Dispersal, Gene Flow, and Adaptive Evolution During Invasion: Testing Range-Limit Theory with the Asian Tiger Mosquito

2012

Kimberly Medley

University of Central Florida

Find similar works at: https://stars.library.ucf.edu/etd

University of Central Florida Libraries http://library.ucf.edu

Part of the Biology Commons

STARS Citation


This Doctoral Dissertation (Open Access) is brought to you for free and open access by STARS. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of STARS. For more information, please contact lee.dotson@ucf.edu.
DISPERSAL, GENE FLOW AND ADAPTIVE EVOLUTION DURING INVASION: TESTING RANGE-LIMIT THEORY WITH THE ASIAN TIGER MOSQUITO

by

KIM A. MEDLEY
A.B. Drury University, 1995
M.S. Missouri State University, 2004

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Biology in the College of Sciences at the University of Central Florida Orlando, Florida

Spring Term
2012

Major Professor: David G. Jenkins
ABSTRACT

Understanding the factors that make non-native species successful invaders is an important step towards mitigating spread. At the same time, species invasions can serve as natural experiments to test range-limit theory. Range-limit theory postulates declines in local abundance (abundant center model) and genetic diversity (central-peripheral hypothesis) towards range edges because of underlying environmental gradients. Such declines constrain adaptation to marginal habitats via gene swamping. However, broader evolutionary theory predicts intermediate rates of immigration into range-edge populations can relieve genetic drift and improve adaptive potential. I tested hypotheses generated from theory while illuminating aspects affecting of the invasion of the Asian tiger mosquito (Aedes albopictus Skuse) into the US. Using reciprocal distribution modeling, I found US populations occupied significantly different climate and habitat than in their native range (SE Asia). Most inconsistencies were found in the northern US range, where Ae. albopictus has recently crept northward, providing an opportunity to test range-limit theory as the range reaches its limit. Because of its limited natural dispersal ability, rapid spread after the 1985 US introduction pointed to human-aided dispersal. I tested the current role of human-aided versus natural dispersal using a landscape genetics framework, and found that natural dispersal dominated current patterns. Some distant localities were highly genetically similar, indicating potential human-aided transport in limited cases. Asymmetric gene flow from core to edge localities supported the abundant center model, but uniformly high genetic diversity contrasted with the central-marginal hypothesis. I detected a significant signature of local adaptation by overwintering diapause-induced eggs in multiple field sites using reciprocal transplants. Surprisingly, most genotypes from throughout the range produced large offspring when overwintered at the range edge. Relative offspring mass between home and away winters
peaked at an intermediate immigration rate. These results show that rapid adaptation has occurred in US populations of *Ae. albopictus* and highlight the potential for further spread. Genetic admixture from multiple introductions may explain high genetic diversity throughout the US range and contribute to high offspring size for all genotypes overwintered at the range edge. Finally, my work highlights the need for a better understanding of contemporary ecological and evolutionary processes leading to range-limits (or expansion) to more accurately reflect processes occurring in a human-dominated world.
ACKNOWLEDGMENTS

Over the years, my development as a scientist has been influenced by many people that I have collaborated with, been advised by, or have simply observed doing great science. Thus, the work presented in this dissertation has been influenced by peers and advisors at UCF, in addition to those I have worked with at other institutions during my time here and long before I arrived. I learned much through countless hours discussing experimental design and doing field work for a project conducted at MacArthur Agro-ecology Research Center, and organized by four very different, but very talented co-PIs: my advisor David G. Jenkins, Pedro F. Quintana-Ascencio, John E. Fauth, and Patrick Bohlen. It was a project for which I am very thankful to have been involved in, and I’m grateful to each co-PI for the wisdom they imparted to me throughout the project, either intentionally or by accident.

I am forever grateful to my advisor, David G. Jenkins, for the creative freedom he allowed me during my UCF career. He encouraged me to develop a research project that I would be proud of, and one that reflected the kinds of scientific questions that I felt needed to be pursued. It was through this process that I was able to unleash my potential not only as a researcher, but as a responsible member of the scientific community. I am also thankful to him for being supportive of my life beyond science, and the flexibility he allowed me to start a family while doing research. I always felt supported in both aspects of my life while working in Dave’s lab. It’s refreshing to see an academic that appreciates hard work but understands that a happy life requires balance.

I also want to thank Pedro Quintana-Ascencio, Eric Hoffman, and Phil Lounibos for their influence as members of my advisory committee. Pedro imparted not only statistical wisdom to
me, but offered philosophical considerations about scientific responsibility that I will always consider when interpreting and disseminating my work. Eric willingly accepted the challenge of advising the genetics aspects of my research, knowing full well I hadn’t picked up a pipette in years. Moreover, Eric consistently boosted my confidence by complimenting my work and my potential as a scientist, and provided great advice regarding the workings of the “business” of science. Phil was the mosquito expert on my committee, and I thank him for allowing me to play in his world even though I wasn’t invited. The quality of work conducted by Phil and his care as a scientist was great to observe, and I will take away from his involvement in my dissertation a constantly more cautious approach.

Many other members of the UCF community assisted with aspects of my research. Chris Parkinson provided lab space and equipment for processing DNA samples, and also provided useful advice on several topics. Bob Banks housed and cared for mice for feeding adult mosquitoes, and provided his expertise regarding handling and restraining mice (an area where I had no previous experience). Numerous undergraduates did many of the daily tasks involved in my research. Rebecca Francois helped extract DNA and weight pupae, and researched a great deal about arborival disease. Eddie Rysak spent many hours building mosquito cages and assisted with feedings and weighing pupae. And the best undergraduate crew at UCF weighed pupae: John Houder IV, Chris Grimont, and Nate Nafzger. Thanks fellas!

The extensive nature of my reciprocal transplant experiment required the assistance of people throughout the country that were willing to watch over a plastic container for a winter. These included Jeff Kimmons (Marshall, MO), Mary Lazar-Rutherford (Columbia, MO), Mark McCauley (St. Louis, MO), Jo McCauley (Springfield, IL), Jack Swanson (Peoria, IL), Mike Sinsko (Indianapolis, IN), Dr. James Joy (Marshall University), Mike Hutchinson (Harrisburg,
PA), Camille Collins (Raleigh, NC), Chris Evans (Columbia, SC), and Aunt Flossie and Uncle Bernie (Atlanta, GA).

