}: 0.213	R_{ST}: 0.428	R_{ST}: 0.231	R_{ST}: 0.219	R_{ST}: 0.153	R_{ST}: 0.168	R_{ST}: 0.245	R_{ST}: 0.07			
	F_{ST}: 0.196	F_{ST}: 0.217	F_{ST}: 0.269	F_{ST}: 0.256	F_{ST}: 0.264	F_{ST}: 0.314	F_{ST}: 0.366	F_{ST}: 0.497	F_{ST}: 0.279	F_{ST}: 0.296	F_{ST}: 0.287	F_{ST}: 0.313	F_{ST}: 0.208	F_{ST}: 0.042			
New Caledonia	R_{ST}: 0.002	R_{ST}: 0.039	R_{ST}: 0.044	R_{ST}: 0.016	R_{ST}: 0.043	R_{ST}: 0.082	R_{ST}: 0.119	R_{ST}: 0.141	R_{ST}: 0.037	R_{ST}: 0.180	R_{ST}: 0.085	R_{ST}: 0.058	R_{ST}: 0.022	R_{ST}: 0.036	R_{ST}: 0.232		
	F_{ST}: 0.138	F_{ST}: 0.162	F_{ST}: 0.216	F_{ST}: 0.192	F_{ST}: 0.181	F_{ST}: 0.262	F_{ST}: 0.306	F_{ST}: 0.412	F_{ST}: 0.207	F_{ST}: 0.268	F_{ST}: 0.244	F_{ST}: 0.313	F_{ST}: 0.126	F_{ST}: 0.009	F_{ST}: 0.075		
Australia	R_{ST}: 0.006	R_{ST}: 0.037	R_{ST}: 0.054	R_{ST}: 0.043	R_{ST}: 0.027	R_{ST}: 0.132	R_{ST}: 0.124	R_{ST}: 0.211	R_{ST}: 0.060	R_{ST}: 0.198	R_{ST}: 0.117	R_{ST}: 0.081	R_{ST}: 0.034	R_{ST}: 0.026	R_{ST}: 0.257	R_{ST}: 0.003	
	F_{ST}: 0.154	F_{ST}: 0.177	F_{ST}: 0.216	F_{ST}: 0.203	F_{ST}: 0.180	F_{ST}: 0.302	F_{ST}: 0.305	F_{ST}: 0.403	F_{ST}: 0.220	F_{ST}: 0.248	F_{ST}: 0.250	F_{ST}: 0.325	F_{ST}: 0.125	F_{ST}: 0.081	F_{ST}: 0.167	F_{ST}: 0.044	
New Zealand	R_{ST}: 0.032	R_{ST}: 0.118	R_{ST}: 0.113	R_{ST}: 0.045	R_{ST}: 0.121	R_{ST}: 0.096	R_{ST}: 0.191	R_{ST}: 0.124	R_{ST}: 0.071	R_{ST}: 0.173	R_{ST}: 0.117	R_{ST}: 0.082	R_{ST}: 0.042	R_{ST}: 0.098	R_{ST}: 0.345	R_{ST}: 0.020	R_{ST}: 0.066
	F_{ST}: 0.157	F_{ST}: 0.179	F_{ST}: 0.214	F_{ST}: 0.189	F_{ST}: 0.161	F_{ST}: 0.324	F_{ST}: 0.278	F_{ST}: 0.398	F_{ST}: 0.213	F_{ST}: 0.245	F_{ST}: 0.225	F_{ST}: 0.328	F_{ST}: 0.132	F_{ST}: 0.147	F_{ST}: 0.253	F_{ST}: 0.093	F_{ST}: 0.084

Table S4.3 Pairwise F_{ST} and R_{ST} values between 18 sampling locations, as calculated by Genepop version 4.1.0 (Rousset 2008). Values based on corrected allele and genotype frequencies as determined by MICRO-CHECKER version 2.2.3 (Van Oosterhout et al. 2004). Shading indicates values significantly different from 0.

	USA	Mexico	Belize	Costa Rica	S. Florida	Puerto Rico	Aruba	Ecuador	Bermuda	Spain	Portugal	Morocco	Hawaii	Samoa	Fiji	New Caledonia	Australia
Mexico	R_{ST}: 0.013 F_{ST}: 0.001																
Belize	R_{ST}: -0.005 F_{ST}: 0.007	R_{ST}: -0.010 F_{ST}: 0.005															
Costa Rica	R_{ST}: 0.008 F_{ST}: 0.002	R_{ST}: 0.044 F_{ST}: 0.010	R_{ST}: 0.041 F_{ST}: 0.011														
S. Florida	R_{ST}: 0.004 F_{ST}: 0.029	R_{ST}: -0.003 F_{ST}: 0.015	R_{ST}: 0.001 F_{ST}: 0.037	R_{ST}: 0.048 F_{ST}: 0.037													
Puerto Rico	R_{ST}: 0.057 F_{ST}: 0.073	R_{ST}: 0.201 F_{ST}: 0.084	R_{ST}: 0.177 F_{ST}: 0.058	R_{ST}: 0.067 F_{ST}: 0.084	R_{ST}: 0.192 F_{ST}: 0.129												
Aruba	R_{ST}: 0.029 F_{ST}: 0.075	R_{ST}: 0.021 F_{ST}: 0.094	R_{ST}: 0.018 F_{ST}: 0.064	R_{ST}: 0.076 F_{ST}: 0.074	R_{ST}: 0.038 F_{ST}: 0.108	R_{ST}: 0.267 F_{ST}: 0.148											
Ecuador	R_{ST}: 0.081 F_{ST}: 0.165	R_{ST}: 0.203 F_{ST}: 0.207	R_{ST}: 0.165 F_{ST}: 0.194	R_{ST}: 0.152 F_{ST}: 0.175	R_{ST}: 0.221 F_{ST}: 0.254	R_{ST}: 0.192 F_{ST}: 0.332	R_{ST}: 0.266 F_{ST}: 0.214										
Bermuda	R_{ST}: -0.014 F_{ST}: 0.071	R_{ST}: 0.012 F_{ST}: 0.071	R_{ST}: 0.024 F_{ST}: 0.112	R_{ST}: -0.005 F_{ST}: 0.096	R_{ST}: 0.012 F_{ST}: 0.075	R_{ST}: 0.125 F_{ST}: 0.155	R_{ST}: 0.064 F_{ST}: 0.175	R_{ST}: 0.174 F_{ST}: 0.301									
Spain	R_{ST}: 0.082 F_{ST}: 0.077	R_{ST}: 0.084 F_{ST}: 0.081	R_{ST}: 0.057 F_{ST}: 0.082	R_{ST}: 0.127 F_{ST}: 0.083	R_{ST}: 0.085 F_{ST}: 0.133	R_{ST}: 0.224 F_{ST}: 0.201	R_{ST}: 0.056 F_{ST}: 0.184	R_{ST}: 0.130 F_{ST}: 0.179	R_{ST}: 0.079 F_{ST}: 0.194								
Portugal	R_{ST}: 0.078	R_{ST}: 0.040	R_{ST}: 0.083	R_{ST}: 0.092	R_{ST}: 0.095	R_{ST}: 0.252	R_{ST}: 0.082	R_{ST}: 0.208	R_{ST}: 0.090	R_{ST}: 0.160							

	F_{ST}: 0.095	F_{ST}: 0.115	F_{ST}: 0.128	F_{ST}: 0.085	F_{ST}: 0.165	F_{ST}: 0.220	F_{ST}: 0.201	F_{ST}: 0.273	F_{ST}: 0.213	F_{ST}: 0.135							
Morocco	R_{ST}: 0.019	R_{ST}: - 0.002	R_{ST}: 0.005	R_{ST}: 0.017	R_{ST}: 0.023	R_{ST}: 0.127	R_{ST}: 0.036	R_{ST}: 0.114	R_{ST}: 0.015	R_{ST}: 0.105	R_{ST}: 0.010						
	F_{ST}: 0.107	F_{ST}: 0.121	F_{ST}: 0.112	F_{ST}: 0.115	F_{ST}: 0.188	F_{ST}: 0.198	F_{ST}: 0.240	F_{ST}: 0.290	F_{ST}: 0.258	F_{ST}: 0.083	F_{ST}: 0.091						
Hawaii	R_{ST}: 0.003	R_{ST}: 0.038	R_{ST}: 0.0178	R_{ST}: 0.006	R_{ST}: 0.016	R_{ST}: 0.062	R_{ST}: 0.037	R_{ST}: 0.099	R_{ST}: - 0.006	R_{ST}: 0.077	R_{ST}: 0.106	R_{ST}: 0.038					
	F_{ST}: 0.042	F_{ST}: 0.041	F_{ST}: 0.056	F_{ST}: 0.060	F_{ST}: 0.058	F_{ST}: 0.123	F_{ST}: 0.153	F_{ST}: 0.256	F_{ST}: 0.152	F_{ST}: 0.140	F_{ST}: 0.146	F_{ST}: 0.166					
Samoa	R_{ST}: 0.034	R_{ST}: 0.008	R_{ST}: 0.012	R_{ST}: 0.074	R_{ST}: 0.023	R_{ST}: 0.204	R_{ST}: 0.020	R_{ST}: 0.171	R_{ST}: 0.027	R_{ST}: 0.075	R_{ST}: 0.034	R_{ST}: 0.025	R_{ST}: 0.064				
	F_{ST}: 0.165	F_{ST}: 0.189	F_{ST}: 0.263	F_{ST}: 0.227	F_{ST}: 0.223	F_{ST}: 0.294	F_{ST}: 0.360	F_{ST}: 0.471	F_{ST}: 0.225	F_{ST}: 0.298	F_{ST}: 0.272	F_{ST}: 0.315	F_{ST}: 0.181				
Fiji	R_{ST}: 0.216	R_{ST}: 0.151	R_{ST}: 0.204	R_{ST}: 0.316	R_{ST}: 0.232	R_{ST}: 0.508	R_{ST}: 0.218	R_{ST}: 0.439	R_{ST}: 0.242	R_{ST}: 0.193	R_{ST}: 0.150	R_{ST}: 0.157	R_{ST}: 0.273	R_{ST}: 0.074			
	F_{ST}: 0.226	F_{ST}: 0.241	F_{ST}: 0.304	F_{ST}: 0.284	F_{ST}: 0.290	F_{ST}: 0.344	F_{ST}: 0.395	F_{ST}: 0.517	F_{ST}: 0.262	F_{ST}: 0.329	F_{ST}: 0.312	F_{ST}: 0.335	F_{ST}: 0.259	F_{ST}: 0.067			
New Caledonia	R_{ST}: 0.005	R_{ST}: 0.045	R_{ST}: 0.046	R_{ST}: 0.016	R_{ST}: 0.045	R_{ST}: 0.088	R_{ST}: 0.077	R_{ST}: 0.123	R_{ST}: - 0.001	R_{ST}: 0.118	R_{ST}: 0.069	R_{ST}: 0.023	R_{ST}: 0.018	R_{ST}: 0.036	R_{ST}: 0.235		
	F_{ST}: 0.158	F_{ST}: 0.185	F_{ST}: 0.260	F_{ST}: 0.226	F_{ST}: 0.210	F_{ST}: 0.297	F_{ST}: 0.357	F_{ST}: 0.460	F_{ST}: 0.208	F_{ST}: 0.304	F_{ST}: 0.277	F_{ST}: 0.332	F_{ST}: 0.164	F_{ST}: 0.006	F_{ST}: 0.101		
Australia	R_{ST}: 0.010	R_{ST}: 0.046	R_{ST}: 0.054	R_{ST}: 0.049	R_{ST}: 0.029	R_{ST}: 0.149	R_{ST}: 0.088	R_{ST}: 0.206	R_{ST}: 0.016	R_{ST}: 0.138	R_{ST}: 0.106	R_{ST}: 0.044	R_{ST}: 0.026	R_{ST}: 0.026	R_{ST}: 0.261	R_{ST}: 0.001	
	F_{ST}: 0.175	F_{ST}: 0.201	F_{ST}: 0.258	F_{ST}: 0.234	F_{ST}: 0.215	F_{ST}: 0.307	F_{ST}: 0.356	F_{ST}: 0.447	F_{ST}: 0.249	F_{ST}: 0.285	F_{ST}: 0.279	F_{ST}: 0.320	F_{ST}: 0.171	F_{ST}: 0.079	F_{ST}: 0.186	F_{ST}: 0.043	
New Zealand	R_{ST}: 0.044	R_{ST}: 0.140	R_{ST}: 0.137	R_{ST}: 0.057	R_{ST}: 0.146	R_{ST}: 0.114	R_{ST}: 0.157	R_{ST}: 0.117	R_{ST}: 0.070	R_{ST}: 0.156	R_{ST}: 0.129	R_{ST}: 0.077	R_{ST}: 0.045	R_{ST}: 0.098	R_{ST}: 0.359	R_{ST}: 0.016	R_{ST}: 0.065
	F_{ST}: 0.145	F_{ST}: 0.163	F_{ST}: 0.225	F_{ST}: 0.184	F_{ST}: 0.174	F_{ST}: 0.310	F_{ST}: 0.311	F_{ST}: 0.404	F_{ST}: 0.206	F_{ST}: 0.250	F_{ST}: 0.229	F_{ST}: 0.297	F_{ST}: 0.142	F_{ST}: 0.114	F_{ST}: 0.250	F_{ST}: 0.082	F_{ST}: 0.087

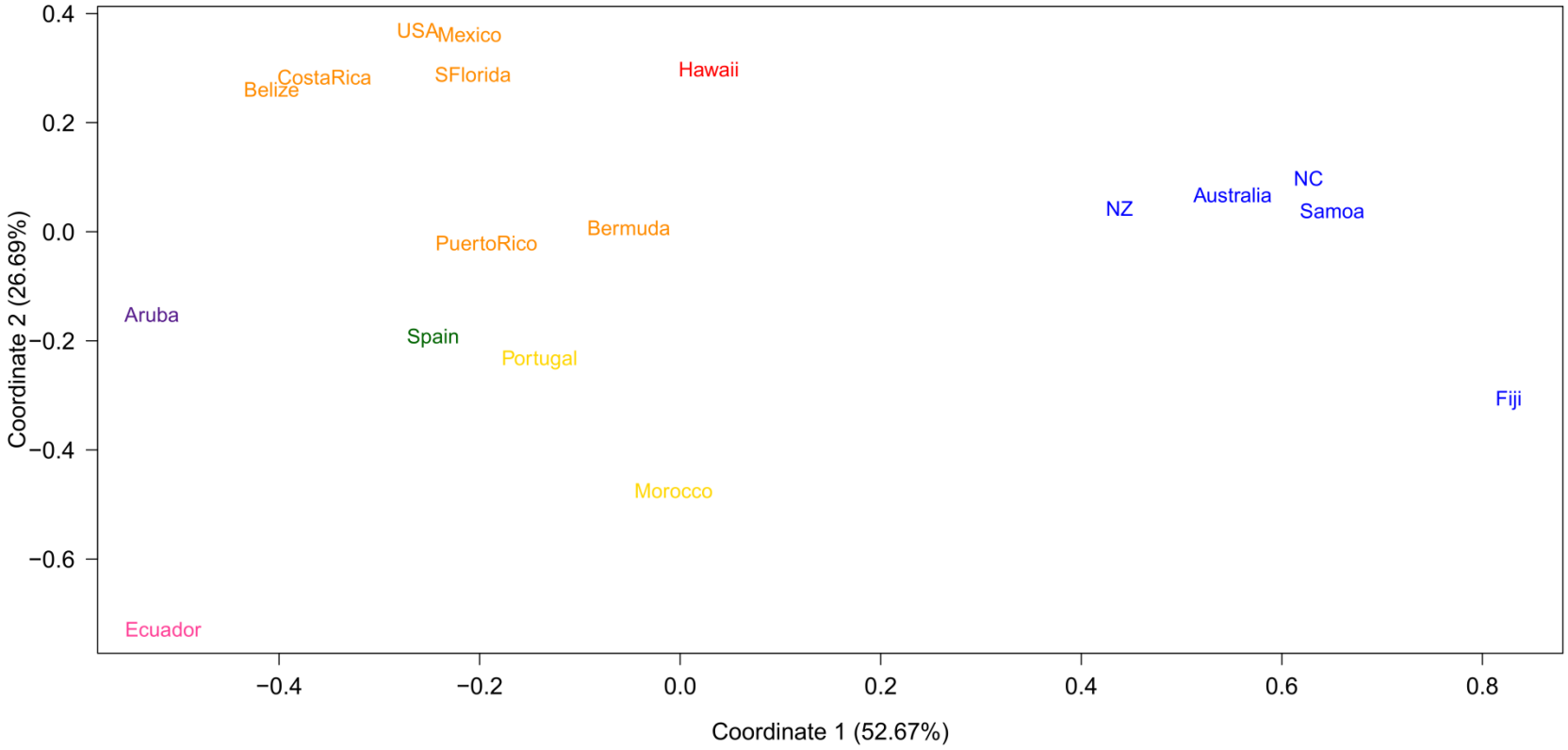

Figure S4.1 Principal Coordinate Analysis Sampling locations are color coordinated to correspond with populations as depicted in Fig 4.2 and Fig 4.3. Variance explained by each axis is given in parantheses.