Finally, I want to thank my family and friends for the support they have provided during the trying process of pursuing a doctoral degree. Lisa McCauley, a fellow Jenkins lab doctoral student, was able to keep it real and made life in the lab fun, but also served as a sounding board for ideas, design issues, and theoretical conversations. Without the constant support of my education by my parents, Jeff and Shirley Medley, I would not have understood so early on in my life that education is one key to a happy, successful, and fulfilling life. I thank them both for that and for emotional and financial support when times were tough. I thank Bri Kaiser for being an un-questioning and extremely supportive partner during the past six years in Florida. Without her constant hard work and support, I could not have accomplished a fraction of what I was able to accomplish while at UCF, nor would I have made the decision to pursue my doctorate without her encouragement. I also thank her for listening to me work through scientific problems, and providing those “duh” moments that made me realized when I was making things too complicated. Finally, I dedicate this dissertation to my son, Bryce Anthony. It was only through his eyes that I was able to gain perspective. He made all the hard work worthwhile, and made life so much richer and more fulfilling. I love you little man, and I can’t wait to see you appreciate our natural world. I hope I made this place a little bit better just for you.

*I sincerely thank you all.*
# TABLE OF CONTENTS

LIST OF FIGURES .............................................................................................................. 1

LIST OF TABLES ................................................................................................................... 5

CHAPTER 1: INTRODUCTION ............................................................................................... 8

  Range Limit Theory ........................................................................................................... 8

  Species Invasions as Tests of Range-Limit Theory ............................................................. 11

  Asian tiger mosquito, Aedes albopictus (Skuse) ................................................................ 12

  Research Questions and Hypotheses .................................................................................. 16

CHAPTER 2: NICHE SHIFTS DURING THE GLOBAL INVASION OF THE ASIAN TIGER MOSQUITO, Aedes albopictus Skuse (Culicidae) REVEALED BY RECIPROCAL DISTRIBUTION MODELS .................................................................................................................. 19

  Introduction ....................................................................................................................... 19

  Methods ............................................................................................................................. 21

    Species Occurrence Data ................................................................................................. 21

    Climate Layers ................................................................................................................ 22

    Modeling Approach ........................................................................................................ 23

  Results ............................................................................................................................... 27

  Discussion ......................................................................................................................... 37

  Conclusions ....................................................................................................................... 43

CHAPTER 3: LANDSCAPE GENETICS REVEAL UNEXPECTED POST-INVASION DISPERSAL PATTERNS FOR THE ASIAN TIGER MOSQUITO .................................................................................................................. 44
# Introduction

Methods .................................................................................................................. 48

- Study Species and Sampling ............................................................................. 48
- Lab Husbandry .................................................................................................. 48
- Genotyping .......................................................................................................... 49
- Range Expansion ................................................................................................. 50
- Genetic Distance ................................................................................................. 50
- Least-cost Paths and Resistance Analyses ......................................................... 51
- Model Selection and Parameter Estimation ....................................................... 53

Results .................................................................................................................... 55

- Genotypic Data .................................................................................................... 55
- Landscape Modeling ............................................................................................ 55

Discussion ............................................................................................................... 64

## CHAPTER 4: RAPID LOCAL ADAPTATION DURING INVASION DESPITE

ASYMMETRIC IMMIGRATION AT AN EXPANDING RANGE EDGE ...................... 71

Introduction ........................................................................................................... 71

Methods ................................................................................................................ 76

- Study species ...................................................................................................... 76
- Field collections ................................................................................................. 76
- Lab husbandry ................................................................................................. 77
- Experiment ......................................................................................................... 77
- Statistical analyses ............................................................................................ 79
LIST OF FIGURES

Figure 1. Map of expansion of *Ae. albopictus* from Houston, TX since 1986. Site of first established population indicated with black star. ................................................................. 18

Figure 2. Distributions based upon maximum entropy niche models using environmental data and point occurrences for *Aedes albopictus*. Color darkens on a ramp from low to high probability of occurrence (0–100%). White dots represent occurrences for *A. albopictus*. A: Native model generated with occurrences in the native range (Asia, Japan, Indonesia). B: Native model (Asia, Japan, Indonesia) projected onto South America. C: Model generated with occurrences in South America. D: South American model projected onto the native range. Arrows indicate the direction of model projections......................................................... 32

Figure 3. Distributions based upon maximum entropy niche models using environmental data and point occurrences for *Aedes albopictus*. Colour darkens on a ramp from low to high probability of occurrence (0–100%). White dots represent occurrences for A. albopictus. A: Native model generated with occurrences in the native range (Asia, Japan, Indonesia). B: Native model (Asia, Japan, Indonesia) projected onto North America. C: Model generated with occurrences in North America. D: North American model projected onto the native range. Arrows indicate the direction of model projections......................................................... 33

Figure 4. Distributions based upon maximum entropy niche models using environmental data and point occurrences for *Aedes albopictus*. Colour darkens on a ramp from low to high probability of occurrence (0–100%). White dots represent occurrences for *A. albopictus*. A: Native model generated with occurrences in the native range (Asia, Japan, Indonesia). B: Native model (Asia, Japan, Indonesia) projected onto Europe. C: Model generated with
occurrences in Europe. D: European model projected onto the native range. Arrows indicate the direction of model projections.

Figure 5. Principal components analysis (PCA) ordination of 18 environmental characteristics at known occurrences for *Aedes albopictus*. Symbols represent continents: black triangle, South America; grey triangle, North America; hollow circle, Asia; square, Europe. The three axes shown explain 62% of the variance for the PCA (axis 1 = 31%, axis 2 = 19%, axis 3 = 12%). Multi-response permutation procedures revealed that all paired comparisons were significantly different (P << 0.001).

Figure 6. Models from non-native distributions projected across the globe: South American (A), North American (B) and European (C). The darkest shading shows areas at risk of invasion from propagules originating from non-native distributions. The ellipse in (A) shows the approximate distribution for *Aedes albopictus* in Africa. The arrow in (B) shows a population postulated to have been founded from a temperate source (Lounibos et al., 2003). Maps are projected using the Robinson projection.

Figure 7. Sample localities in 26 US cities within the US range for *Ae. albopictus*. Range-edge localities represented by circles; range core localities represented by squares. Inset: location of study area. Darker shading indicates higher habitat suitability (Medley 2010). Inset: location of study area.