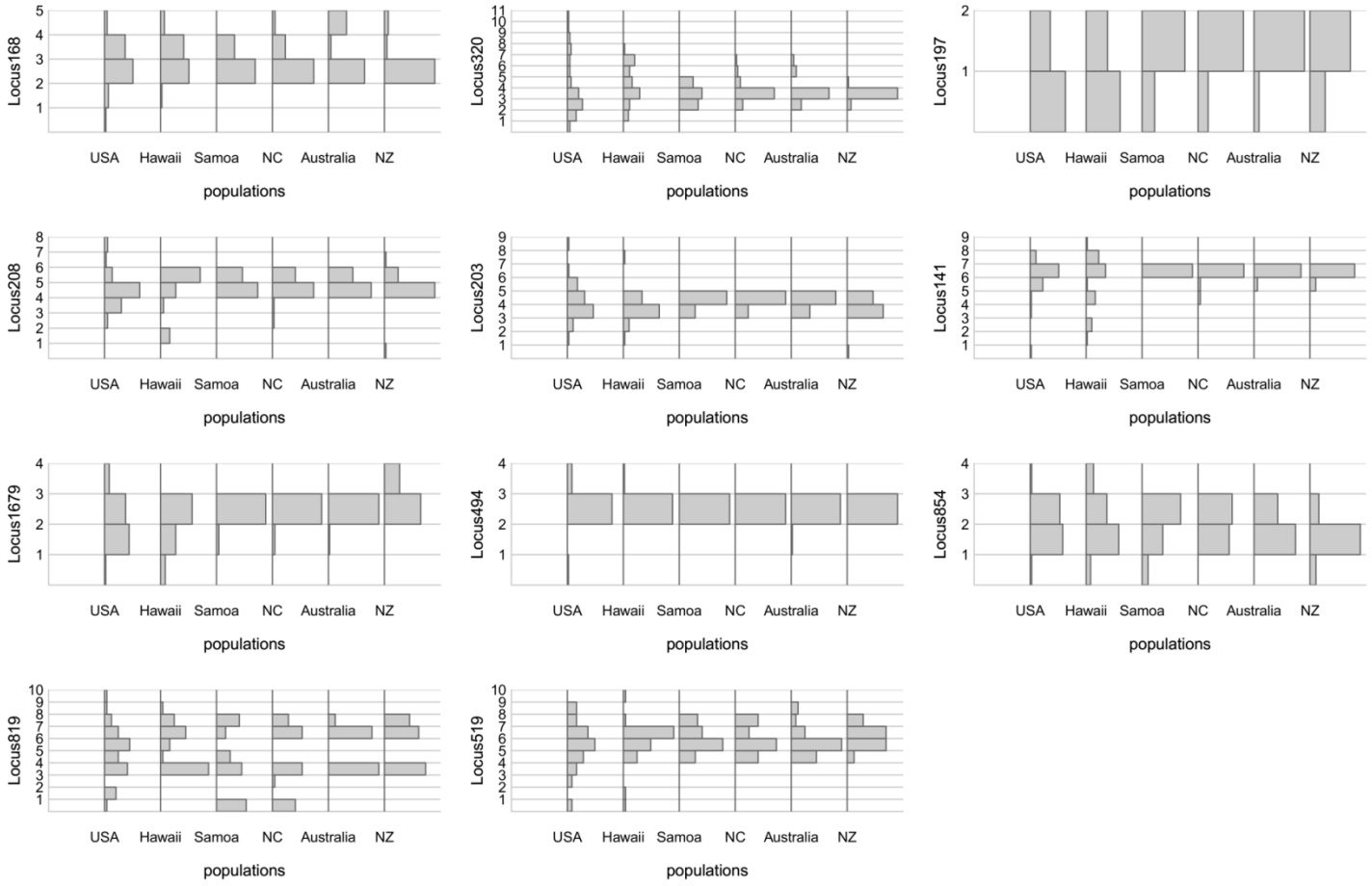

Figure S4.2 Allelic clinal pattern across Pacific. The change in the frequency of alleles across 11 loci as colonies move further west from the United States. Allele frequency values have been corrected to account for differences in sample sizes. The increase of some alleles with the disappearance of others along the wave front of the range expansion is indicative of gene surfing and genetic drift.

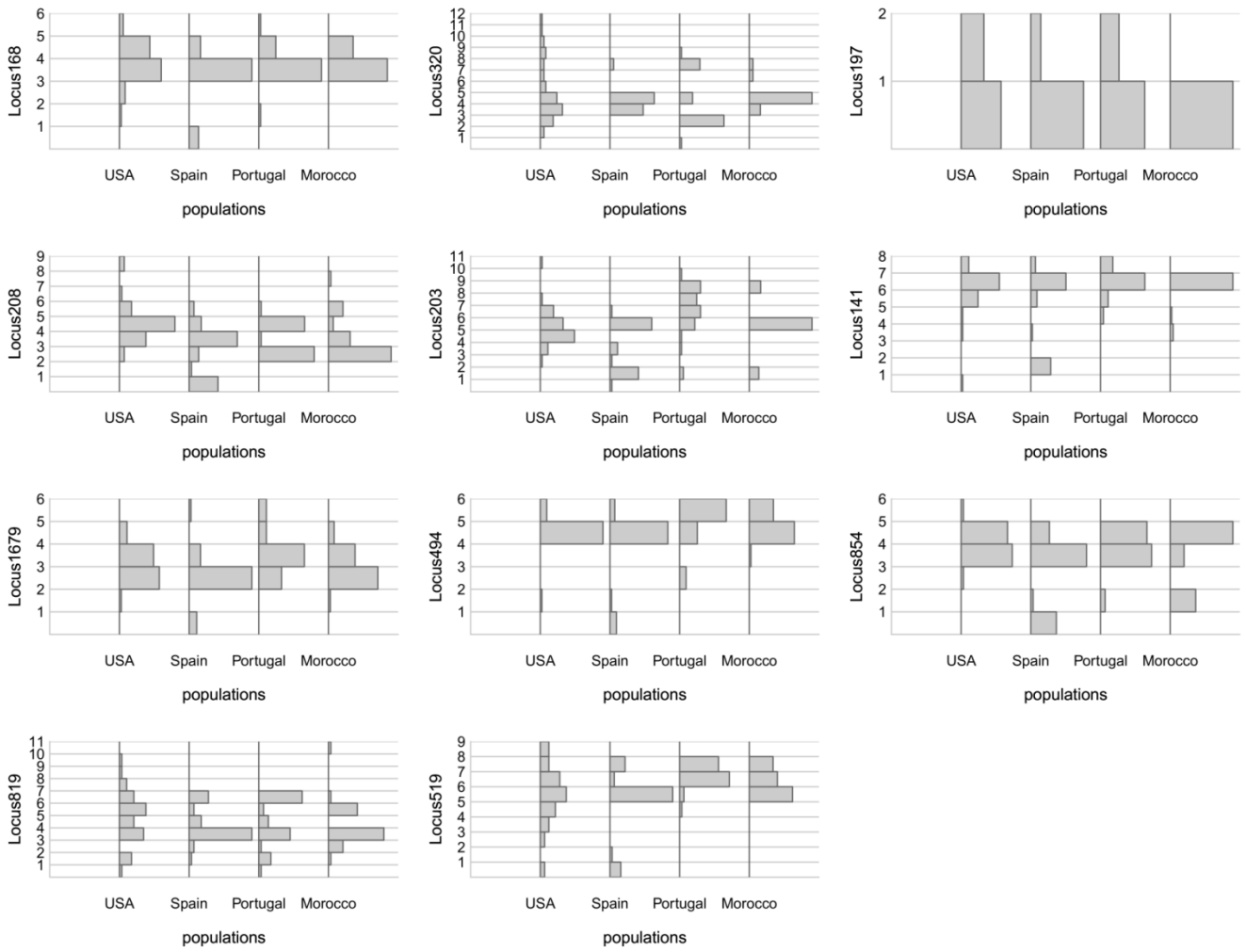

Figure S4.3 Allele frequencies across the Atlantic. The change in the frequency of alleles across 11 loci as colonies move further east from the United States. Values based on corrected allele frequencies to account for differences in sample sizes. Allelic clinal patterns demonstrate gene surfing but the relatively high frequency of rare alleles in more eastern colonies indicates multiple colonizations and founder effect.

Chapter 5

Extreme heterogeneity in parasitism despite low population genetic structure among monarch butterflies inhabiting the Hawaiian Islands

Reprinted material from: A.A. Pierce, J. C. de Roode, S.M. Altizer, R.A. Bartel. *PLoS ONE*. 9,6:e100061(June, 2014)

Introduction

Much work during the past two decades has focused on understanding the spatial ecology of host-pathogen interactions. Some studies have shown that genetic variation in traits affecting host resistance and pathogen virulence can generate spatial variation in infection patterns (Ravensdale *et al.* 2011). Other work demonstrated that landscape-level heterogeneity in factors such as habitat quality, the relative abundance of host species, and geographic features such as rivers and mountains, can affect the spatial spread and prevalence of pathogens (Allan *et al.* 2003; Ostfeld *et al.* 2005; Smith *et al.* 2002). Understanding the pattern of spatial heterogeneity in infection is crucial for identifying key drivers of pathogen persistence and for predicting and managing disease risk.

Host dispersal patterns can have important consequences for spatial processes and the ecology and evolution of host-pathogen interactions (Cronin 2009; Ostfeld *et al.* 2005; Riley 2007; Thrall & Antonovics 1995; Thrall & Burdon 1997). Some studies have shown that host movement among patches can facilitate pathogen persistence at the landscape level (Altizer *et al.* 2011; Hassell *et al.* 1991; Johnson *et al.* 2011). On the other hand, directed seasonal movement (i.e., long distance migration) can lower parasite transmission by allowing hosts to escape from parasitized locations (Altizer *et al.* 2011), as has been suggested for warble flies affecting reindeer (Folstad *et al.* 1991), and protozoan parasites infecting monarch butterflies

(Bartel *et al.* 2011). Movement can further result in gene flow and the spread of host resistance alleles across a landscape, with studies of anther-smut in plants and viruses in moths showing that limited host movement or gene flow can generate high spatial heterogeneity in prevalence, allowing some patches to become heavily infected while others remain disease-free (Best *et al.* 2011; Carlsson-Graner 2006; Carlsson-Graner & Thrall 2002).

Here, we examined spatial heterogeneity in the occurrence of an obligate protozoan parasite (*Ophryocystis elektroscirrha*, hereafter called *OE*) infecting monarch butterflies (*Danaus plexippus*) on the island chain of Hawaii. Monarchs inhabit islands and continents worldwide and occupy a subset of the range of their larval milkweed host plants (Ackery & Vane-Wright 1984). Monarchs are best known for undertaking a spectacular long-distance migration (up to 5000 km roundtrip) in eastern North America (Brower & Malcolm 1991; Urquhart & Urquhart 1978), but they also form non-migratory populations that breed year-round in tropical and subtropical locations such as the Caribbean Islands, Central America and Hawaii. Monarchs colonized Hawaii and other Pacific Islands in the mid-1800s (Shephard *et al.* 2002; Zalucki & Clarke 2004) following the introduction of their host plants, and now occupy most of the eight Hawaiian islands (Ackery & Vane-Wright 1984). Monarchs in Hawaii breed year-round in habitats containing introduced larval host plants, especially *Asclepias physocarpa, Calotropis gigantea,* and *C. procera*. Hawaiian monarchs are smaller than North American migratory monarchs (Altizer & Davis 2010), and microsatellite markers showed that Hawaiian monarchs are genetically distinct from those in North America and New Zealand (Lyons *et al.* 2012).

All monarch populations examined to date are parasitized by *OE*, and prevalence varies widely among regions (Altizer & De Roode 2013). Prevalence reaches the highest levels in monarch populations that breed year-round (e.g., South Florida, Cuba) and is much lower in populations that migrate long distances (Altizer *et al.* 2000; Leong *et al.* 1997a; Leong *et al.* 1997b). In particular, non-migratory monarchs likely experience higher rates of transmission due to continuous breeding activity and extended use of the same host plants for egg deposition

(Altizer *et al.* 2011; Altizer 2004), as parasites are transmitted when infected adults scatter spores onto milkweed leaves (Leong *et al.* 1997b; McLaughlin & Myers 1970). Larvae ingest the spores, parasites replicate internally, and adults emerge with millions of dormant spores on the outsides of their bodies (de Roode *et al.* 2007; Leong *et al.* 1997b). While no further parasite replication occurs at the monarch adult stage, infected adults suffer from decreased body size, eclosion success, lifespan, flight performance and migration success (Altizer & Oberhauser 1999; Bartel *et al.* 2011; Bradley & Altizer 2005).

In this study, we sampled monarchs and recorded *OE* infection across replicate sites within each of four Hawaiian Islands over multiple years. Based on previously documented associations between monarch migratory ecology and parasite prevalence, we expected that *OE* prevalence would reach high levels across all sites sampled owing to year-round breeding and the limited potential for long-distance movement among monarchs inhabiting these oceanic islands. Because our field analysis showed extreme heterogeneity in *OE* prevalence within and among islands (and lower than expected prevalence overall), we further used neutral genetic (microsatellite) markers to examine evidence for host population structure. In particular, we asked whether genetic evidence indicates that host movement within and among islands might be limited, such that between-site variation in prevalence could be attributed to locally structured host sub-populations that are isolated from other patches. Finally, we asked whether measures of host neutral diversity (as indicators of genome-wide heterozygosity) might correlate negatively with parasite infection probability at the individual or patch level, as suggested by prior work in Soay sheep, sea lions and several other species, whereby animals with greater genome-wide diversity can better resist parasite infections than more inbred hosts (Acevedo-Whitehouse *et al.* 2009; Coltman *et al.* 1999; Smith *et al.* 2012; Whitehorn *et al.* 2011).

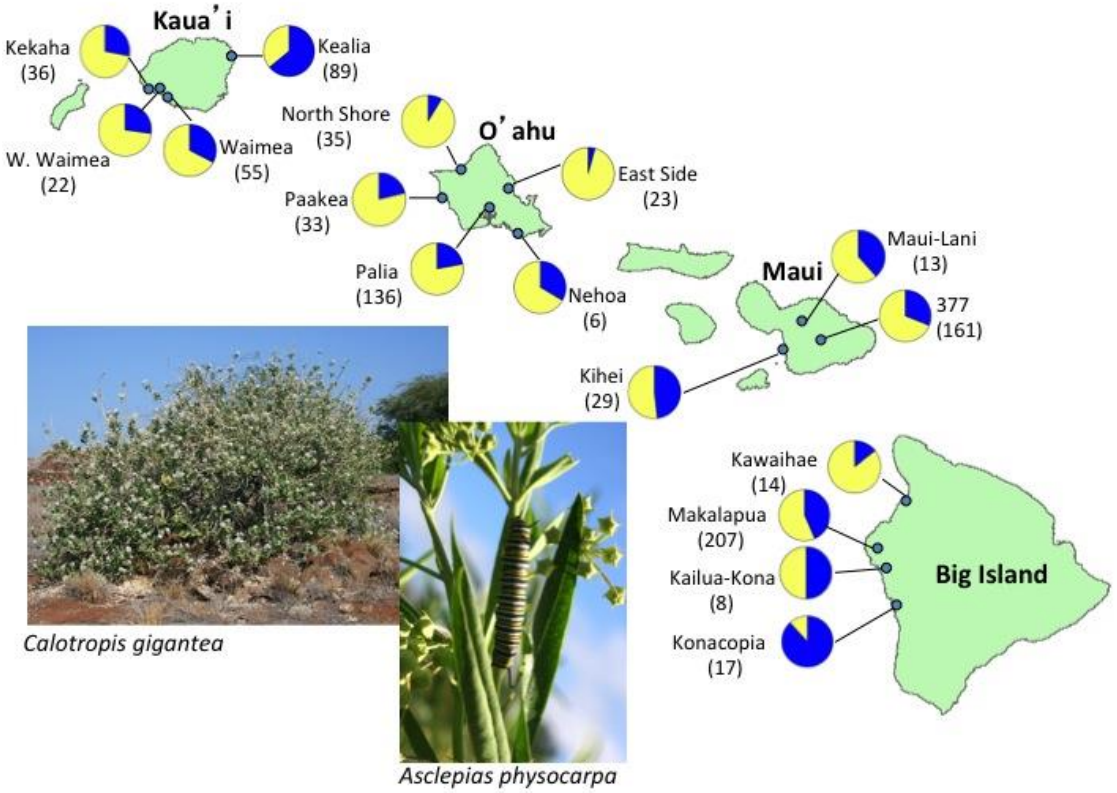

Figure 5.1. Variation in parasite prevalence on four islands of Hawaii based on field sampling from 2007-2010. Dark shading indicates the proportion of monarch infected with *OE* within subpopulations. Sample sizes are indicated within parentheses. Red dots indicate sites from which samples were further analyzed for microsatellite markers (Table S2). Photographs show two representative host plant species common throughout most islands.

Materials and Methods

Field sampling. We sampled monarchs and their parasites once per year in each of three years (2007, 2009, 2010) across four islands in Hawaii: Hawaii (Big Island), Oahu, Maui, and Kauai (Figure 1; Table 1). These islands differ in their total area and human population density. On each island, we identified 3-5 representative habitat patches where monarchs and their

milkweed host plants (*Calotropis* or *Asclepias spp.*) occur (Table S1). Sites were separated by a minimum of 5 km and early site visits indicated that monarch adults and larvae were concentrated in host plant patches, as has been shown before for monarchs (Zalucki & Kitching 1982). Field surveys occurred during the rainy season (Jan-Feb), and with each progressive year, we identified additional sampling sites (Table 1). In 2007, only the Big Island and Oahu were visited and sampled (N = 117 monarchs, 3 sites). In 2009, we expanded field efforts to Kauai and Maui and included more sites on the Big Island and Oahu (N = 388 monarchs, 10 sites), and in 2010, we added sites on all islands (N = 380, 15 sites).

Table 5.1. Monarchs sampled in Hawaii by collection site and year, with sample sizes (count) and the proportion of monarchs heavily infected with *O. elektroscirrha*. Sites in boldface were also examined for microsatellite markers (Table S3).

	2007		**2009**		**2010**		**Total Count**	**Total Proportion Infected**
	Count	Average Proportion Infected	Count	Average Proportion Infected	Count	Average Proportion Infected		
Big Island	65	0.48	100	0.36	81	0.54	246	0.45
Kailua-Kona					8	0.50	8	0.50
Kawaihae					14	0.14	14	0.14
Konacopia Farms					17	0.88	17	0.88
Makalapua	65	0.48	100	0.36	42	0.55	207	0.43
Kauai			102	0.45	101	0.45	203	0.45
Kealia Beach			43	0.81	46	0.48	89	0.64
Kekaha Beach			26	0.23	11	0.36	37	0.27
Waimea			22	0.18	33	0.42	55	0.33
West Waimea			11	0.09	11	0.45	22	0.27
Maui			106	0.20	97	0.49	203	0.34
Kihei					29	0.48	29	0.48
Maui 377			100	0.20	61	0.49	161	0.31
Maui-Lani			6	0.17	7	0.57	13	0.38

Oahu	52	0.13	80	0.21	101	0.19	233	0.18
East Side					23	0.04	23	0.04
Nehoa St.					6	0.33	6	0.33
North Shore			5	0.00	30	0.10	35	0.09
Paakea Rd.	21	0.24	12	0.17			33	0.21
Palai St.	31	0.06	63	0.24	42	0.31	136	0.22
Grand Total	**117**	**0.32**	**388**	**0.31**	**380**	**0.41**	**885**	**0.35**

The field collections for this project did not involve endangered or protected species. We collected at three different private sites (Palia, Nehoa, and Konacopia) after receiving permission from S. and A. Montgomery, S. Marques, and E. Kilpatrick. The remaining collection sites consisted of roadsides, parks, or unprotected areas. No permits were necessary to collect these monarchs in Hawaii (collecting non-endangered butterflies in public areas is not prohibited in the United States, and monarchs themselves are not native to the Hawaiian Islands). All butterflies were transported to the University of Georgia, Athens, GA, under permission from the United States Department of Agriculture (USDA PPQ-526 Permit #11-04112 and Permit #06-01690 to S. Altizer).

Adult monarchs were captured using an aerial net between 0900 and 1600 hr. Following capture, monarchs were stored individually in glassine envelopes and held at 14°C for up to 6 hr prior to sampling. We recorded sex and forewing length to the nearest 0.01mm. Wing condition, which qualitatively reflects age or distance traveled, was recorded in two ways. First, we recorded wing damage on a 0-4 scale, based on the number of wings with evidence of tears or other physical damage as might be caused by predators or contact with hard surfaces. Second, we recorded wing wear on an ordinal scale of 1-5, based on the level of scale loss (from newly emerged to nearly transparent wings) following Cockrell et al. (Cockrell *et al.* 1993).

***Measuring parasite prevalence and transmission*:** Adult monarchs captured at each site were scored for parasite infection status based on the number of *OE* spores transferred to a 2.5 cm-diameter transparent sticker pressed against adult abdomens (described in Altizer et al. (Altizer *et al.* 2000)). Samples were examined at 63X magnification to record infection scores on a 0-5 scale. This method is highly sensitive and past work showed that categorical scores are highly correlated with Log_{10} of quantitative spore loads (de Roode *et al.* 2009) measured using an agitation and hemocytometer counting chamber method as described in Leong et al. [25] and Altizer et al. [26]. Samples with more than 100 spores were considered heavily infected; this classification includes the two highest spore load categories defined by Altizer *et al.* (2000). Importantly, heavily infected monarchs are those with infections likely caused by the ingestion of one or more spores as larvae, thus resulting in these individuals experiencing negative consequences of within-host replication (de Roode *et al.* 2007). In contrast, lower spore numbers can result from passive transfer of spores between adult butterflies (Altizer 2004; de Roode *et al.* 2009; de Roode *et al.* 2007); these dormant spores cannot directly infect adults and must be ingested by a larva to cause a new infection. Following scoring infection status, we released the majority of monarchs at the collection site and kept a subset for genetic analysis (Table S2).

Microsatellite analyses: We used polymorphic microsatellite markers to determine whether monarchs were genetically differentiated between sites, or whether extensive gene flow occurs. Microsatellite marker development and PCR protocol were as described in Lyons et al 2012. Briefly, DNA for PCR was extracted from a 0.5 mm section of butterfly abdomen (male butterflies) or thorax (female butterflies) using the UltraClean DNA Isolation Kit from Mo-Bio (Carlsbad, CA, USA) and quantified using a Nanodrop 2000. We did not use female abdominal tissue as this could possibly contain DNA from male sperm. PCR was carried out in 15 μl multiplex reactions using the Type-It Microsatellite PCR kit (Qiagen). Only a subset of monarchs scored for infection status were collected for genetic work, so sites with nine or more

samples were chosen to genotype (Table S2). In total, we genotyped 42 butterflies from two sites on the Big Island (Kawaiahea, N = 9; Makalapua, N = 33), 48 from four sites on Oahu (East Side, N = 9; North Shore, N = 9; Paakea, N = 11; Palia, N = 19), and 9 from one site on Maui (Maui377, N = 9) for 16 microsatellite loci (Table S3).

Analysis of field and genetic data: For field-collected samples, we used logistic regression (IBM SPSS Statistics 20.0) to examine the main effects of year, island, and site (as a random effect, nested within island) on variation in monarch infection status (at the individual level) as a binomial variable. We also included the island*year interaction effect, and individual-level predictors of sex, forewing length, wing damage and wing wear in the full model. Prior to analysis we excluded data from sites for which fewer than 5 samples were available. In a separate analysis, we investigated whether site-level variation in patch size, land use type (categorized as urban, suburban or rural), and host plant species explained variation in average prevalence measures (with details provided in Supplementary Materials).

To investigate host genetic differentiation, we used the software Arlequin 3.5.1.2 (Excoffier & Lischer 2010) to calculate observed and expected heterozygosity at each microsatellite locus for each site. We also used Arlequin to calculate deviations from Hardy-Weinberg equilibrium for each locus at each site, and used a sequential Bonferroni correction (Rice 1989) to determine whether observed and expected heterozygosity levels were significantly different (α=0.05). We excluded locus 137, which was not polymorphic or in Hardy-Weinberg in at least 5 out of 7 populations; the remaining 15 loci were used in subsequent analyses (Table S4).

Samples for each site were resampled with replacement using Poptools (Hood 2010) to standardize sample size across sites for comparison of relative levels of genetic diversity. To do this, we calculated genetic diversity (using the value 1-Qinter) using Genepop version 4.1.0 (Rousset 2008) and allelic richness using ADZE-1.0 (Szpiech *et al.* 2008), which utilizes a rarefaction approach to account for differences in sample size. To understand the relative

magnitude of within- and between-population genetic diversity, we carried out a locus by locus analysis of molecular variance using 10,000 permutations in Arlequin 3.5.1.2 (Excoffier & Lischer 2010) for six of the sites (Kawaihae, Big Island; Makalapua, Big Island; East Side, Oahu; North Shore, Oahu; Paakea, Oahu; Palia, Oahu). In this analysis, we combined sites based on island, and compared this to the variation among populations within groups (i.e. variation among sites within the same island) as well as genetic variation within sites.

We used the software STRUCTURE version 2.3.2.1 (Pritchard *et al.* 2000) to investigate population structure. We used an admixture model with uncorrelated allele frequencies to avoid the risk of overestimating the number of populations, K, and used the LOCPRIOR model to include location information for each butterfly. We did the latter to ensure that STRUCTURE would be able to detect subtle population structure. We also included 16 butterflies from New Zealand (Christchurch, Jan 2011) for comparison, as monarch populations in Hawaii and New Zealand were established within the last 170 years, and are thought to originate from North America through trans-Pacific dispersal (Shephard *et al.* 2002; Vane-Wright 1993; Zalucki & Clarke 2004). Therefore, the inclusion of the New Zealand population allows us to determine that our markers are able to detect subtle and newly formed population structure.

We also examined population genetic structure using F_{ST} and R_{ST} statistics. These statistics are commonly used to calculate genetic differentiation, with levels of 0 indicating panmixia, and values higher than 0 indicating genetic differentiation. R_{ST} was developed as a more appropriate statistic for microsatellite markers, based on its use of a stepwise mutation model (Slatkin 1995), rather than the infinite alleles model utilized in F_{ST} statistics (Balloux & Lugon-Moulin 2002). Permutation tests (using 10,000 permutations), as implemented in Arlequin 3.5.1.2 (Excoffier & Lischer 2010) were used to determine whether pairwise F_{ST} and R_{ST} values were significantly different from 0. To further examine population genetic structure, we analyzed the correlation between site collection time and measures of genetic differentiation

using Mantel tests implemented in the vegan library version 2.0-0 (Oksanen *et al.* 2011) in the statistical package R version 3.0.1. Finally, we calculated heterozygosity at the individual level by determining the proportion of heterozygous loci per butterfly. To investigate the effect of heterozygosity on infection status, we treated infection status as a binomial variable and performed a logistic regression using a generalized linear model (GLM with binomial error distribution, logit link) in R version 3.0.1.

Results

Parasite prevalence and transmission: On average, 35.5% of monarchs were heavily infected with *OE* across all sites and years (N = 885; Table 1). We detected high variation in prevalence both within and among islands on the Hawaiian archipelago (Table 1; Figures 1-2), with the average proportion of heavily infected monarchs per site per year ranging from 0.00 to 0.88. Logistic regression showed a significant main effect of island on infection probability (Wald $\chi^2 = 10.17$, d.f. = 3, P = 0.017). In particular, the outer islands of Kauai and the Big Island showed the highest average infection levels (e.g., proportion of heavily infected monarchs on the Big Island, N = 246, and Kauai, N = 203, were both 0.45 when averaged across sites and years). By comparison, average prevalence was much lower on Oahu (proportion infected = 0.19, N = 233). Although we also observed a significant main effect of year (Wald $\chi^2 = 16.13$, d.f. = 2, $P < 0.001$), with infection prevalence higher for 2010 than for 2007 or 2009 (Figure 2), differences in infection probability across islands were generally consistent among years (Table 1; Figure 2), as supported by a non-significant interaction between island and year (Wald $\chi^2 = 5.20$, d.f. = 4, P = 0.26).

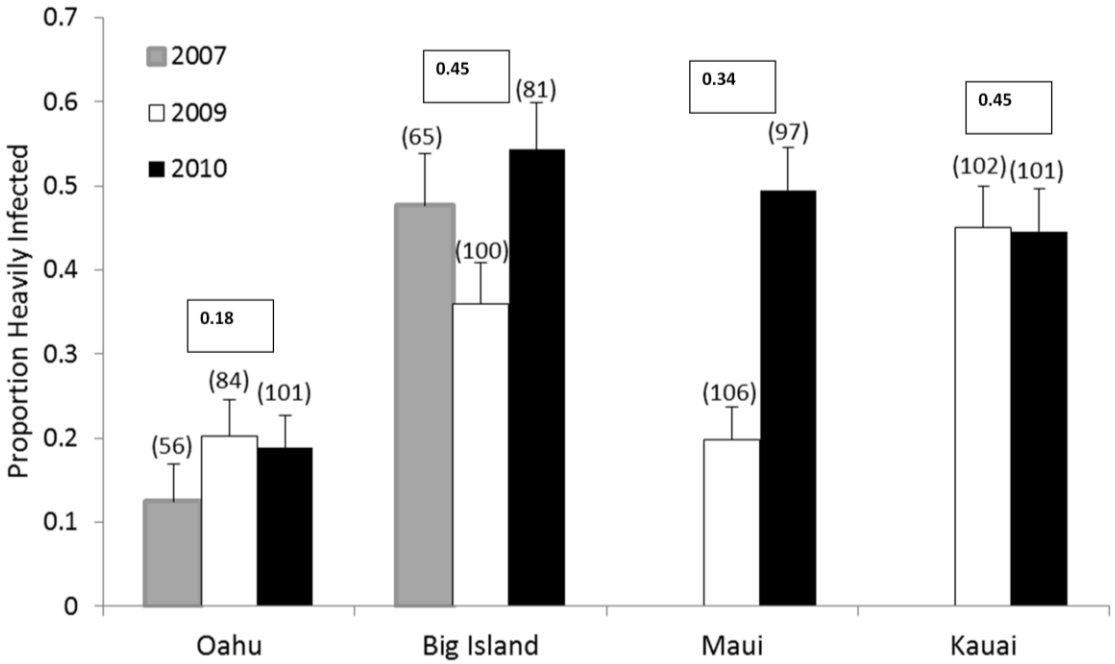

Figure 5.2. **Proportion of monarchs heavily infected with *OE* for four Hawaiian Islands from 2007-2010.** Sample sizes are in parentheses. Average prevalence per island over all sample years is shown in bold type. Field surveys in 2007 focused on the Big Island and Oahu only. Beginning in 2009, we sampled Maui and Kauai, and visited 3-5 sampling sites for each of the 4 islands. Sample sizes per island per year ranged from 56 to 105 (Table 1). Error bars indicate standard errors.