Figure 8. Linear relationship for geographic model (IBD) and for most informative models for genetic distance (Dps) for all locality pairs (n = 325), range-core localities (n = 48), and range-edge localities (n = 117). Dashed lines represent 95% C.I.

Figure 9. Hypothetical patterns of genotype by environment interactions on fitness. Lines represent genotypes (A, B) originating from different habitats (2, 1, respectively). In panel
A, local genotypes fare better than foreign genotypes, and genotypes fare better at home than away. In panel B, genotype A fares worse at home (habitat 2) than it does away, but still better than the foreign genotype. In Panel C, genotype A fares better at home than away, but genotype B still fares better under both conditions. In Panel D, genotype B fares better than the foreign genotype at home, but both genotypes fare worse in habitat 2; differences in fitness are due mostly to environment. Figure modified from Kawecki and Ebert (2004).

Figure 10. Sampling localities for range-edge (circles) and range-core (squares) populations. Red sites were not included as overwintering sites, but genotypes from those areas were overwintered elsewhere. Darker shading indicates higher habitat suitability (Medley 2010). Inset: location of study area.

Figure 11. Immigration from core to edge and from edge to core populations. Genes flow asymmetrically from core to edge localities. Error bars represent standard deviation; ANOVA was performed using 1/immigration as the response.

Figure 12. Average daily mean (top) and minimum (bottom) temperatures for over-wintering sites. Core sites: grey dashed lines; edge sites: black solid lines. Average values across all winter cities in core or edge regions shown with diamonds.

Figure 13. Average daily mean (left) and minimum (right) temperatures for core and edge winter sites. Both mean and minimum temperatures were significantly lower for edge sites (mean: $F_{1,9} = 44.96, p = 0.003$; min: $F_{1,9} = 34.11, p = 0.004$). Error bars represent standard deviation.

Figure 14. Female mean pupal mass by region. Home region shown inside bars. Ln (pupal mass) was significantly higher in edge winters than in core winters, and edge populations had a
significantly higher ln(pupal mass) in both over-wintering regions. Error bars represent standard deviation. ................................................................. 94

Figure 15. Mean pupal mass for local vs. foreign and home vs. away comparisons. Local genotypes had significantly higher (ln) pupal mass than foreigners in their home region. Genotypes had significantly higher (ln) pupal mass in their home region compared to the away region. Paired t-test on ln(pupal mass): local/foreign t = 94.91, home/away t = -112.77. Error bars represent standard deviation. ................................................................. 95

Figure 16. Immigration versus (ln) relative pupal mass for the local/foreigner comparison. Relative mass = pupal mass for female mosquitoes over-wintering in their home city/average female pupal mass for foreigners over-wintering in an away region. Values in the grey box indicate the local genotype has a higher pupal mass than foreign genotypes. Edge localities: hollow circles; core localities: filled circles................................................................. 96

Figure 17. Net gene flow versus (ln) relative pupal mass for the home/away comparison. Relative mass = average female pupal mass for a genotype over-wintering in their home region/female pupal mass for the same genotype in a locality in the away region. Values in the grey box indicate the genotype has a higher pupal mass at home than away. Edge localities: hollow circles; core localities: filled circles. Model for net gene flow vs. (ln) relative pupal mass: \( y = 0.79 + 2.02x + -22.25x^2; F = 5.59, p = 0.008 \)................................................................. 97
LIST OF TABLES

Table 1. Climate and environmental layers used in models. PCA eigenvalues are presented for all variables. Eigenvalues for the most important variables in the PCA in bold. Additional source information and citations are provided in the text. .......................................................... 29

Table 2. Model accuracy results using area under the curve (AUC) and binary tests of omission. AUC is a global assessment of model performance and values range from 0-1, where 1 is a perfect fit. A binomial omission test evaluated the rate of failure for test point prediction for each model. A significant binomial test indicates that the predicted omission rate was less than a random prediction. P values for all omission tests were less than 0.001. ................... 30

Table 3. Values for niche overlap (I) and statistics evaluating niche similarity and niche equivalency between distribution models. Significant values for niche similarity indicate that the two distributions are more similar than random. For niche equivalency, significant values indicate that the two distributions are not equivalent. All comparisons reveal the niche is more similar between native and introduced distributions than expected by chance, but they are not equivalent. For niche similarity, the first I-value reported in the pair represents the comparison between the native and introduced model projected onto the introduced distribution. The second value represents the comparison between the native and introduced model projected onto the native range. .......................................................... 31

Table 4. Geographic coordinates (decimal degrees), range location (Region), and sample size for all sampling localities. .......................................................... 57

Table 5. Cost surfaces for calculating least-cost path and resistance distances representing individual land cover types and combinations within categories. Combined surfaces (all water, all urban, all forest, and all farm) were created by assigning low costs or low
resistances to all individual land-cover types within that category (e.g. woody wetlands, herbaceous wetlands, and open water for the all water surface). Additional combined surfaces assigned relative costs to land cover types. Model 1 assigned low cost to all urban areas, wetlands, and forests and high costs to all other land cover; Model 2 assigned low cost to low-intensity urban areas, higher cost to medium- and high-intensity urban areas, and much higher costs to all other land-cover types; Model 3 assigned low cost to wetlands, open water, and forests (equally) and higher cost across urban areas, while all other land cover was assigned much higher cost; Model 4 assigned costs based upon model output from a suite of statistical models comparing least-cost path distances calculated from each of the single land-cover types to genetic distance.

Table 6. Pairwise cGD (bottom left triangular matrix) and Dps (top right triangular matrix) among localities within the core and along the edge of the US range for *Ae. albopictus*. Comparisons within regions (core or edge) are shaded.

Table 7. Two most informative models and the geographic distance model (representing isolation by distance) for genetic distance (Dps) between all locality pairs, core pairs, and edge pairs. Numerator (ANOVA) degrees of freedom, K = number of parameters, log (£) = log likelihood, AIC: Akaike’s Information Criterion, w_i = model weight, evidence = evidence ratio. The most informative model was chosen as the model with the lowest AIC and the highest model weight. Model parameters provided in Table 8. CD: cost distance along a least-cost path; R: resistance distance.

Table 8. ANOVA table and parameter estimates for most informative model for genetic distance (Dps) between all pairs of localities, core localities only, and edge localities only. Source of variance, df: degrees of freedom, SS: sum of squares, MS: mean square, F-ratio, β_0:
intercept, $\beta_1$: slope, $R^2$: coefficient of variation, $p_{\text{slope}}$: probability of an equal or greater slope of the regression line (10,000 permutations), CD: cost distance along least-cost path, R: resistance distance. ................................................................. 61

Table 9. Geographic coordinates (decimal degrees) and range location (Region) for all sampling localities and overwintering sites. ........................................................................................................ 85

Table 10. ANOVA table for 1/immigration. Significance indicates gene flow is asymmetric; immigration is higher from core to edge than from edge to core localities. $R^2 = 0.16$. ........ 85

Table 11. Three most informative models for each measure of pupal mass (Response).

Numerator (ANOVA) degrees of freedom, $K =$ number of parameters, log ($\ell$) = log likelihood, AIC: Akaike’s Information Criterion, $w_i =$ model weight, evidence = evidence ratio. Most informative model was chosen as the model with the lowest AIC and the highest model weight. Model parameters provided in Tables 12-14. ................................................. 86

Table 12. Most informative model for female pupal mass (ln-transformed). Multiple $R^2_{\text{adj}} = 0.39$, $F_{4,91} = 16.35$, $p < 0.001$. ........................................................................................................................................................................ 87

Table 13. Most informative model for relative pupal mass for the local/foreigner comparison.

Response = ln(female pupal mass in home city/average female pupal mass in that city for all genotypes from the foreign region). Multiple $R^2_{\text{adj}} = 0.63$, $F_{4,35} = 17.8$, $p << 0.001$. ............ 87

Table 14. Most informative model for relative female pupal mass for the home/away comparison. Response = ln(average female pupal mass for a genotype in its home region/mass for that genotype in the away region). Multiple $R^2_{\text{adj}} = 0.69$, $F_{6,33} = 12.06$, $p << 0.001$. min daily: difference in average minimum daily temperature over the experimental winter between a genotype’s home city and away city. ........................................................................ 88
CHAPTER 1: INTRODUCTION

Invasive species potentially threaten global biodiversity, cause economic loss, and contribute to disease spread (Lockwood, Hoopes, & Marchetti 2007; Lounibos 2002). Understanding the factors that make non-native species successful invaders is an important step towards mitigating spread. At the same time, species invasions can serve as natural experiments to illuminate processes leading to geographic range structure as their expanding ranges approach limits (Sax et al. 2007). In this dissertation, I aim to test aspects of range theory while improving knowledge on species invasions by studying a medically important invasive mosquito.

**Range Limit Theory**

Species theoretically reach their distributional limits through a combination of demographic and evolutionary mechanisms related to underlying environmental gradients. As the environment toward the range edge approaches a species’ physiological limits, population abundance declines compared to those in the range core, where environmental conditions are more hospitable. This demographic “abundant center model” predicts that population growth rates continue to decline as Allee effects and fragmentation increase towards the range edge (Andrewartha & Birch 1954; Brown, Mehlman, & Stevens 1995; Gaston 2003; Thomas & Kunin 1999). In extreme environments beyond the specie’s niche, populations can only be maintained with constant immigration.

Demographic clines theoretically influence evolutionary processes, further enforcing range limits. Smaller populations at the range periphery generally have lower genetic diversity,
owed primarily to founder events and genetic drift (the Central-Peripheral hypothesis; Hoffmann and Blows 1994). Additionally, higher-density populations nearer the range core (the Abundant Center Model) contribute more emigrants, so gene flow from central to peripheral populations becomes asymmetric (Kirkpatrick et al. 2008). High rates of asymmetric gene flow into peripheral populations swamp locally adaptive alleles by introducing abundant alleles adapted to conditions in source populations (Bridle & Vines 2006; García-Ramos & Kirkpatrick 1997; Kawecki & Holt 2002; Kirkpatrick & Barton 1997) and edge populations remain maladapted (Lenormand 2002). However, theory is only partially supported by evidence; 39% of range-structure studies in a meta-analysis supported the abundant center model (Sagarin & Gaines 2002). Weak support for theory occurred in part because some studies inadequately sampled species ranges, and often underrepresented range-edge populations. Thus, empirical support for the abundant center model remains equivocal, though range-limit theory continues to provide valid hypotheses for processes operating at range margins given appropriate conditions.

In contrast to range-limit theory, recent theoretical models have shown that gene flow can facilitate adaptation in harsh range-edge (i.e. sink) environments (e.g., Holt et al. 2003). When peripheral populations are small enough to be affected by genetic drift, immigration can provide the genetic variation and demographic stability necessary for adaptation to occur, providing an “evolutionary release”. In “black-hole” sink environments (i.e. no emigration), simulation models find that genetic variation and adaptive potential depend on immigration rate (Holt et al. 2003; Vucetich & Waite 2003). This result contrasts with analyses in which simulated immigration patterns that are uniform across a species’ range or that decrease from core to edge populations show a maladaptive response to immigration (Holt & Gomulkiewicz 1997; Kirkpatrick & Barton 1997). Temporal variation in the sink habitat can also facilitate adaptation
by extending periods of positive population growth (Holt, Barfield, & Gomulkiewicz 2004). Adaptation is also more likely to occur in sink environments when the optimal phenotype in the sink is not extremely different from the optimal phenotype in source environments (Holt et al. 2005).

Whether gene flow facilitates or constrains adaptation in peripheral environments has not been widely tested. However, recent reviews of the role of gene flow in local adaptation (Garant, Forde, & Hendry 2007; Kawecki & Ebert 2004; Parmesan et al. 2005) suggest that immigration facilitates adaptation by boosting local population size when initial population size is very low. Population persistence leads to adaptive divergence, which can indirectly influence gene flow. When population size is larger, immigration could exceed local carrying capacity, reducing fitness in the sink and constraining adaptation. Thus, intermediate levels of gene flow are ideal for adaptation to occur in peripheral populations.