At a finer scale, we detected strong within-island heterogeneity in the proportion of infected monarchs (Figure 1), and the effect of site nested within island was highly significant (Wald χ^2 = 40.38, d.f. = 4, $P < 0.001$). Of the 16 sites monitored through 2010, 9 were assessed for two or more consecutive years for monarch presence and parasite prevalence (Figure 2). Although some sites showed consistently low or high prevalence (Table 1), a separate simple linear regression analysis showed that *OE* prevalence per site in a given year was not predictive of prevalence the following year ($R^2 = 0.144$, $t_{11} = 1.30$, $P = 0.224$).

Our multivariate logistic regression analysis further controlled for individual-level variables that might explain variation in *OE* infection. At the individual level, males (proportion infected = 0.49, N = 548) had higher infection prevalence than females (proportion infected = 0.44, N = 337) and this effect was highly significant (Wald χ^2 = 16.0, d.f. = 1, $P < 0.001$).

Forewing length was negatively associated with infection status, such that infected monarchs had smaller wings than healthy butterflies (Wald $\chi^2 = 9.95$, d.f. = 1, P = 0.002). Wing wear (reflecting wing scale loss) also predicted variation in infection probability (Wald $\chi^2 = 10.51$, d.f. = 1, P = 0.001), such that infected monarchs were more likely to show greater wing scale loss. Wing damage (as an index of tatter), however, was not associated with monarch infection status (Wald $\chi^2 = 0.32$, d.f. = 1, P = 0.57).

A separate analyses of variance based on average prevalence by site and year (N = 28) showed that no site-level measurements (e.g., patch area, host plant species, habitat type) were significant predictors of variation in parasitism (results presented in Supplementary Materials). Although collection times (within the 0900 – 1600 hr range) varied among sites, there was no correlation between collection time and average infection prevalence (P=0.52 for 2009, when detailed collection times were recorded).

Table 5.2. Results of analysis of molecular variance (AMOVA) comparing samples from locations within two islands (Kawaihae, Makalapua, East Side, North Shore, Paakea and Palia)

Source of variation	d.f.	Sum of squares	Variance components	Percentage variation	*P*-value
Among groups	1	1416.861 (17.408)	9.701 (0.134)	4.08 (3.66)	0.14467 (0.06940)
Among pop. within groups	4	1636.393 (18.399)	7.752 (0.046)	3.26 (1.27)	0.01564* (0.12219)
Among individuals within pop.	174	38319.251 (603.755)	220.226 (3.470)	92.66 (95.07)	0.00000* (0.00000)*
Total	179	41372.506 (639.561)	237.679 (3.650)		

In this analysis, Kawaihae and Makalapua were grouped into the same group (Big Island) while East Side, North Shore, Paakea and Palia formed another group (Oahu). The analysis was carried out based on R_{ST} and F_{ST} values; results for the latter are shown in parentheses. Significant *P*-values, based on permutation tests in Arlequin v3.5.1.2, are indicated with asterisks.

Neutral genetic variation and population structure: An AMOVA analysis using R_{ST} demonstrated that differences among sites, rather than among islands, are responsible for much of the observed variation in allele frequencies (Table 2). Therefore, our subsequent analyses

were performed on the site scale. F_{ST} and R_{ST} analysis of site comparisons revealed moderate clustering based on island with the sites on Oahu differentiated from those on the Big Island (Table 3). The sites within Oahu were not significantly differentiated from one another except for one pairwise comparison (East Side and Palia). According to F_{ST} calculations, the Maui site was significantly different when compared to one of the Big Island sites (Kawaihae) and one of the Oahu sites (East Side). However, R_{ST} values for these comparisons were not significant.

Table 5.3. Pairwise R_{ST} and F_{ST} values between seven monarch buttterfly populations, as calculated in Arlequin v3.5.1.2

	Kawaihae	**Makalapua**	**Maui 377**	**East Side**	**North Shore**	**Paakea**
Makalapua	R_{ST}: 0.01886					
	F_{ST}: 0.00524					
Maui 377	R_{ST}: 0.02309	R_{ST}: 0.01747				
	F_{ST}: 0.06087*	F_{ST}: 0.03410				
East Side	R_{ST}: 0.07261*	R_{ST}: 0.11108*	R_{ST}: 0.04457			
	F_{ST}: 0.03481*	F_{ST}: 0.05370*	F_{ST}: 0.06896*			
North Shore	R_{ST}: 0.01119	R_{ST}: 0.04047*	R_{ST}: -0.01476	R_{ST}: 0.03341		
	F_{ST}: 0.04989*	F_{ST}: 0.02963*	F_{ST}: 0.00106	F_{ST}: 0.02524		
Paakea	R_{ST}: 0.11897*	R_{ST}: 0.10292*	R_{ST}: 0.00858	R_{ST}: 0.03400	R_{ST}: 0.00520	
	F_{ST}: 0.08489*	F_{ST}: 0.06238*	F_{ST}: 0.01235	F_{ST}: 0.03855	F_{ST}: -0.00344	
Palia	R_{ST}: 0.11233*	R_{ST}: 0.03965*	R_{ST}: 0.01722	R_{ST}: 0.08108*	R_{ST}: 0.03968	R_{ST}: 0.04376
	F_{ST}: 0.08310*	F_{ST}: 0.03434*	F_{ST}: 0.00090	F_{ST}: 0.05444*	F_{ST}: -0.01063	F_{ST}: 0.00153

Asterisks and shading denote values that are significantly different from zero. Note that all values are less than 0.12, and that significance is at the 0.05 level.

For thoroughness, we also looked at differentiation among islands and found similar results, with slight genetic differentiation detected between Oahu and the Big Island, as well as slight differentiation detected using F_{ST} between the Big Island and the other islands (Table S5). Thus, although there were some significant differences between sites, the observed levels of differentiation were low. This low level of genetic differentiation was confirmed with the analysis

in STRUCTURE, which did not indicate any significant population structure (Figure 3). The lack of genetic structure is unlikely to be an artifact of our microsatellite markers as they clearly detect genetic differentiation between Hawaii and New Zealand butterflies. We also ran STRUCTURE without the inclusion of New Zealand, and still found a lack of population structure among the Hawaiian sites (Figure S1). We performed a Mantel test to determine whether genetic differentiation correlated with differences in site collection time and found that the relationship was not significant for F_{ST} (r=0.04, P=0.54) or R_{ST} (r=0.22, P=0.43). This lack of a correlation indicates that differences in collection times are not responsible for the small amount of genetic variation found.

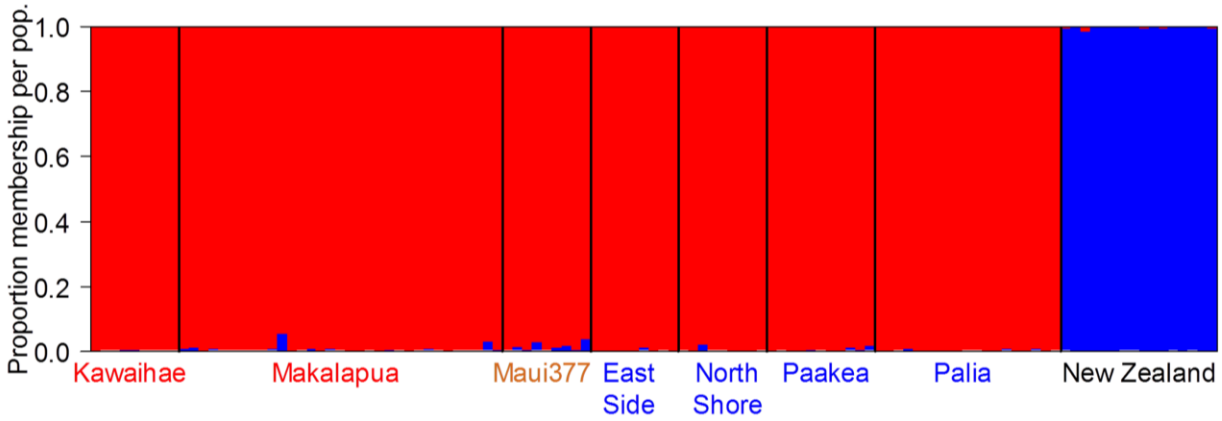

Figure 5.3. Structure plot showing that K (number of distinct populations) = 2. Monarchs on the Hawaiian Islands form one admixed genetic population (red text=Big Island, orange text=Maui, blue text=Oahu). New Zealand monarchs are differentiated from Hawaii and form their own genetic group.

Mean heterozygosity levels among sites ranged from a low of 0.333 in Palia to a high of 0.474 in North Shore (both of these sites are within Oahu) and did not significantly differ among sites within islands ($F_{6,98}$=0.65, P=0.69; Figure 4A). Allelic richness ranged from a low of 2.702 (Paakea, on Oahu) to a high of 3.266 (Makalapua, on the Big Island) but did not significantly differ among sites ($F_{6,98}$=0.44, P=0.85; Figure 4B). Genetic diversity ranged from 0.385 in Palia

to 0.522 in East Side (both of these sites are within Oahu) and did not differ significantly ($F_{6,98}$=0.66, P=0.68; Figure 4C).

Associations between genetic diversity and infection status. No site-level measures of genetic diversity (mean heterozygosity, P=0.22; allelic richness, P=0.30; genetic diversity, P=0.42) were found to correlate significantly with site-level averages of parasite prevalence (Figures 4D, 4E, 4F). At the individual level, average microsatellite heterozygosity was not found to predict infection status (χ^2= 126, d.f. = 1, P= 0.67).

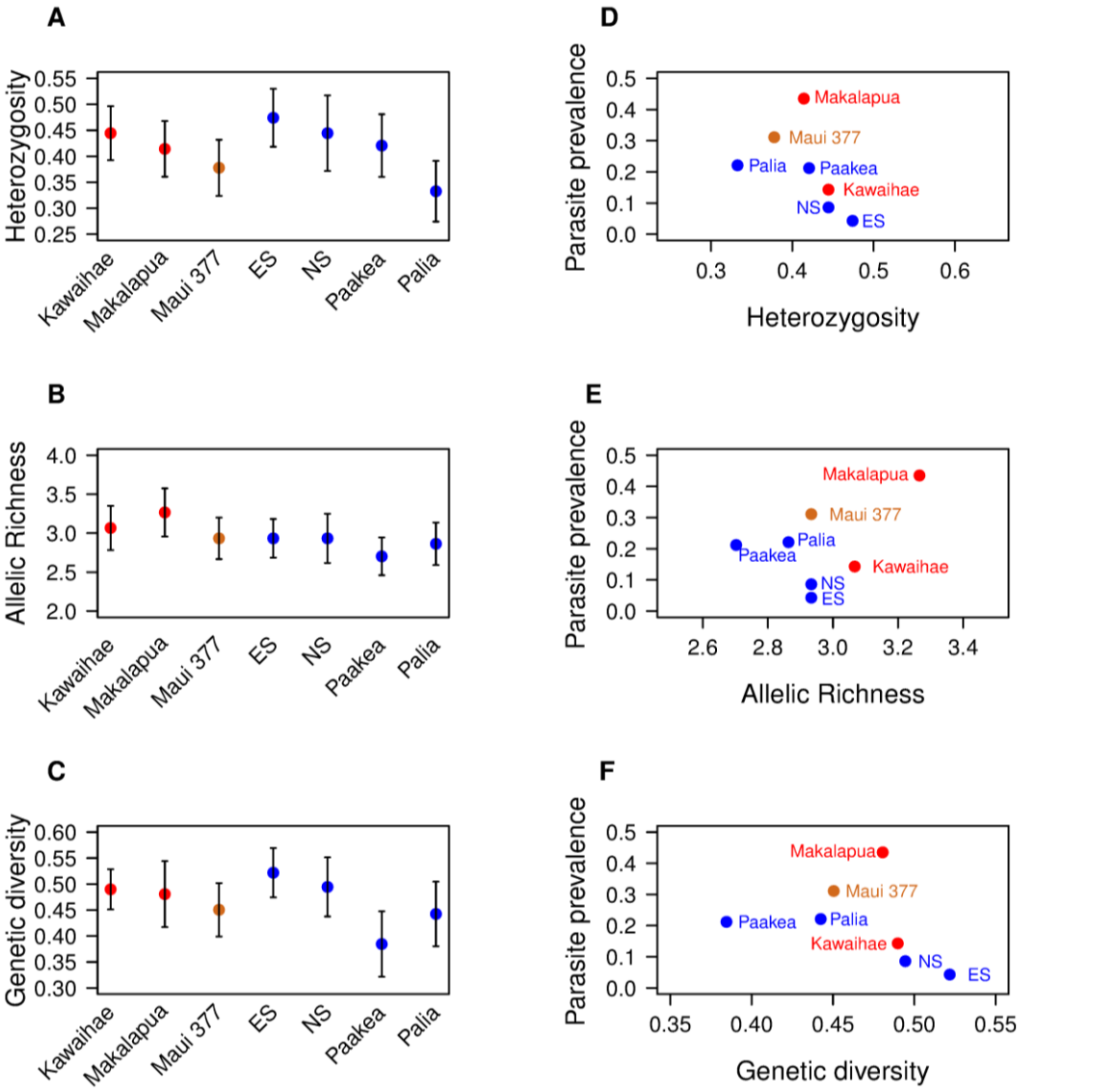

Figure 5.4. Measures of genetic diversity for monarchs from seven sites in Hawaii (red=Big Island, brown=Maui, blue=Oahu). ES refers to East Side and NS refers to North Shore, both of which are located on Oahu (**A**) Heterozygosity was found to be similar among the sites. (**B**) Allelic richness was similar amongst the sites. (**C**) Genetic diversity (using the value 1-Qinter, the inter-individual diversity within populations) was also found to be similar. (**D**) Heterozygosity did not correlate with parasite prevalence (r=-0.53; p=0.22). (**E**) Allelic richness was not found to correlate with parasite prevalence (r= 0.46; p=0.30). (**F**) Genetic diversity was not found to correlate with parasite prevalence (r= -0.36; p=0.42). Error bars in panels A-C show ±1 SE across loci.

Discussion

Parasite prevalence was highly variable among and within the Hawaiian Islands. These results are unexpected, because the non-migratory status of monarchs of the Hawaiian Islands would lead us to predict that parasite prevalence should be relatively high across all sites. Instead, patterns identified here suggest that factors other than migratory behavior can play a major role in driving heterogeneity in parasite prevalence in this system.

In this paper, we examined whether population sub-structuring might be responsible for the among-site variation in parasite prevalence. In particular, limited host movement might allow for local inbreeding and the loss of genetic diversity in sites with small populations, which is known to increase pathogen susceptibility in other systems (Acevedo-Whitehouse *et al.* 2009; Coltman *et al.* 1999; Smith *et al.* 2012; Whitehorn *et al.* 2011), and could more generally allow for the spatial segregation of host resistance alleles, leading to some sites with high resistance to infection and other sites with high susceptibility (Carlsson-Graner 2006; Carlsson-Graner & Thrall 2002). Although we found slight to moderate genetic differentiation among sites and islands when using R_{ST} and F_{ST} statistics, we found no evidence of population structure using the program STRUCTURE. Moreover, parasite prevalence was not explained by variation in genetic diversity, heterozygosity, and allelic richness among sites. Thus, it appears that population genetic variation cannot explain the observed heterogeneity in parasite prevalence in this

system. Instead, spatial environmental heterogeneity or ecological metapopulation processes might play stronger roles in determining infection heterogeneity in this host-parasite interaction.