In summary, theoretical and empirical research to date suggest range limits result from a series of demographic and evolutionary processes set in motion by environmental gradients. Populations at range edges are thus expected to be small, genetically depauperate, and potentially maladapted if they are swamped by genes adapted for range-center environments, but can potentially adapt given intermediate gene flow. Empirical tests of theory are difficult because evolutionary processes typically occur over long temporal scales. However, as invasive species expand and their ranges reach equilibrium, observations combined with experiments may test range-limit theory over contemporary time scales.
Species Invasions as Tests of Range-Limit Theory

Species invasions uniquely fill new ranges, and can provide opportunities to study range evolution in real time. To be viable models for testing range theory, evolution for invasive species must be observable. Theory predicts that traits beneficial to invasion can evolve rapidly given sufficient genetic variation (Jones & Gomulkiewicz 2012). Indeed, empirical evidence for rapid evolution during invasion is growing (Kinnison, Unvin, & Quinn 2008; Lambrinos 2004; Leger et al. 2009; Novak 2007; Rodríguez 2002; Sax et al. 2007; Willis, Memmott, & Forrester 2000). For instance, the cane toad invasion of Australia corresponded with greater hind leg length for toads at the invasion front, suggesting an adaptive advantage to colonizing new habitats (Phillips, Brown, et al. 2006). This work supported hypotheses generated from simulation models at expanding range edges (Phillips et al. 2008; Travis & Dytham 2002). More recently, reciprocal transplants with invasive cheatgrass in the US revealed patterns of local adaptation, although range-edge populations were not specifically adapted to edge habitat (Leger et al. 2009). However, genetic analyses revealed high levels of genetic variation at the range edge, and highlighted the potential for adaptation to occur and for range expansion to continue (Leger et al. 2009). Species invasions may include repeated introductions from multiple sources, which may contribute to high genetic variation and may require extensions of current range limit theory.

Other examples reveal responses to range expansion during climate change. In England, two butterfly species increased their niche breadth in order to track climate change (Thomas et al. 2001). Similarly, selection favored flight in a dimorphic cricket during a northward, climate-induced range expansion (Thomas et al. 2001). For the spotted knapweed (Centaurea maculosa)
invasion of western North America, Broennimann et al. (2007) detected a niche shift during invasion. Interestingly, a climate envelope model predicted sites of introduction successfully, but did not predict the expanded range now suitable as a likely result of adaptive evolution. Thus, species invasions can provide the opportunity to observe adaptation and to tease apart the mechanisms that facilitate adaptive evolution (Lambrinos 2004; Sax, Stachowicz, & Gaines 2005). I studied an ongoing species invasion to evaluate gene flow and adaptation at an expanding range edge. Research questions and hypotheses are described in detail below, following an introduction to the study organism.

**Asian tiger mosquito, Aedes albopictus (Skuse)**

The Asian tiger mosquito is an ideal organism for testing theories of adaptive evolution at a range edge. Its biology is well-known (Armbruster & Hutchinson 2002; Armbruster & Conn 2006; Costanzo, Mormann, & Juliano 2005; Gratz 2004; Hanson et al. 1996; Pumpuni, Knepler, & Craig Jr. 1992; Richards et al. 2006; Vitek & Livdahl 2006), its invasion history in the U.S. is well-documented (Benedict et al. 2007; Crans et al. 1996; Jamieson, Olson, & Wilhide 1994; Janousek, Plagge, & Kramer 2001; Moore 1999; Moore et al. 1988; Moore & Mitchell 1997; O’Meara et al. 1992; Sprenger & Wuthiranyagool 1986), and genetic work within the genus has resulted in a library of genetic markers for population genetic analyses (Behbahani et al. 2003; Porretta et al. 2006; Slotman et al. 2007).

Native to Southeast Asia, *Ae. albopictus* has invaded North and South America, Europe, and Africa since the late-1800s (Lounibos 2002) and is one of the 100 worst invasive species in the world (“Global Invasive Species Database” 2008). After its putative US introduction in
Houston, TX in 1985 (Hawley et al. 1987), it rapidly expanded: by 1986, it had already spread across 12 degrees latitude (Fig. 1) and has since coalesced into one large range. Recent surveillance indicates it is still expanding its range in the US (Bennett et al. 2005; Crans et al. 1996; Farajollahi & Nelder 2009; Powers et al. 2006), although some recent detections may not reflect established populations. Evolution of fitness-related traits could explain recent expansion for established populations.

Multiple studies with *Ae. albopictus* from geographically separate locations have evaluated life history differences among populations. Repeatedly, inter-population variation has been observed in various life-history traits, including egg desiccation resistance (Sota & Mogi 1992), photoperiod-induced egg diapause (Leisnham, Towler, & Juliano 2011; Lounibos, Escher, & Lourenço-de-Oliveira 2003; Pumpuni et al. 1992), larval growth (Armbruster & Conn 2006), and adult survival and reproductive output (Leisnham and Sala 2008). Among the different life-history stages, the strongest evidence for adaptive evolution has been for egg diapause (Leisnham et al. 2011; Urbanski et al. 2012). Diapause was detected in populations that established early in the US, leading researchers to conclude that temperate Asian populations were putative sources for the initial introduction in the U.S. (Hawley et al. 1987). Subsequent studies have shown evolution of this trait since initial introduction. Focks et al. (1994) detected a cline in critical photoperiod (day length at which 50% of eggs are diapausing) across the U.S. range for *Ae. albopictus*. Although the cline was shallower than in the native range, they postulated that it would likely steepen in another decade post-invasion. Using a “common garden” approach, Lounibos et al. (2003) reared mosquitoes from populations in the U.S. and Brazil under short-day photoperiods. They found a reduction in diapause response southward along the Florida peninsula and the first evidence of this trait in tropical populations in Brazil.
Development of a cline in diapause response in the US suggests evolution of the trait (Lounibos et al. 2003). Recent work by Leisnham et al. (2011) showed strong differences in diapause incidence between northern and southern US populations, supporting the evolutionary significance of this trait. These results are important to adaptation for *Ae. albopictus* because diapause is the life history trait that will enable populations to persist in regions that freeze and/or have extended dry seasons. Furthermore, traits related to photoperiod and circadian rhythms are theoretically critical for adaptation to seasonal climate regimes (Bradshaw & Holzapfel 2001b; a; Bradshaw, Zani, & Holzapfel 2004; Emerson, Bradshaw, & Holzapfel 2008). The current range edge for *Ae. albopictus* in the U.S. is defined by areas that are either relatively dry (e.g. western Texas and Oklahoma) or have a shorter growing season (e.g. Pennsylvania) than experienced in its native range, so successful populations at these edges should diapause as well.