Spatial variations in patch size, isolation and quality have been shown to alter parasite transmission and spatial spread in other host-pathogen systems. For example, empirical work demonstrated that landscape-level heterogeneity in habitat quality, host species diversity, and major geographic features such as water bodies affected the spatial spread and prevalence of pathogens ranging from rabies virus in raccoons to Lyme disease in white-footed mice (Allan *et al.* 2003; Almberg 2010; Real 2005; Smith *et al.* 2002). More generally, the dynamic structure of landscapes, particularly as a result of habitat fragmentation and other anthropogenic effects, can impact infectious diseases by affecting host species vital rates, density and distribution (Hess 2002). The Hawaiian Islands are known to vary in key ecological parameters such as total area, elevation, and human population density. Moreover, patches examined here differed in host plant species, patch size, and surrounding urban development. Although basic site level measurements collected here (Table S1) were not found to be correlated to infection prevalence, more comprehensive site-level data should be collected in the future, including actual numbers and distribution of host plants, elevation, temperature, precipitation, and monarch larval and adult densities. Host density in particular might correlate positively with parasite prevalence, as demonstrated by previous work on parasite infection in summer breeding North American monarchs (Bartel *et al.* 2011).

Metapopulation ecology could offer a different perspective for understanding how spatial processes cause infection heterogeneity in the monarch-pathogen interaction (Hanski & Gilpin 1991; Hess 1996, 2002). Specifically, extinction and colonization processes across interconnected patches might generate spatial variation in prevalence (even in the absence of other environmental gradients) simply because sites differ in the timing of host and pathogen

colonization (Gog *et al.* 2002; McCallum & Dobson 2002). Here, we considered the possibility that patch age might predict infection probability, if older habitat patches are more likely to be colonized by the pathogen. One specific prediction might be that patches with older monarchs (with more worn wings) should be more likely to harbor infected butterflies. Unfortunately, known pathogen effects on monarch wing characteristics make testing this idea challenging, because patches with more infected monarchs might have higher average wing wear measures simply because parasites negatively affect wing development. Indeed, at the individual level, infected monarchs sampled here had smaller wings and greater wing scale loss (but not greater wing damage); consistent with prior studies showing that infection lowers monarch wing area and body size (Altizer & Oberhauser 1999; de Roode *et al.* 2007) and reduces the density of black pigmentation on monarch wings (Lindsey & Altizer 2009). Thus additional studies examining host patch age could provide insight into the potential role of metapopulation ecology in this host-parasite dynamic.

In conclusion, we observed drastically varying prevalence of a protozoan parasite in monarchs inhabiting the Hawaiian Islands, despite high levels of butterfly gene flow and a lack of host population structure. The impact of site-level characteristics and landscape heterogeneity, in addition to colonization-extinction processes, are promising directions that could provide insight into the dynamics of this host-parasite interaction.

Supplementary Materials

Analysis for site-level characteristics and infection prevalence. To test whether site-level variables predicted infection prevalence, we calculated average prevalence per site per year, and used arcsin-square root-transformed values (to normalize the error variance) as the dependent variable in an analysis of variance. The following main effects were included in the model: Island, Year, Host Plant Species and Patch Type (treated as categorical variables), and Log_{10}-transformed Patch Area (treated as a continuous covariate). Each site was assigned a Patch Type (urban, suburban, rural) based on qualitative assessment of distance to nearest city, impervious surface and human activity. We measured Patch Area as the area of coverage of milkweed plants per site in m^2. Host plant species was represented per site as presented in Table S1. The analysis was weighted using sample size per year-site combination (to reflect better confidence in samples for which more monarchs were tested). Results showed that infection prevalence was not predicted by any independent variable included in the multivariate model, with significance tests as follows: LogArea ($F_{1,17} = 2.90$; $P = 0.11$); Island ($F_{3,17} = 1.52$; $P = 0.25$); Year ($F_{2,17} = 1.04$; $P = 0.38$); HostPlant ($F_{2,17} = 2.58$; $P = 0.11$); PatchType ($F_{2,17} = 1.52$; $P = 0.25$). The lack of significant effects can partly be explained by limited statistical power from low sample size (N = 28 site by year combinations) and by the high variation in prevalence among sites relative to any mean differences in levels of each independent variable.

Table S5.1. Field collection site variables, including site latitude and longitude, site type, site area and perimeter (based on the estimated area of the actual plants and not the entire patch), and milkweed species (*Asclepias physocarpa, Calotropis gigantea,* and *Calotropis procera*) were recorded. Only sites with 5 or more monarchs sampled are shown below, as sites with fewer monarchs were excluded from analyses.

Site	Island	Latitude	Longitude	Site Type	Area (m^2)	Perimeter (m)	Milkweed Species
Kailua-Kona	Big Island	19.64	-155.99	Urban	2535.78	203.35	*C. gigantea*
Kawaihae	Big Island	20.04	-155.82	Rural	15352.53	575.54	*C. gigantea*
Konacopia Farms	Big Island	19.46	-155.89	Rural	817.71	145.91	*C. gigantea*
Makalapua	Big Island	19.65	-156.00	Urban	297.15	119.17	*C. gigantea*
Kealia Beach	Kauai	22.09	-159.31	Rural	14903.32	496.41	*C. gigantea*
Kekaha Beach	Kauai	21.97	-159.73	Suburban	918.60	156.50	*C. procera*
Waimea	Kauai	21.96	-159.67	Urban	103623.17	1382.26	*C. gigantea*
West Waimea	Kauai	21.96	-159.69	Suburban	84.26	37.14	*C. procera*

Kihei	Maui	20.73	-156.45	Suburban	365.35	74.91	*C. gigantea*
Maui 377	Maui	20.83	-156.32	Rural	396.98	76.22	*A. physocarpa*
Maui-Lani	Maui	20.86	-156.48	Suburban	46.42	27.42	*C. procera*
East Side	Oahu	21.55	-157.85	Suburban	488.42	105.42	*C. gigantea*
Nehoa St.	Oahu	21.31	-157.83	Suburban	758.39	111.76	*C. gigantea*
North Shore	Oahu	21.68	-158.03	Suburban	56168.23	1034.63	*C. gigantea*
Palia St.	Oahu	21.40	-158.02	Suburban	225.27	60.91	*C. gigantea*
Paakea Rd.	Oahu	21.45	-158.18	Suburban	215	45.25	*C. gigantea*

Table S5.2. Monarchs used for genetic analysis by sampling site and sampling year.

	2007	**2009**	**2010**	**Total**	**Proportion Infected**
Big Island					
Kawaihae			9	9	0.11
Makalapua	11	5	17	33	0.55
Oahu					
East Side			9	9	0.11
North Shore			9	9	0.33
Paakea	11			11	0.00
Palia	7	7	5	19	0.63
Maui					
377			9	9	0.00

Table S5.3. Microsatellite loci used in this study, showing locus name, multiplex reaction, fluorescent label, primer sequences, repeat motif and primer annealing temperature (T_A). Number of alleles and allele size range were determined by Lyons et al (2012).

Locus	Multiplex reaction	Label	Primer sequence	Repeat motif	T_A	No. alleles	Allele size range
168	1	FAM	F: AGTTCAGGGTTTACGTGAGCA R: CATTATGTGAAGTGTTGCATGG	tcata	57°C	6	143-168
153	1	FAM	F: TGCGAAAASTGGTTTGAGGT R: TTATCGCCAAGTAAGTAATTTCG	ta	57°C	10	228-258
320	2	HEX	F: AATTTCTTGAGCGCTTTATCC R: CTGATCCTCGTCATCTCTCG	at	57°C	18	153-187
197	2	FAM	F: TGTCATTTCGATGTCGGCTA R: CAGAGAGAGCCTCGGGTAAA	att	57°C	4	174-183
208	3	FAM	F: TTTAGGACCCCAATCGGATTTTCG R: CGCGGACATTTTCACTTTCACGAT	at	60°C	19	178-242
203	3	HEX	F: TGACATACTTTATGTTCGTGGAAGG R: CCGCTCGCCTATATACAGGACACA	at	60°C	14	196-222
141	4	FAM	F: TCAAACCCGCATCCCTAGTGGTA R: TGGCAACGTACAGGGACGTGA	tc	60°C	13	150-178
1679	4	FAM	F: ATAGCCCTTCGACTTGTCGTTTCTC R:TCGACTGATGTTTTCGGGACTACGA	tat	60°C	4	215-224
137	5	HEX	F: AAGGTGGCGGTAAAAAGGCACAGA R: TCGCTTTCTTCCTCTTCCTCCTCA	aag	60°C	3	239-248
122	5	FAM	F: TTATAAGACCTCAACACCCACGAA R: CGCCGCTTCTAAATGAGTGGGATT	tta	60°C	6	228-252
494	6	HEX	F: CCGCGCTAGTCATTGTGTGAATGT R: CCTCGACTGATAGCCTTCGAAACG	att	60°C	7	160-181
983	6	FAM	F:AGACGCTTTGTTCAGCTTCGACCAC R: TTTACGATCACTCATACGAAACGGTA	ac	60°C	15	223-257
854	8	HEX	F:AACGTCATCTGCACACGCCATACTA R:TCCAATTAAACGTGACGCCATTTTG	at	67°C	8	230-254
165	8	FAM	F:CCTCcGGAACCTGTCAAGAAAAaGA R:CACTCATCAGAACTGAAAAGTTCGAGACC	tat	67°C	8	189-213
819	8	FAM	F:GACTCGGAGACATGAGATCGACGAC R:TCGTCAGACAATTGCTCAAAATGGA	cacga	67°C	11	213-263
519	9	FAM	F:GTGGCGGGGCTTTGTGTAAATAAGA R:CAGGGTTCCATACAAACGTGTGATACAATA	att	63°C	15	221-263

Table S5.4. Observed (H_o) and expected (H_e) heterozygosity at the seven Hawaiian sites at each locus as calculated by Arlequin 3.5.1.2

	Kawaihae		Makalapua		Maui 377		EastSide		NorthShore		Paakea		Palia	
Locus	H_0	H_e	H_0	H_e	H_0	H_e	H_0	H_e	H_0	H_e	H_0	H_e	H_0	H_e
168	0.778	0.542	0.424	0.516	0.778	0.529	0.667	0.582	0.889	0.739	0.909	0.671	0.474	0.650
153	0.667	0.627	0.364	0.553	0.333	0.386	0.444	0.627	0.333	0.464	0.273	0.385	0.053	0.235
320	0.444	0.673	0.667	0.723	0.222	0.627*	0.667	0.712	0.667	0.660	0.364	0.671	0.556	0.641*
197	0.444	0.471	0.424	0.429	0.333	0.307	0.444	0.471	0.444	0.366	0.455	0.368	0.222	0.203
208	0.444	0.542	0.438	0.546	0.556	0.569	0.778	0.699	0.556	0.699	0.455	0.541	0.211	0.568
203	0.444	0.529	0.594	0.604	0.556	0.712	0.556	0.634	0.444	0.752	0.364	0.498	0.789	0.596
141	0.444	0.556	0.485	0.502	0.444	0.399	0.778	0.778	0.889	0.667	0.455	0.593	0.474	0.538
1679	0.222	0.471	0.545	0.516	0.222	0.209	0.222	0.582	0.333	0.425	0.545	0.589	0.474	0.494
137	-	-	-	-	-	-	-	-	-	-	-	-	-	-
122	0.333	0.294	0.061	0.060	-	-	0.222	0.209	0.111	0.111	-	-	-	-
494	0.333	0.294	0.030	0.030	0.222	0.209	0.111	0.111	-	-	-	-	-	-
983	0.222	0.209	0.242	0.322	0.111	0.503	0.556	0.503	0.333	0.425	0.545	0.610	0.261	0.627
854	0.333	0.503	0.303	0.505*	0.444	0.399	0.333	0.425	0.000	0.366	0.300	0.268	0.389	0.417
165	0.111	0.307	0.273	0.373	0.333	0.542	0.222	0.366	0.333	0.503	0.400	0.505	0.167	0.475
819	0.778	0.569	0.727	0.669	0.444	0.569	0.444	0.503	0.667	0.627	0.700	0.595	0.500	0.422
519	0.667	0.719	0.636	0.729	0.667	0.654	0.667	0.582	0.667	0.569	0.545	0.515	0.421	0.632

Loci for which observed and expected heterozygosities are significantly different are indicated with asterisks; significance was determined using an α of 0.05 and a sequential Bonferroni correction. Loci for which at least four populations were in Hardy-Weinberg equilibirum and which were polymorphic were used for subsequent analyses and are indicated in shading. Dashes indicate monomorphic loci.

Table S5.5. Pairwise R_{ST} and F_{ST} values between four islands, as calculated in Arlequin version 3.5.1.2

	Big Island	Kauai	Maui
Kauai	R_{ST}: 0.00000		
	F_{ST}: 0.02133*		
Maui	R_{ST}: 0.01155	R_{ST}: 0.00000	
	F_{ST}: 0.03755*	F_{ST}: 0.00593	
Oahu	R_{ST}: 0.02054*	R_{ST}: 0.00000	R_{ST}: 0.00165
	F_{ST}: 0.04943*	F_{ST}: 0.01376*	F_{ST}: 0.00729

Asterisks and shading denote values that are significantly different from zero.

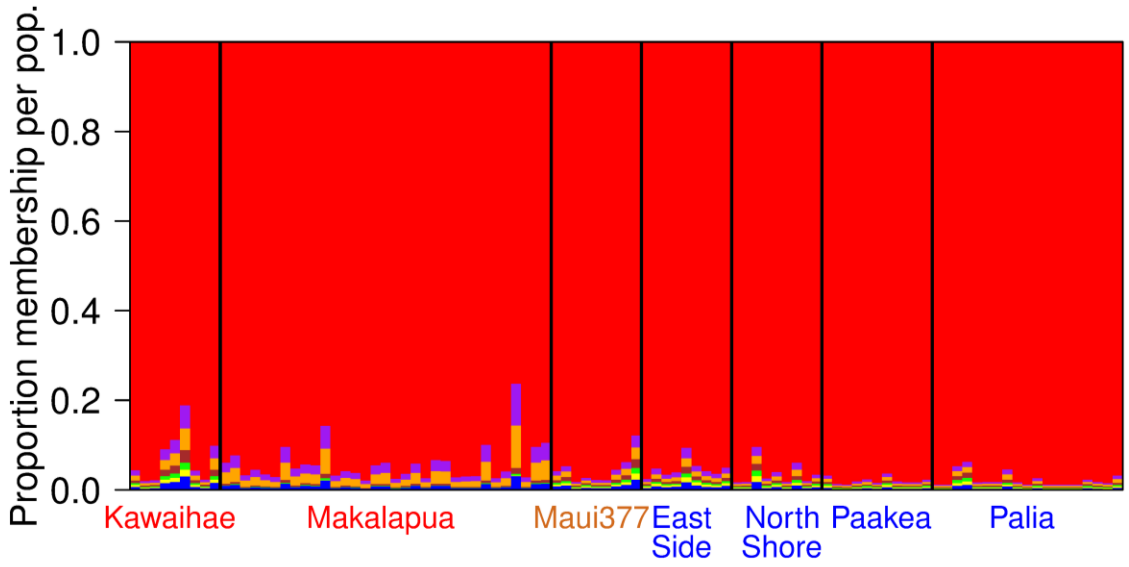

Figure S5.1. Structure plot showing that K (number of distinct populations) = 1. Monarchs on the Hawaiian Islands for one admixed genetic population (red text = Big Island, orange text = Maui, blue text = Oahu).

Chapter 6

Summary and Conclusions

6.1 Animal dispersal and its effects on population genetics

One of the aims of this dissertation was to determine the role of animal dispersal and range expansion in shaping population genetics. Chapter 4 addresses this by examining 18 worldwide populations of monarch butterflies. Monarchs are hypothesized to have expanded their range from North America to form stable colonies around the globe within the last 200 years (Vane-Wright 1993; Zalucki & Clarke 2004) and thus provide an excellent system to examine long-distance dispersal and its subsequent effects on population genetics.

By using microsatellites, I found that among 18 sampling sites, which were distributed as far west as Australia and as far east as the Iberian Peninsula, there are seven genetically distinct monarch populations. Additionally, it appeared the colonies across the Pacific Ocean were formed via a stepwise dispersal in which serial founder effects occurred. This was demonstrated by decreasing genetic diversity and allelic richness with increasing distance from the North American source population. Moreover, the frequencies of some alleles heavily declined with increasing distance from the source population while others were enriched, producing a clinal pattern across all loci, which is indicative of gene surfing, a spatial analog of genetic drift, playing a large role in shaping population genetics of these colonies. For colonies across the Atlantic Ocean, I found evidence supporting multiple colonization events shown by genetic diversity remaining relatively high. The colonies across the Atlantic likely also suffered from multiple bottlenecks due to excess larval density or destruction of host plants by flooding rivers (Fernandez Haeger *et al.* 2013). This was shown through the relatively high frequency of rare alleles in the colonies in the Iberian Peninsula and northern Africa. All of these results indicate that despite the relatively recent dispersal, I was still able to detect genetic differentiation and

population genetic trends. Furthermore, this genetic differentiation appeared to be the result of genetic drift through serial founder effects and bottlenecks.

Previous empirical studies have shown that genetic drift can be a strong force in determining allele frequencies after a range expansion, but these studies have been primarily limited to species with low-dispersal capabilities, like the spur-thighed tortoise in Spain or the bank vole in Ireland (Gracia *et al.* 2013; White *et al.* 2013). In these situations, we again saw genetic diversity loss and greater differentiation with increasing distance. However, theoretical studies have shown that these effects can be reduced or removed through long distance dispersal (Bialozyt *et al.* 2006; Fayard *et al.* 2009; Ray & Excoffier 2010). Some empirical studies have confirmed these predictions. In European Starlings inhabiting South Africa, for example, and peat mosses in the Stockholm archipelago, long-distance dispersal events have maintained genetic diversity (Berthouly-Salazar *et al.* 2013; Szovenyi *et al.* 2012) and multiple colonization events are hypothesized to have maintained relatively high genetic diversity in expanding populations of the Mediterranean damselfly (Swaegers *et al.* 2013). In contrast, in Chapter 4, I present one of the few studies that indicate that even in a species with high-dispersal capabilities, genetic drift can still be a leading force in shaping population genetics.

While it appears that genetic drift, rather than natural selection, has played the primary role in determining allele frequencies after this range expansion, I am limited from further understanding population differentiation by having examined only neutral sites. These sites indicate general patterns shown across the genome (assuming our microsatellite markers are distributed around the genome), but do not exclude differentiation or selection on certain genes or genomic regions. It is possible that local adaptation has also played a role in the genetic differentiation of these colonies, and that selection is acting on specific genes. For example, many of these locations differ in the milkweed host plant species present, as well as the prevalence and virulence of the parasite *Ophryocystis elektroscirrha* (Sternberg *et al.* 2012; Sternberg *et al.* 2013), both of which could result in selection on particular genes involved in

chemical processing and disease resistance. Additionally, while I found evidence for serial founder effects, it is possible that some monarchs have a higher affinity for dispersal and were more likely to be founders, thereby affecting subsequent allele frequencies. One candidate gene to investigate is the phosphoglucose isomerase (*Pgi*) gene; *Pgi* is a central enzyme in glycolysis and is involved in the flight and dispersal ability of monarchs and other butterflies (Hughes & Zalucki 1993; Niitepõld *et al.* 2009). Preliminary analyses have shown that migratory and non-migratory monarch populations from North and South America are nearly fixed for different haplotypes around *Pgi* (N. Chamberlain & M. Kronforst, unpublished data). It would therefore be interesting to examine if there are greater levels of the "migratory" *Pgi* haplotype in the non-migratory colonies than expected due to monarchs with a higher propensity for dispersal acting as founders.

6.2 Animal migration and its effects on population genetics

A second major aim of this dissertation was to examine the effects of animal migration on population genetics. This was addressed in Chapters 2, 3 and 5. In Chapter 2, I used microsatellites to examine two North American groups of migratory monarchs which travel to different locations to overwinter, with eastern North American monarchs overwintering in Mexico and western North American monarchs overwintering in California (Urquhart & Urquhart 1977, 1978). During the summer breeding season these eastern and western monarchs are separated by the Rocky Mountains and it has long been assumed that these two groups form distinct populations. However, even though these two groups are separated by a geographic barrier and overwinter in different countries, my work found that they are not genetically differentiated based on neutral genetic markers and are considered to form one large panmictic population. This work indicates that divergent migratory strategies are not dependent on whole genome differentiation. Instead, migratory differences may be driven by few genetically

differentiated loci or differences in gene expression subject to varying environmental conditions (Liedvogel *et al.* 2011).

In Chapter 3, I expanded the study to include non-migratory samples from sites in the Caribbean, South America, and locations in Central America that are south of the Mexican overwintering sites. With this I found that while locations south of the overwintering sites, such as Belize and Costa Rica, exhibited slight genetic differentiation, they were still considered part of the same panmictic population as eastern and western North American monarchs based on microsatellite markers. Again, while particular genes may differ between the migratory and non-migratory groups, this result indicates that sufficient gene flow is occurring to prevent population structure formation. A hypothesis as to how this occurs is that rather than a strict two way migration in which monarchs travel from the northern US and Canada to Mexico in the Fall, and then return to the north in the Spring, there is some portion of monarchs radiating outwards from the Mexican overwintering sites. This would result in the necessary gene flow to prevent the non-migratory monarchs in Belize and Costa Rica to significantly differentiate genetically.

As mentioned, though neutral markers do not distinguish between groups with various migratory strategies studied in Chapters 2 and 3, monarchs from these different locations could still be differentiated at loci that are under selection in their respective geographic areas. These monarchs could be differentiated at loci that influence traits such as flight ability or metabolism, where certain haplotypes would be favored in populations that migrate annually, but not in populations that do not migrate. Thus, it is possible that locally adapted variants of specific genes are selected for during the breeding and migration seasons despite a high influx of neutral genes via gene flow. Comparative genomics studies, which I address in the following section, offer a way to examine this.

Finally, in Chapter 5, I studied the effects of migration on population genetics from a disease dynamics perspective. Previous research has found that migratory populations of

monarchs have lower parasite prevalence than their non-migratory counterparts (Altizer *et al.* 2000; Bartel *et al.* 2011). Therefore, it is hypothesized that in locations where all individuals were equally non-migratory, one would expect to see equal levels of infection. However, I found that in monarchs among the Hawaiian Island chain, there is extreme heterogeneity in parasite prevalence despite all monarchs in this location being non-migratory. Using microsatellite markers, I found that despite the variation in parasite prevalence, there was no host population structure. Furthermore, I found no correlation between host neutral-site heterozygosity and infection status. It seems that in Hawaii, host-parasite dynamics may instead be driven by environmental factors, parasite genetics, or specific host genes linked to immunity rather than by host population structure. Each of these possibilities has been shown in many other host-parasite systems. For example, environmental factors, such as snail habitat and agricultural and sanitation practices, were found to significantly impact schistosomiasis in western China (Liang *et al.* 2007). Additionally, parasite genetics, like variation in copy number of a specific genic region in *Wolbachia*, was shown to alter virulence and affect host-parasite interactions (Chrostek & Teixeira 2015). Conversely, rather than genome-wide heterogeneity, specific host genes or regions, like the major histocompatibility complex in vertebrates, may be under selection (Tracy *et al.* 2015). Finally, parasite dynamics often result in variable levels of prevalence during different parts of an epidemic, both seasonally and over years. While there was some consistency over years in parasite prevalence in this study, more data are needed to determine how dynamic the system really is.

6.3 Future directions: from population genetics to population genomics

Through my dissertation work I have advanced our current knowledge in a number of ways. Specifically in the field of monarch butterfly research, I have dispelled the assumption that eastern and western North American monarchs form two distinct populations. I have also shown high levels of gene flow occurring between migratory North American monarchs and

non-migratory monarchs south of the Mexican overwintering sites, which challenges our view of a strict two-way migration. From a broader perspective, I demonstrated that even a species with a high capacity for dispersal is subject to the influences of genetic drift after a range expansion. I have also shown that divergent migratory strategies are not dependent on whole genome differentiation. However, to push this work and the field forward even further, additional experiments must be completed and next generation technologies can be used to allow us to gain a more complete picture.

The monarch butterfly genome has recently been sequenced (Zhan *et al.* 2011); however, the genetic basis of traits related to migration is still unknown. Previous genomics studies have shown differential expression of specific genes such as *turtle*, which affects locomotion (in fruit flies), and *rosy*, related to increased longevity in fruit flies, between summer-breeding and fall-migratory monarch butterflies (Zhu *et al.* 2009). Other genes have been identified, like those associated with the formation and efficient function of flight muscles (Zhan *et al.* 2014); however, we still lack a functional understanding of how migration genes ultimately enable the spectacular flight of several thousand kilometers (Ffrench-Constant 2014). Although the exact mechanisms by which eastern North American monarchs navigate to the overwintering sites remain a puzzle, monarch butterfly navigation involves a time-compensated sun compass, polarized light and circadian rhythms (Froy *et al.* 2003; Merlin *et al.* 2009; Reppert *et al.* 2004; Zhan *et al.* 2011; Zhu *et al.* 2008). Indeed, whether monarch butterflies are true navigators (meaning they are able to correct their directionality to reach a specific target), or instead employ vector navigation (meaning they simply fly in a southerly direction), is still hotly debated (Mouritsen *et al.* 2013a; Mouritsen *et al.* 2013b; Oberhauser *et al.* 2013).

Recent work on phenotypic variation across wild populations supports the idea that populations experiencing gene flow can continue to diverge at traits under selection. For example, a study examining wing morphology across multiple wild monarch populations showed that forewings were larger and more elongated in North American migratory monarchs

relative to non-migratory monarchs in South Florida (Altizer & Davis 2010); indeed, this study also found differences in wing morphology between eastern and western North American monarchs, which were not genetically differentiated based on neutral markers as shown in Chapter 2. These results suggest that migratory differences or traits linked to migration may instead be driven by few genetically differentiated loci or differences in gene expression subject to varying environmental conditions (Liedvogel *et al.* 2011).

With this in mind, a genome-wide analysis of genetic differences between groups with divergent migratory strategies might be the key to understanding migration genetics, and the published sequence of the full monarch genome (Zhan *et al.* 2011) makes this approach feasible. I am currently revisiting the eastern and western North American migratory groups, but rather than using microsatellite markers, I have performed whole genome resequencing. This approach allows me to investigate whether certain genes or genomic regions are under selection, or whether divergent migratory pathways are maintained despite a lack of genetic differentiation.

Such sequencing will provide a powerful way to uncover novel migration genes, and could confirm whether monarchs with different migration strategies have variations in genes that are involved in flight ability, such as *Pgi*, or genes involved in circadian rhythm, such as *cry1* and *cry2* (Zhu *et al.* 2008). I am also examining multiple overwintering sites along the California coast. While the eastern monarch overwintering sites are fairly centralized and result in large-scale genetic mixing (Eanes & Koehn 1978), California wintering sites are found along the coast hundreds of miles apart. It has been hypothesized that during the return migration north, some eastern monarchs fly in a north westerly direction and reach the western range, resulting in gene flow between the two migratory groups. However, due to the distribution of the California overwintering sites over a large area, it is possible that some western overwintering sites, specifically more southern ones closer to Mexico, receive a larger influx of eastern migrants than other sites (Brower & Pyle 2004). The comparative genomics approach will allow us to determine if gene flow between eastern and western monarchs are evenly distributed

throughout the western range or whether certain overwintering locations receive a higher influx of eastern migrants.

Additional research combining genome techniques and between-population crosses could also prove useful in identifying genetic regions associated with migration. For example, newly colonized monarch populations described in Chapter 4 are phenotypically distinct from the migratory population in terms of migratory status and wing shape; moreover, I have shown that these populations are differentiated at neutral genetic loci. Future projects can take advantage of the differences in migratory status of monarchs from these locations: in particular, the populations in Australia and Spain experience differing levels of gene flow with the North American migratory source population. A future project could be crossing monarchs from multiple non-migratory populations with monarchs from the migratory population from eastern North America and perform quantitative trait loci analyses to determine genetic regions affecting migratory traits. By including several non-migratory populations, this provides a unique opportunity to examine multiple instances of independent loss of migration. In addition, varying levels of connectivity between migratory and non-migratory monarchs allow investigators to determine the effects of differential gene flow on the genetics of migration loss.

The field of population genomics has rapidly expanded and this growth will continue with improving technologies. Recent years have seen the transition from the use of traditional population genetics approaches to the use of genomics and the investigation of genome-wide gene expression. The use of these techniques to address other important questions in ecology and evolution, like migration and dispersal behavior, will also open the door to new findings. The enticing aspect of genetic studies is that they not only can show us what is currently happening with a species, but can also give insights into their past as differing allele frequencies and genetic diversity levels provide glimpses into historical movement and evolutionary history. Undoubtedly, future work in population genomics, coupled with traditional studies and observation, will lead to exciting new breakthroughs.

Bibliography

Acevedo-Whitehouse K, Petetti L, Duignan P, Castinel A (2009) Hookworm infection, anaemia and genetic variability of the New Zealand sea lion. *Proceedings of the Royal Society B-Biological Sciences* **276**, 3523-3529.

Ackery PR, Vane-Wright RI (1984) *Milkweed Butterflies: their Cladistics and Biology* Cornell University Press, Ithaca, NY.

Åkesson S, Walinder G, Karlsson L, Ehnbom S (2001) Reed warbler orientation: initiation of nocturnal migratory flights in relation to visibility of celestial cues at dusk. *Animal Behaviour* **61**, 181-189.

Alerstam T (2006) Conflicting evidence about long-distance animal navigation. *Science* **313**, 791-794.

Alerstam T, Hedenstrom A, Akesson S (2003) Long-distance migration: evolution and determinants. *Oikos* **103**, 247-260.

Allan BF, Keesing F, Ostfeld RS (2003) Effect of forest fragmentation on Lyme disease risk. *Conservation Biology* **17**, 267-272.

Almberg ES, P. C. Cross, and D. W. Smith (2010) Persistence of canine distemper virus in the Greater Yellowstone Ecosystem's carnivore community. *Ecological Applications* **20**, 2058-2074.

Altizer S, Bartel R, Han BA (2011) Animal migration and infectious disease risk. *Science* **331**, 296-302.

Altizer S, Davis AK (2010) Populations of monarch butterflies with different migratory behaviors show divergence in wing morphology. *Evolution* **64**, 1018-1028.

Altizer S, De Roode JC (2013) Monarch defense against a debilitating parasite: resistance, immunity and self-medication. In: *Monarchs in a Changing World: Biology and*

Conservation of an Iconic Insect (eds. Oberhauser K, Altizer S, Nail KR), p. *In press*. Cornell University.

Altizer SM, K. Oberhauser, and K. A. Geurts (2004) Transmission of the protozoan parasite, Ophryocystis elektroscirrha, in monarch butterfly populations. In: *The Monarch Butterfly: Biology and Conservation* (ed. Solensky KOaM), pp. 203-218. Cornell University Press, Ithaca, New York.

Altizer SM, Oberhauser KS (1999) Effects of the protozoan parasite Ophryocystis elektroscirrha on the fitness of monarch butterflies (Danaus plexippus). *Journal of Invertebrate Pathology* **74**, 76-88.

Altizer SM, Oberhauser KS, Brower LP (2000) Associations between host migration and the prevalence of a protozoan parasite in natural populations of adult monarch butterflies. *Ecological Entomology* **25**, 125-139.

Austerlitz F, JungMuller B, Godelle B, Gouyon PH (1997) Evolution of coalescence times, genetic diversity and structure during colonization. *Theoretical Population Biology* **51**, 148-164.

Balloux F, Lugon-Moulin N (2002) The estimation of population differentiation with microsatellite markers. *Molecular Ecology* **11**, 155-165.

Bartel RA, Oberhauser KS, de Roode JC, Altizer SM (2011) Monarch butterfly migration and parasite transmission in eastern North America. *Ecology* **92**, 342-351.

Beerli P (2006) Comparison of Bayesian and maximum-likelihood inference of population genetic parameters. *Bioinformatics* **22**, 341-345.