Genetic reconstruction of the invasion history of *Ae. albopictus* indicates that sufficient genetic variation for adaptive evolution is present in U.S. populations. Studies have revealed high levels of genetic diversity among populations at allozyme loci (Black IV et al. 1988; Kambhampati et al. 1990) and non-coding nuclear DNA (Black IV, McLain, & Rai 1989), indicating several US populations were likely founded by multiple relatively large, independent introductions (Black IV et al. 1988). Multiple introductions have been implicated in the rapid spread of several invasive species because genetic admixture creates novel genotypes and increased genetic variation (Gillis et al. 2009; Kolbe et al. 2004; Marrs, Sforza, & Hufbauer 2008; Roman & Darling 2007). Evidence of multiple introductions for *Ae. albopictus* contradicts with genetic patterns of mitochondrial DNA (mtDNA; Birungi and Munstermann 2002) but patterns of mtDNA variation are confounded by the global infection of *Ae. albopictus* by
Wolbachia (Armbruster et al. 2003). Thus, evidence for multiple introductions based on mtDNA exists when accounting for Wolbachia infection (Khatchikian & Livdahl 2009). Furthermore, recent work comparing mtDNA variation with nuclear markers shows higher variation among US populations when using nuclear DNA (Khatchikian & Livdahl 2009).

Dispersal patterns for *Ae. albopictus* across the US range could provide ideal conditions for adaptation at the range periphery. In its natural environment, this mosquito is a relatively poor disperser, travelling ≤ 800 m from its natal habitat in a lifetime (Honório et al. 2003; Liew & Curtis 2004; Niebylski & Craig Jr. 1994). Population genetic analyses in other countries reveal high genetic differentiation due to its short dispersal distances (Ayres et al. 2002; Honório et al. 2003; Urbanelli et al. 2000). However, larvae have been detected in commerce and adults can travel in motorized vehicles (Enserink 2008; Phil Lounibos, pers. comm), suggesting that dispersal is inflated by human-aided movement. Thus, I expected the combination of natural and human-aided dispersal to effectively result in intermediate gene flow rates that would relieve drift, improve genetic variation, and lead to adaptation in range-edge populations (Garant et al. 2007). Despite evidence for sufficient genetic variation across the range, range-edge populations are expected to be subject to drift if they were founded from sources within the US range core. In addition to testing effects on adaptation, gene flow estimates can illuminate current rates of dispersal by analyzing emerging genetic patterns across the U.S. range.

Studying adaptation for *Ae. albopictus* also has direct applications for human health. The Asian tiger mosquito is a medically important invasive species in the United States because of its potential as a disease vector (Charrel, de Lamballerie, & Raoult 2007; Enserink 2008; Francy et al. 1990; Gratz 2004; Ibáñez-Bernal et al. 1997; Morens & Fauci 2008; Rai 1991; Rezza, Nicoletti, Angelini, Romi, Finarelli, Panning, Cordioli, Fortuna, Boros, Magurano, Silvi, et al. 2000).
2007; Shroyer 1986). Under lab conditions, *Ae. albopictus* is capable of vectoring at least 22 arboviruses and several viruses have been isolated from wild-caught individuals (Gratz 2004). In the US, West Nile virus (Turrell et al. 2001), Potosi virus (Francy et al. 1990), Cache Valley and LaCrosse virus (Mitchell et al. 1998), Eastern Equine Encephalitis (Niebylski et al. 1992), Tensaw virus and Keystone virus (Gratz 2004) have been detected. However, evidence for involvement of *Ae. albopictus* in US outbreaks is limited to LaCrosse virus (Gerhardt et al. 2001). *Ae. albopictus* was the main vector for an outbreak of dengue virus in Hawaii in 2001-2002 (Effler et al. 2005) and is the primary vector in central Africa, Bangladesh, south India, and parts of Thailand. Moreover, *Ae. albopictus* is responsible for outbreaks of the emerging chikungunya virus in Africa, the Indian Ocean, Asia, and Europe (Charrel et al. 2007; Delatte et al. 2008; Dubrulle et al. 2009; Rezza, Nicoletti, Angelini, Romi, Finarelli, Panning, Cordioli, Fortuna, Boros, Magurano, Silvi, et al. 2007). Recent genetic work revealed adaptive mutations within the virus led to a selective advantage for transmission by *Ae. albopictus* (DeLamballerie et al. 2008), and has elevated *Ae. albopictus* as a threatening vector for a potentially lethal disease. chikungunya has not been detected in mosquitoes in the US, but illuminating patterns of adaptation for *Ae. albopictus* during its US invasion serves to improve predictive ability for health agencies charged with assessing risk for disease spread globally.

**Research Questions and Hypotheses**

My overall hypothesis is that human-aided and natural dispersal combine to create intermediate gene flow rates that facilitate adaptation along the northern US range edge for *Ae.*
*albopictus*. Within the overall hypothesis, I addressed five questions organized into three chapters:

1) CHAPTER 2: Do niche-based distribution models for *Ae. albopictus* suggest US populations of *Ae. albopictus* occupy significantly different climate and habitat than native populations?

2) CHAPTER 3: Are dispersal patterns (as inferred by gene flow) explained better by human-aided dispersal, by natural means (i.e. flight), or by a combination of both?

3) CHAPTER 4: Has local adaptation occurred in US populations of *Ae. albopictus*?

4) CHAPTER 4: What is the relationship between gene flow and local adaptation, if detected?

5) CHAPTER 4: Are gene flow patterns and genetic diversity consistent with range-limit theory (abundant core model and central-peripheral hypothesis)?
Figure 1. Map of expansion of *Ae. albopictus* from Houston, TX since 1986. Site of first established population indicated with black star.
CHAPTER 2: NICHE SHIFTS DURING THE GLOBAL INVASION OF THE ASIAN TIGER MOSQUITO, Aedes albopictus Skuse (Culicidae) REVEALED BY RECIPROCAL DISTRIBUTION MODELS

Introduction

Invasive species potentially threaten global biodiversity, cause economic loss, and contribute to the spread of disease (Lockwood et al. 2007; Lounibos 2002). Understanding the factors that make non-native species successful invaders is an important step to managing geographic spread. Niche-based models predict species distributions based on occurrence points and climate and environmental datasets (Guisan & Thuiller 2005; Phillips, Anderson, & Schapire 2006), and have recently been extended to predict the geography of species invasions (Peterson 2003; Reed et al. 2008). A central assumption of such models is that the species’ niche is conserved during invasion.