Beerli P (2009) How to use migrate or why are markov chain monte carlo programs difficult to use? In: *Population Genetics for Animals Conservation* (eds. Bertorelle G, Bruford MW, Hauffe HC, Rizzoli A, Vernesi C), pp. 42-79. Cambridge University Press, Cambridge, UK.

Beerli P, Felsenstein J (2001) Maximum likelihood estimation of a migration matrix and effective population sizes in n subpopulations by using a coalescent approach. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 4563-4568.

Benhamou S, Sudre J, Bourjea J*, et al.* (2011) The role of geomagnetic cues in green turtle open sea navigation. *Plos One* **6**.

Bensch S, Akesson S, Irwin DE (2002) The use of AFLP to find an informative SNP: genetic differences across a migratory divide in willow warblers. *Molecular Ecology* **11**, 2359-2366.

Bensch S, Andersson T, Akesson S (1999) Morphological and molecular variation across a migratory divide in willow warblers, Phylloscopus trochilus. *Evolution* **53**, 1925-1935.

Berthouly-Salazar C, Hui C, Blackburn TM*, et al.* (2013) Long-distance dispersal maximizes evolutionary potential during rapid geographic range expansion. *Molecular Ecology* **22**, 5793-5804.

Best A, Webb S, White A, Boots M (2011) Host resistance and coevolution in spatially structured populations. *Proceedings of the Royal Society B-Biological Sciences* **278**, 2216-2222.

Bialozyt R, Ziegenhagen B, Petit RJ (2006) Contrasting effects of long distance seed dispersal on genetic diversity during range expansion. *Journal of Evolutionary Biology* **19**, 12-20.

Biek R, Henderson JC, Waller LA, Rupprecht CE, Real LA (2007) A high-resolution genetic signature of demographic and spatial expansion in epizootic rabies virus. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 7993-7998.

Birand A, Vose A, Gavrilets S (2012) Patterns of species ranges, speciation, and extinction. *American Naturalist* **179**, 1-21.

Bowlin MS, Bisson IA, Shamoun-Baranes J*, et al.* (2010) Grand challenges in migration biology. *Integrative and Comparative Biology* **50**, 261-279.

Bradley CA, Altizer S (2005) Parasites hinder monarch butterfly flight: implications for disease spread in migratory hosts. *Ecology Letters* **8**, 290-300.

Brower AVZ, Boyce TM (1991) Mitochondrial DNA variation in monarch butterflies. *Evolution* **45**, 1281-1286.

Brower AVZ, Jeansonne MM (2004) Geographical populations and "subspecies" of new world monarch butterflies (Nymphalidae) share a recent origin and are not phylogenetically distinct. *Annals of the Entomological Society of America* **97**, 519-523.

Brower LP (1995) Understanding and misunderstanding the migration of the monarch butterfly (Nymphalidae) in North America: 1857-1995. *Journal of the Lepidopterists' Society* **49**, 304-385.

Brower LP (1996) Monarch butterfly orientation: Missing pieces of a magnificent puzzle. *Journal of Experimental Biology* **199**, 93-103.

Brower LP, Fink LS, Brower AV*, et al.* (1995) On the dangers of interpopulational transfers of monarch butterflies. *Bioscience* **45**, 540-544.

Brower LP, Malcolm SB (1991) Animal migrations- endangered phenomena. *American Zoologist* **31**, 265-276.

Brower LP, Pyle RM (2004) The interchange of migratory monarchs between Mexico and the western United States, and the importance of floral corridors to the fall and spring migrations. In: *Conserving Migratory Pollinators and Nectar Corridors in Western North America* (ed. Nabhan GP), pp. 144-166. University of Arizona Press.

Buerkle CA (1999) The historical pattern of gene flow among migratory and nonmigratory populations of prairie warblers (Aves : parulinae). *Evolution* **53**, 1915-1924.

Carlsson-Graner U (2006) Disease dynamics, host specificity and pathogen persistence in isolated host populations. *Oikos* **112**, 174-184.

Carlsson-Graner U, Thrall PH (2002) The spatial distribution of plant populations, disease dynamics and evolution of resistance. *Oikos* **97**, 97-110.

Chrostek E, Teixeira L (2015) Mutualism Breakdown by Amplification of Wolbachia Genes. *Plos Biology* **13**.

Clarke AR, Zalucki MP (2004) Monarchs in Australia: on the winds of a storm? *Biological Invasions* **6**, 123-127.

Cockrell BJ, Malcolm SB, Brower LP (1993) Time, temperature, and latitudinal constraints on the annual recolonization of eastern North America by the monarch butterfly. *Natural History Museum of Los Angeles County Science Series* **0**, 233-251.

Coltman DW, Pilkington JG, Smith JA, Pemberton JM (1999) Parasite-mediated selection against inbred Soay sheep in a free-living, island population. *Evolution* **53**, 1259-1267.

Cronin JT (2009) Habitat edges, within-patch dispersion of hosts, and parasitoid oviposition behavior. *Ecology* **90**, 196-207.

Dallimer M, Jones PJ (2002) Migration orientation behaviour of the red-billed quelea *Quelea quelea*. *Journal of Avian Biology* **33**, 89-94.

Dallimer M, Jones PJ, Pemberton JM, Cheke RA (2003) Lack of genetic and plumage differentiation in the red-billed quelea *Quelea quelea* across a migratory divide in southern Africa. *Molecular Ecology* **12**, 345-353.

De Roode JC, Altizer S (2010) Host-parasite genetic interactions and virulence-transmission relationships in natural populations of monarch butterflies. *Evolution* **64**, 502-514.

de Roode JC, Chi J, Rarick RM, Altizer S (2009) Strength in numbers: high parasite burdens increase transmission of a protozoan parasite of monarch butterflies (Danaus plexippus). *Oecologia* **161**, 67-75.

de Roode JC, Gold LR, Altizer S (2007) Virulence determinants in a natural butterfly-parasite system. *Parasitology* **134**, 657-668.

De Roode JC, Yates AJ, Altizer S (2008) Virulence-transmission trade-offs and population divergence in virulence in a naturally occurring butterfly parasite. *Proceedings of the National Academy of Sciences, USA* **105**, 7489-7494.

Deshpande O, Batzoglou S, Feldman MW, Cavalli-Sforza LL (2009) A serial founder effect model for human settlement out of Africa. *Proceedings of the Royal Society B-Biological Sciences* **276**, 291-300.

Dingle H (1972) Migration strategies of insects. *Science* **175**, 1327-&.

Dingle H (1996) *Migration: The Biology of Life on the Move* Oxford University Press, Oxford.

Dingle H, Zalucki MP, Rochester WA (1999) Season-specific directional movement in migratory Australian butterflies. *Australian Journal of Entomology* **38**, 323-329.

Dingle H, Zalucki MP, Rochester WA, Armijo-Prewitt T (2005) Distribution of the monarch butterfly, *Danaus plexippus* (L.) (Lepidoptera : Nymphalidae), in western North America. *Biological Journal of the Linnean Society* **85**, 491-500.

Dobler S, Dalla S, Wagschal V, Agrawal AA (2012) Community-wide convergent evolution in insect adaptation to toxic cardenolides by substitutions in the Na,K-ATPase. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 13040-13045.

Dockx C (2007) Directional and stabilizing selection on wing size and shape in migrant and resident monarch butterflies, *Danaus plexippus* (L.), in Cuba. *Biological Journal of the Linnean Society* **92**, 605-616.

Dockx C (2012) Differences in phenotypic traits and migratory strategies between eastern North American monarch butterflies, *Danaus plexippus* (L.). *Biological Journal of the Linnean Society* **106**, 717-736.

Eanes WF, Koehn RK (1978) Analysis of genetic structure in the monarch butterfly, *Danaus plexippus* L. *Evolution* **32**, 784-797.

Ebert D, Haag C, Kirkpatrick M, *et al.* (2002) A selective advantage to immigrant genes in a Daphnia metapopulation. *Science* **295**, 485-488.

Eckert CG, Samis KE, Lougheed SC (2008) Genetic variation across species' geographical ranges: the central-marginal hypothesis and beyond. *Molecular Ecology* **17**, 1170-1188.

Edmonds CA, Lillie AS, Cavalli-Sforza LL (2004) Mutations arising in the wave front of an expanding population. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 975-979.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* **14**, 2611-2620.

Excoffier L, Foll M, Petit RJ (2009) Genetic consequences of range expansions. *Annual Review of Ecology, Evolution, and Systematics* **40**, 481-501.

Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* **10**, 564-567.

Excoffier L, Ray N (2008) Surfing during population expansions promotes genetic revolutions and structuration. *Trends in Ecology & Evolution* **23**, 347-351.

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes - application to human mitochondrial DNA restriction data. *Genetics* **131**, 479-491.

Fayard J, Klein EK, Lefevre F (2009) Long distance dispersal and the fate of a gene from the colonization front. *Journal of Evolutionary Biology* **22**, 2171-2182.

Fernandez Haeger J, Jordano Barbudo D (2009) The Monarch butterfly *Danaus plexippus* (Linnaeus, 1758) in the Strait of Gibraltar (Lepidoptera: Danaidae). *Shilap-Revista De Lepidopterologia* **37**, 421-438.

Fernandez Haeger J, Jordano D, Leon Melendez M (2011) Status and conservation of Asclepiadaceae and Danaus in southern Spain. *Journal of Insect Conservation* **15**, 361-365.

Fernandez Haeger J, Jordano D, Zalucki MP (2013) Monarchs across the Atlantic Ocean: what's happening on the other shore? In: *Monarhcs in a Changing World: Biology and*

Conservation of an Iconic Insect (eds. Oberhauser K, Altizer S, Nail KR), p. *In press*. Cornell University.

Ffrench-Constant RH (2014) Genomics: Of monarchs and migration. *Nature* **514**, 314-315.

Flaxman SM (2013) Surfing downhill: when should population range expansion be characterized by reductions in fitness? *Molecular Ecology* **22**, 5963-5965.

Folstad I, Nilssen AC, Halvorsen O, Andersen J (1991) Parasite avoidance: the cause of post-valing migrations in *Rangifer*? *Canadian Journal of Zoology* **69**, 2423-2429.

Francois O, Blum MGB, Jakobsson M, Rosenberg NA (2008) Demographic history of European populations of Arabidopsis thaliana. *Plos Genetics* **4**.

Fricke HC, Hencecroth J, Hoerner ME (2011) Lowland-upland migration of sauropod dinosaurs during the late Jurassic epoch. *Nature* **478**, doi:10.1038.

Froy O, Gotter AL, Casselman AL, Reppert SM (2003) Illuminating the circadian clock in monarch butterfly migration. *Science* **300**, 1303-1305.

Gog J, Woodroffe R, Swinton J (2002) Disease in endangered metapopulations: the importance of alternative hosts. *Proceedings of the Royal Society B-Biological Sciences* **269**, 671-676.

Gracia E, Botella F, Anadon JD*, et al.* (2013) Surfing in tortoises? Empirical signs of genetic structuring owing to range expansion. *Biology Letters* **9**, 20121091.

Guerra PA, Merlin C, Gegear RJ, Reppert SM (2012) Discordant timing between antennae disrupts sun compass orientation in migratory monarch butterflies. *Nature Communications* **3**.

Haig SM, GrattoTrevor CL, Mullins TD, Colwell MA (1997) Population identification of western hemisphere shorebirds throughout the annual cycle. *Molecular Ecology* **6**, 413-427.

Hallatschek O, Hersen P, Ramanathan S, Nelson DR (2007) Genetic drift at expanding frontiers promotes gene segregation. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 19926-19930.

Hallatschek O, Nelson DR (2008) Gene surfing in expanding populations. *Theoretical Population Biology* **73**, 158-170.

Hanski I, Gilpin M (1991) Metapopulation dynamics- brief history and conceptual domain. *Biological Journal of the Linnean Society* **42**, 3-16.

Hassell MP, Comins HN, May RM (1991) Spatial structure and chaos in insect population-dynamics. *Nature* **353**, 255-258.

Hess G (1996) Disease in metapopulation models: Implications for conservation. *Ecology* **77**, 1617-1632.

Hess G, S. Randolph, P. Arneberg, C. Chemini, C. Furlanello, J. Harwood, M. Roberts, and J. Swinton (2002) Spatial aspects of disease dynamics. In: *Ecology of Wildlife Diseases* (ed. P. Hudson AR, B. T. Grenfell, H. Heesterbeek, and A. P. Dobson), pp. 102-119. Oxford University Press, Oxford.

Holdo RM, Holt RD, Sinclair ARE, Godley BJ, Thirgood S (2011) Migration impacts on communites and ecosystems: empirical evidence and theoretical insights. In: *Animal Migrations: A Synthesis* (eds. Milner- Gulland EJ, Fryxell JM, Sinclair ARE), pp. 131-143. Oxford univeresity Press.

Holsinger KE, Weir BS (2009) Genetics in geographically structured populations: defining, estimating and interpreting F(ST). *Nature Reviews Genetics* **10**, 639-650.

Hood GM (2010) PopTools version 3.2.5. Available on the internet. *http://www.poptools.org*.

Hughes JM, Zalucki MP (1993) The relationship between the *Pgi* locus and the ability to fly at low temperatures in the monarch butterfly *Danaus plexippus*. *Biochemical Genetics* **31**, 521-532.

Irwin DE, Irwin JH (2005) Siberian migratory divides: the role of seasonal migration in speciation. In: *Birds of Two Worlds: The Ecology and Evolution of Migration* (eds. Greenberg R, Marra PP), pp. 27-40. Johns Hopkins University Press.

James DG (1993) Migration biology of the monarch butterfly in Australia. In: *Biology and conservation of the monarch butterfly* (eds. Malcolm SB, Zalucki MP), pp. 189-200. Natural History Museum of Los Angeles County, Los Angeles, CA.

Johnson TL, Cully JF, Collinge SK*, et al.* (2011) Spread of Plague Among Black-Tailed Prairie Dogs Is Associated With Colony Spatial Characteristics. *Journal of Wildlife Management* **75**, 357-368.

Johst K, Brandl R (1997) The effect of dispersal on local population dynamics. *Ecological Modelling* **104**, 87-101.

Klopfstein S, Currat M, Excoffier L (2006) The fate of mutations surfing on the wave of a range expansion. *Molecular Biology and Evolution* **23**, 482-490.

Kolbe JJ, Leal M, Schoener TW, Spiller DA, Losos JB (2012) Founder Effects Persist Despite Adaptive Differentiation: A Field Experiment with Lizards. *Science* **335**, 1086-1089.

Lenormand T (2002) Gene flow and the limits to natural selection. *Trends in Ecology & Evolution* **17**, 183-189.

Leong KLH, Yoshimura MA, Kaya HK (1997a) Occurrence of a neogregarine protozoan, Ophryocystis elektroscirrha McLaughlin and Myers, in populations of monarch and queen butterflies. *Pan-Pacific Entomologist* **73**, 49-51.

Leong KLH, Yoshimura MA, Kaya HK, Williams H (1997b) Instar susceptibility of the monarch butterfly (Danaus plexippus) to the neogregarine parasite, Ophryocystis elektroscirrha. *Journal of Invertebrate Pathology* **69**, 79-83.

Li JZ, Absher DM, Tang H*, et al.* (2008) Worldwide human relationships inferred from genome-wide patterns of variation. *Science* **319**, 1100-1104.

Liang S, Seto EYW, Remais JV*, et al.* (2007) Environmental effects on parasitic disease transmission exemplified by schistosomiasis in western China. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 7110-7115.

Liedvogel M, Akesson S, Bensch S (2011) The genetics of migration on the move. *Trends in Ecology & Evolution* **26**, 561-569.

Lindsey E, Altizer S (2009) Sex differences in immune defenses and response to parasitism in monarch butterflies. *Evolutionary Ecology* **23**, 607-620.

Lyons JI, Pierce AA, Barribeau SM*, et al.* (2012) Lack of genetic differentiation between monarch butterflies with divergent migration destinations. *Molecular Ecology* **21**, 3433-3444.