Niche conservatism is the tendency for a species to retain ancestral niche characteristics. This property is useful for understanding historical biogeography, patterns of species diversity, community ecology and speciation, and has been assumed when predicting the spread of invasive species (Wiens & Graham 2005). While niche conservatism is likely the rule in several of these frameworks, the ability of non-native species to adapt to novel conditions could explain successful species invasion over contemporary time scales. If this is the case, niche-based models will not accurately predict the spread of invasive species, and characteristics of the niche for non-native distributions will not accurately predict the native range. Indeed, this “reciprocal

---

distribution modeling” (RDM) approach revealed a niche shift for spotted knapweed and fire ant invasions in the United States (Broennimann et al. 2007; Fitzpatrick et al. 2007). RDM proceeds by creating a model using native occurrences and projecting the model onto invaded regions to compare to the actual invasion. A second model using invasive occurrences is created and projected back onto the native distribution (Fitzpatrick et al. 2007). If the native model accurately predicts the introduced distribution and vice versa, the niche has been conserved. Discrepancies between models can facilitate hypothesis generation and elucidation of processes influencing successful invasions (Fitzpatrick et al. 2007).

One explanation for incongruence between model predictions is the “enemy release hypothesis” (ERH). ERH predicts that release from native biotic interactions could allow non-native species to colonize novel habitats, provided the invader is already adapted to local abiotic conditions (Keane & Crawley 2002; MacIsaac, Grigorovich, & Ricciardi 2001; Richardson et al. 2000). Thus, to best test ideas about niche conservation during invasion, species must be selected that minimize the likelihood of enemy release, such as a dominant competitor.

This study tests hypotheses using the invasion of the Asian tiger mosquito (Aedes albopectus Skuse). This species is native to Southeast Asia, and has invaded five continents in the last (ca.) 25 years. Experiments evaluating competitive interactions between U.S. populations of Ae. albopictus and the introduced Ae. aegypti have revealed Ae. albopictus consistently emerges as the superior competitor (Braks et al. 2004, Juliano and Lounibos 2005, and references therein). Furthermore, the native range of Ae. albopictus overlaps with other native and non-native mosquitoes in Southeast Asia. Thus, it is unlikely that competition constrains the extent of its native range.
In the current study, I used an RDM approach with maximum entropy models to test the null hypothesis that the niche for the Asian tiger mosquito (*Ae. albopictus*) has been conserved during its invasion of three continents. Niche conservation is indicated if the native model accurately predicts non-native distributions and non-native models accurately predict the native distribution. Alternatively, if reciprocal models poorly predict one another, a niche shift is indicated. To support these reciprocal comparisons, I used two recently developed metrics of niche overlap (equivalency vs. similarity, Warren et al. 2008) to compare RDMs. I also applied multivariate analyses to visualize and evaluate statistical differences in multi-dimensional climate and environmental characteristics among continental distributions. Finally, I evaluated the risk of range expansion in Southeast Asia using non-native model predictions of the native range, and highlight the potential for dispersal from invasive populations to other continents to increase the global invasiveness of the species.

**Methods**

**Species Occurrence Data**

All recorded occurrences of *Ae. albopictus* were compiled from literature (Aranda, Eritja, & Roiz 2006; Benedict et al. 2007; Bennett et al. 2005; Chadee, Hong Fat, & Persad 2003; Klobucar et al. 2006; Krueger & Hagen 2007; Rossi, Pascual, & Krsticevic 1999; Schaffner, Karch, & Culicidae 2000). Occurrences were represented as points (latitude and longitude) except in the U.S., where data were county-level. These data were converted to points by digitizing the centroid of each positive county in a GIS (ArcGIS, Environmental Systems Research Institute, Redlands, CA, USA). Literature reports of *Ae. albopictus* occurrence without
evidence of establishment were not included in the dataset. In total, 6599 occurrence points were compiled: 236 points for the native range (Asia, Indonesia, and Japan), 73 points for Europe, 1052 points for North America, and 5238 points for South America. To reduce sampling bias and autocorrelation of climatic data during model generation, I randomly selected 200 occurrence points from each of Asia, North America, and South America using Hawth’s Tools (http://www.spatalecology.com). All 73 points were used for Europe because sample locations were not as clustered as for other continents, and to improve balance in the number of data points between continents for multivariate analyses (see “Niche comparisons among continents” below). Comparable occurrence data were not available for Africa.

Climate Layers

Climatic and environmental datasets were chosen based upon their ability to affect the physiological ecology of *Ae. albopictus* (Pumpuni et al. 1992, Alto and Juliano 2001, Lounibos et al. 2002, e.g. Armbruster and Conn 2006a, Table 1). Variables that were highly correlated ($r \geq 0.70$) were excluded from the final models, resulting in 7 variables representing temperature and precipitation from the WORLDCLIM database (Hijmans et al. 2005), 8 layers representing relative humidity and photoperiod from the IWMI database (dw.iwmi.org), land cover (Global Land Cover Facility), and mean frost days and mean wind speed obtained from the CRU CL 2.0 database (Table 1, New et al. 2002). WORLDCLIM and land cover data were obtained at 5 arc-minute resolution and the remaining four layers were obtained at 10 arc-minute resolution. To maintain the highest resolution of the data, data at 10 arc-minutes were resampled to 5 arc-minute resolution in ArcGIS using bilinear interpolation, and were clipped to the same extent as the WORLDCLIM layers. All layers were exported as ASCII grids. Although the datasets
include both climate and local environmental data, the complete dataset will be referred to as “environmental data” hereafter for simplicity.

Modeling Approach

All models were developed using a maximum entropy algorithm implemented using Maxent software (Phillips, Anderson, et al. 2006; Phillips, Dudik, & Schapire 2004). Maximum entropy is a machine-learning technique that predicts species distributions using detailed climatic and environmental datasets together with species occurrence data, and generally performs better than other algorithms in tests of model performance (Elith et al. 2006; Ortega-Huerta & Peterson 2008; Phillips, Anderson, et al. 2006). Maximum entropy is more robust to spatial errors in occurrence data and uses presence-only datasets to predict the suitability of habitat (Graham et al. 2008; Phillips, Anderson, et al. 2006).