Malpica A, Ornelas JF (2014) Postglacial northward expansion and genetic differentiation between migratory and sedentary populations of the broad-tailed hummingbird (Selasphorus platycercus). *Molecular Ecology* **23**, 435-452.

McCallum H, Dobson A (2002) Disease, habitat fragmentation and conservation. *Proceedings of the Royal Society of London Series B-Biological Sciences* **269**, 2041-2049.

McKinnon L, Smith PA, Nol E*, et al.* (2010) Lower Predation Risk for Migratory Birds at High Latitudes. *Science* **327**, 326-327.

McLaughlin R, Myers J (1970) *Ophryocystis elektroscriffha* sp. n., a neogregarine pathogen of the monarch butterfly *Danaus plexippus* (L.) and the Florida queen butterfly *D. gilippus berenice* Cramer. *Journal of Protozoology* **17**, 300-&.

Merlin C, Gegear RJ, Reppert SM (2009) Antennal circadian clocks coordinate sun compass orientation in migratory monarch butterflies. *Science* **325**, 1700-1704.

Mettler R, Schaefer HM, Chernetsov N*, et al.* (2013) Contrasting Patterns of Genetic Differentiation among Blackcaps (Sylvia atricapilla) with Divergent Migratory Orientations in Europe. *Plos One* **8**.

Meyer CFJ, Kalko EKV, Kerth G (2009) Small-Scale Fragmentation Effects on Local Genetic Diversity in Two Phyllostomid Bats with Different Dispersal Abilities in Panama. *Biotropica* **41**, 95-102.

Miller NG, Wassenaar LI, Hobson KA, Norris DR (2012) Migratory Connectivity of the Monarch Butterfly (*Danaus plexippus*): Patterns of Spring Re-Colonization in Eastern North America. *Plos One* **7**, e31891.

MonarchWatch (2011) *http://www.monarchwatch.org/biology/westpop.htm*.

Mouritsen H, Derbyshire R, Stalleicken J*, et al.* (2013a) Reply to Oberhauser et al.: The experimental evidence clearly shows that monarch butterflies are almost certainly not true navigators. *Proceedings of the National Academy of Sciences of the United States of America* **110**, e3681.

Mouritsen H, Derbyshire R, Stalleicken J*, et al.* (2013b) An experimental displacement and over 50 years of tag-recoveries show that monarch butterflies are not true navigators. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 7348-7353.

Muller MJI, Neugeboren BI, Nelson DR, Murray AW (2014) Genetic drift opposes mutualism during spatial population expansion. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 1037-1042.

Nei M (1973) Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences of the United States of America* **70**, 3321-3323.

Neves VC, J.C. Fraga, H. Schafer, V. Vieira, A. Bivar de Sousa, and P.V. Borges (2001) The occurence of the monarch butterfly *Danaus plexippus* L. in the Azores, with a brief review of its biology. *Arquipelago - Life and Marine Sciences* **18A**, 17-24.

Niitepõld K, Smith AD, Osborne JL*, et al.* (2009) Flight metabolic rate and *Pgi* genotype influence butterfly dispersal rate in the field. *Ecology* **90**, 2223-2232.

O'Corry-Crowe GM, Suydam RS, Rosenberg A, Frost KJ, Dizon AE (1997) Phylogeography, population structure and dispersal patterns of the beluga whale *Delphinapterus leucas* in the western Nearctic revealed by mitochondrial DNA. *Molecular Ecology* **6**, 955-970.

Oberhauser KS, Taylor OR, Reppert SM*, et al.* (2013) Are monarch butterflies true navigators? The jury is still out. *Proceedings of the National Academy of Sciences of the United States of America* **110**, e3680.

Oksanen JF, Blanchet FG, Kindt R*, et al.* (2011) vegan: Community Ecology Package. R package version 2.0-0. *http://CRAN.R-project.org/package=vegan*.

Ostfeld RS, Glass GE, Keesing F (2005) Spatial epidemiology: an emerging (or re-emerging) discipline. *Trends in Ecology & Evolution* **20**, 328-336.

Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics* **28**, 2537-2539.

Peter BM, Slatkin M (2013) Detecting range expansions from genetic data. *Evolution* **67**, 3274-3289.

Petit E, Mayer F (2000) A population genetic analysis of migration: the case of the noctule bat (*Nyctalus noctula*). *Molecular Ecology* **9**, 683-690.

Pierce AA, Altizer S, Chamberlain NL, Kronforst MR, De Roode JC (2015) Unraveling the mysteries of monarch migration and global dispersal through molecular genetic techniques. In: *Monarchs in a changing world: Biology and conservation of an iconic butterfly* (eds. Oberhauser K, Nail KR, Altizer S). Cornell University Press, Ithaca, NY.

Pierce AA, de Roode JC, Altizer S, Bartel RA (2014a) Extreme Heterogeneity in Parasitism Despite Low Population Genetic Structure among Monarch Butterflies Inhabiting the Hawaiian Islands. *Plos One* **9**.

Pierce AA, Zalucki MP, Bangura M*, et al.* (2014b) Serial founder effects and genetic differentiation during worldwide range expansion of monarch butterflies. *Proceedings of the Royal Society B: Biological Sciences* **281**.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* **155**, 945-959.

Prugnolle F, Manica A, Balloux F (2005) Geography predicts neutral genetic diversity of human populations. *Current Biology* **15**, R159-R160.

Pusey A, Wolf M (1996) Inbreeding avoidance in animals. *Trends in Ecology & Evolution* **11**, 201-206.

Ramachandran S, Deshpande O, Roseman CC*, et al.* (2005) Support from the relationship of genetic and geographic distance in human populations for a serial founder effect originating in Africa. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 15942-15947.

Ravensdale M, Nemri A, Thrall PH, Ellis JG, Dodds PN (2011) Co-evolutionary interactions between host resistance and pathogen effector genes in flax rust disease. *Molecular Plant Pathology* **12**, 93-102.

Ray N, Excoffier L (2010) A first step towards inferring levels of long-distance dispersal during past expansions. *Molecular Ecology Resources* **10**, 902-914.

Raymond L, Plantegenest M, Gauffre B, Sarthou JP, Vialatte A (2013) Lack of Genetic Differentiation between Contrasted Overwintering Strategies of a Major Pest Predator Episyrphus balteatus (Diptera: Syrphidae): Implications for Biocontrol. *Plos One* **8**.

Real LA, and J. E. Childs (2005) Spatial-temporal dynamics of rabies in ecological communities. In: *Disease Ecology: Community Structure and Pathogen Dynamics* (ed. Ray SKCaC), pp. 168–185. Oxford University Press, Oxford.

Reppert SM, Gegear RJ, Merlin C (2010) Navigational mechanisms of migrating monarch butterflies. *Trends in Neurosciences* **33**, 399-406.

Reppert SM, Zhu HS, White RH (2004) Polarized light helps monarch butterflies navigate. *Current Biology* **14**, 155-158.

Rice WR (1989) Analyzing tables of statistical tests. *Evolution* **43**, 223-225.

Riley S (2007) Large-scale spatial-transmission models of infectious disease. *Science* **316**, 1298-1301.

Rousset F (2008) GENEPOP ' 007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources* **8**, 103-106.

Rozen S, Skaletsky HJ (2000) Primer3 on the WWW for general users and biologist programmers. In: *Bioinformatics Methods and Protocols: Methods in Molecular Biology* (eds. Krawetz S, Misener S), pp. 365-386. Humana Press, Totowa NJ.

Russell AL, Medellín RA, McCracken GF (2005) Genetic variation and migration in the Mexican free-tailed bat (*Tadarida brasiliensis mexicana*). *Molecular Ecology* **14**, 2207-2222.

Ryman N, Palm S (2006) POWSIM: a computer program for assessing statistical power when testing for genetic differentiation. *Molecular Ecology Notes* **6**, 600-602.

Schulte U, Veith M, Mingo V, Modica C, Hochkirch A (2013) Strong genetic differentiation due to multiple founder events during a recent range expansion of an introduced wall lizard population. *Biological Invasions* **15**, 2639-2649.

Servedio MR, Van Doorn GS, Kopp M, Frame AM, Nosil P (2011) Magic traits in speciation: 'magic' but not rare? *Trends in Ecology & Evolution* **26**, 389-397.

Shephard JM, Hughes JM, Zalucki MP (2002) Genetic differentiation between Australian and North American populations of the monarch butterfly *Danaus plexippus* (L.) (Lepidoptera : Nymphalidae): an exploration using allozyme electrophoresis. *Biological Journal of the Linnean Society* **75**, 437-452.

Slatkin M (1987) Gene flow and the geographic structure of natural populations. *Science* **236**, 787-792.

Slatkin M (1995) A measure of population subdivision based on microsatellite allele frequencies. *Genetics* **139**, 457-462.

Slatkin M, Excoffier L (2012) Serial Founder Effects During Range Expansion: A Spatial Analog of Genetic Drift. *Genetics* **191**, 171-181.

Smith DL, Lucey B, Waller LA, Childs JE, Real LA (2002) Predicting the spatial dynamics of rabies epidemics on heterogeneous landscapes. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 3668-3672.

Smith EM, Hoffman JI, Green LE, Amos W (2012) Preliminary association of microsatellite heterozygosity with footrot in domestic sheep. *Livestock Science* **143**, 293-299.

Solensky MJ, Oberhauser KS (2009) Sperm precedence in monarch butterflies (*Danaus plexippus*). *Behavioral Ecology* **20**, 328-334.

Sternberg ED, Lefevre T, Li J*, et al.* (2012) Food plant derived disease tolerance and resistance in a natural butterfly-plant-parasite interactions. *Evolution* **66**, 3367-3376.

Sternberg ED, Li H, Wang R, Gowler C, de Roode JC (2013) Patterns of Host-Parasite Adaptation in Three Populations of Monarch Butterflies Infected with a Naturally Occurring Protozoan Disease: Virulence, Resistance, and Tolerance. *American Naturalist* **182**, E235-E248.

Swaegers J, Mergeay J, Therry L*, et al.* (2013) Rapid range expansion increases genetic differentiation while causing limited reduction in genetic diversity in a damselfly. *Heredity* **111**, 422-429.

Szovenyi P, Sundberg S, Shaw AJ (2012) Long-distance dispersal and genetic structure of natural populations: an assessment of the inverse isolation hypothesis in peat mosses. *Molecular Ecology* **21**, 5461-5472.

Szpiech Z, Jakobsson M, Rosenberg N (2008) ADZE: a rarefaction approach for counting alleles private to combinations of. *Bioinformatics* **24**, 1367-4811.

Taberlet P, Fumagalli L, Wust-Saucy AG, Cosson JF (1998) Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology* **7**, 453-464.

Tamura K, Peterson D, Peterson N*, et al.* (2011) MEGA5: Molecular Evolutionary Genetics Analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution* **28**, 2731-2739.

Tanabe K, Mita T, Jombart T*, et al.* (2010) Plasmodium falciparum Accompanied the Human Expansion out of Africa. *Current Biology* **20**, 1283-1289.

Thomas CD, Bodsworth EJ, Wilson RJ*, et al.* (2001) Ecological and evolutionary processes at expanding range margins. *Nature* **411**, 577-581.

Thrall PH, Antonovics J (1995) Theoretical and empirical studies of metapopulations-population and genetic dynamics of the silene-usfilago system. *Canadian Journal of Botany-Revue Canadienne De Botanique* **73**, S1249-S1258.

Thrall PH, Burdon JJ (1997) Host-pathogen dynamics in a metapopulation context: the ecological and evolutionary consequences of being spatial. *Journal of Ecology* **85**, 743-753.

Tracy KE, Kiemnec-Tyburczy KM, DeWoody JA, Parra-Olea G, Zamudio KR (2015) Positive selection drives the evolution of a major histocompatibility complex gene in an endangered Mexican salamander species complex. *Immunogenetics* **67**, 323-335.

Tuskes PM, Brower LP (1978) Overwintering ecology of the monarch butterfly, *Danaus plexippus* L., in California. *Ecological Entomology* **3**, 141-153.

Urquhart FA (1976) Found at last: the monarch's winter home. *National Geographic*, 161-173.

Urquhart FA, Urquhart NR (1977) Overwintering areas and migratory routes of the monarch butterfly (*Danaus p. plexippus*, Lepidoptera: Danaidae) in North America, with special reference to the western population. *Canadian Entomologist* **109**, 1583-1589.

Urquhart FA, Urquhart NR (1978) Autumnal migration routes of the eastern population of the monarch butterfly (*Danaus p. plexippus* L.; Danaidae; Lepidoptera) in North America to the overwintering site in the Neovolcanic Plateau of Mexico. *Canadian Journal of Zoology* **56**, 1759-1764.

Urquhart FA, Urquhart NR (1979) Vernal migration of the monarch butterfly (*Danaus plexippus* Lepidoptera, Daniadae) in North America from the overwintering site in the neo-volcanic plateau of Mexico. *Canadian Entomologist* **111**, 15-18.

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* **4**, 535-538.

Vandenbosch R (2007) What do monarch population time series tell us about eastern and western population mixing? *Journal of the Lepidopterists' Society* **61**, 28-31.

Vandewoestijne S, Van Dyck H (2010) Population Genetic Differences along a Latitudinal Cline between Original and Recently Colonized Habitat in a Butterfly. *Plos One* **5**.

Vane-Wright RI (1993) The Columbus hypothesis: an explanation for the dramatic 19th century range expansion of the monarch butterfly. In: *Biology and Conservation of the Monarch Butterfly* (eds. Malcolm SB, Zalucki MP), pp. 179-187. Natural History Museum of Los Angeles County, Los Angeles.

Wenner AM, Harris AM (1993) Do California monarchs undergo long-distance directed migration? In: *Biology and conservation of the monarch butterfly* (eds. Macolm SB, Zalucki MP), pp. 209-218. Natural History Museum of Los Angeles County, Los Angeles, CA.

White TA, Perkins SE, Heckel G, Searle JB (2013) Adaptive evolution during an ongoing range expansion: the invasive bank vole (Myodes glareolus) in Ireland. *Molecular Ecology* **22**, 2971-2985.

Whitehorn PR, Tinsley MC, Brown MJF, Darvill B, Goulson D (2011) Genetic diversity, parasite prevalence and immunity in wild bumblebees. *Proceedings of the Royal Society B-Biological Sciences* **278**, 1195-1202.

Wolf JBW, Bayer T, Haubold B*, et al.* (2010) Nucleotide divergence vs. gene expression differentiation: comparative transcriptome sequencing in natural isolates from the carrion crow and its hybrid zone with the hooded crow. *Molecular Ecology* **19**, 162-175.

Zalucki MP, Clarke AR (2004) Monarchs across the Pacific: the Columbus hypothesis revisited. *Biological Journal of the Linnean Society* **82**, 111-121.

Zalucki MP, Hughes JM, Arthur JM, Carter PA (1993) Seasonal variation at four loci in a continuously breeding population of *Danaus plexippus* L. *Heredity* **70**, 205-213.

Zalucki MP, Kitching RL (1982) The analysis and description of movement in adult *Danaus plexippus* L (Lepidotera, Danaiae). *Behaviour* **80**, 174-198.

Zhan S, Merlin C, Boore JL, Reppert SM (2011) The monarch butterfly genome yields insights into long-distance migration. *Cell* **147**, 1171-1185.

Zhan S, Zhang W, Niitepold K, *et al.* (2014) The genetics of monarch butterfly migration and warning colouration. *Nature* **514**, 317-321.

Zhen Y, Aardema ML, Medina EM, Schumer M, Andolfatto P (2012) Parallel molecular evolution in an herbivore community. *Science* **337**, 1634-1637.

Zhu HS, Gegear RJ, Casselman A, Kanginakudru S, Reppert SM (2009) Defining behavioral and molecular differences between summer and migratory monarch butterflies. *Bmc Biology* **7**, doi: 10.1186/1741-7007-1187-1114.

Zhu HS, Sauman I, Yuan Q, *et al.* (2008) Cryptochromes define a novel circadian clock mechanism in monarch butterflies that may underlie sun compass navigation. *Plos Biology* **6**, 138-155.