Reciprocal Models

I generated reciprocal models by first creating a model for the native range and a model for each invasion (South America, North America, and Europe) using occurrence points and environmental data clipped to the appropriate continent. I then projected the native model onto each invasive distribution and compared it to the model generated using occurrences from each invaded range. I also projected each invasive distribution onto the native range and compared them to the distribution generated with native occurrences. Finally, to explore areas at risk of invasion globally from propagules originating from invasive distributions, I projected each invasive model onto all continents. To improve the transferability of models across space, I used a regularization modifier of 1 (Phillips & Dudik 2008). Regularization reduces the likelihood of
overfitting models, thus increasing predictive ability of models beyond the training region (Phillips & Dudik 2008).

To evaluate model accuracy, I randomly selected 40% of occurrence points (from the set of 200 points selected for Asia, North America, and South America and from the full set of 73 points for Europe) to train each model and used the remainder to test each model using both binary tests of omission and analysis of the area under the curve (AUC) of the receiver operating characteristic (ROC) plot (Phillips, Anderson, et al. 2006). I calculated binary omission rates as the proportion of test points that were not predicted at a threshold probability that equaled the minimum probability of any pixel containing an occurrence point. AUC is a composite measure of model performance, and provides a global comparison of model fit to that of a random prediction. AUC values range 0-1, where 1 is a perfect fit. Useful models produce AUC values of 0.7-0.9, and excellent models produce AUC values above 0.9 (Swets 1988).

**Niche Comparisons Among Continents**

I used Principle Components Analysis (PCA) to compare environmental data among distributions using PC-ORD 5 (MjM Software, Gleneden Beach, Oregon, USA). I extracted environmental data at each occurrence point in a GIS (ArcGIS, Environmental Systems Research Institute, Redlands, CA, USA). To conform to the assumption of normality, I log-transformed Bio 3 and Bio 4 WORLDCLIM data and all relative humidity data from IWMI, and square-root transformed mean number of frost days. To assess significance, I compared variation explained by each PCA axis to axis variation obtained from 1000 PCAs conducted on matrices containing random values. A significant result indicates PCA axes are significantly better than a random configuration (McCune & Grace 2002). I used Multi-Response Permutation Procedures (MRPP)
in PC-ORD 5 to test the null hypothesis that environmental data at occurrence points in the native range for *Ae. albopictus* were no different than those data at occurrence points on other continents. MRPP is a non-parametric procedure that tests differences between groups, and is conceptually similar to ANOVA in that it assesses within-group similarity compared to among-group similarity (McCune & Grace 2002). Finally, I correlated each environmental variable with scores from the most important PCA axes to evaluate most important variable(s) in the PCA.

In addition to evaluating environmental differences among continents at occurrence points, I used two new approaches to evaluate niche characteristics between modeled distributions (Warren et al. 2008). Evaluating differences only at known occurrences biases environmental values towards sampling locations. Comparisons between entire distributions provide a broader estimate of the niche for each distribution, including potential occurrences within distributions. I used a permutation-based approach to evaluate niche similarity and niche equivalency between distributions. Both values compare niche overlap between a pair of real models to niche overlap between a real model and a model generated using either randomly generated occurrence points (niche similarity) or a randomly selected subset of the pooled occurrences for both distributions in the comparison (niche equivalency). For both comparisons, I calculated niche overlap between pairs of models using the metric $I$ (Warren et al. 2008), which ranges from 0 (no overlap) to 1 (identical). The metric compares probability values for individual pixels between two distributions. Thus, $I$ evaluates differences in potential occupancy predictions between two models. I provide specific methods for niche similarity and niche equivalency below.
**Niche Similarity**

Niche similarity compares models created with real occurrences to models created with randomly generated occurrences over the same geographic area (Warren et al. 2008). To evaluate niche similarity, I calculated $I$ between two models based upon real occurrences: one (focal) distribution generated with occurrences in that range (e.g., the native model) and another model generated with occurrences in another range (e.g. the South American model) that was projected onto the geographic space of the focal distribution. Then I created models with 200 randomly generated points in each distribution and replicated this 100 times in a script written in R 2.8.1 using the sp package. I calculated $I$ between each random model and the focal model, and compared the $I$ value calculated between real models to a histogram of $I$ values between the focal model and random models. A significant (one-tailed) test indicates niche overlap between real models is higher than niche overlap between real and random models.

**Niche Equivalency**

Niche equivalency compares models created with real occurrences to models created with occurrences randomly selected from real occurrences (Warren et al. 2008). For this study, the geographic extent of each distribution in the comparison was the combined size of two distributions. The metric $I$ was calculated in the same way as for niche similarity, except that permutations used a random subset of 200 of the actual pooled occurrence points for both distributions included in each calculation. The probability that distributions are significantly different is indicated by the calculated $I$ relative to a histogram of $I$ values calculated from random distributions. A significant (two-tailed) test indicates niche overlap between real models is not equivalent to niche overlap between real and random models.
Results

Models for individual distributions were all significantly better than random and predicted occurrences within the training region with low omission and high AUC values, indicating excellent model performance (Table 2). However, modeled distributions for one continent did not accurately predict other continents’ distributions. The native model (Fig. 2a) predicted a broad distribution for South America with high-risk areas in the southeast and northwest of the continent, and failed to predict many occurrences along the northeast coastal region and Amazon basin (Fig. 2b). Reciprocally, the South American model (Fig. 2c) predicted a portion of the native distribution, but also predicted high-risk areas in eastern India and northern Australia (Fig. 3d). For North America, the native model (Fig. 3a) predicted a large portion of actual occurrences (~80%, Fig. 3b), but failed to predict its northward and westward expansion (Fig. 3c). Rather, the native model shifted high-risk areas into Mexico and the Caribbean islands (Fig. 3b). Reciprocally, the North American model (Fig. 3c) predicted a very small native distribution (Fig. 3d) that was a subset of the actual native distribution and extended slightly northeast beyond the distribution predicted with the native model. Finally, the native model (Fig. 4a) predicted a European distribution along the border between Italy and France, Switzerland, and Germany, and areas along the eastern Adriatic coast (Fig. 4b), but failed to predict nearly all actual occurrences in Italy. The European model predicted much higher local probabilities of occurrence in Italy (Fig. 4c). When projected onto the native range, the European model predicted a wide band of high risk near the northern border of the native distribution (south-central China), and failed to accurately predict the distribution in south-east Asia (Fig. 4d).
Statistical comparisons between continents supported differences among modeled distributions. PCA ordination revealed three axes that were significantly different from random (p <<0.001), and indicated clear shifts in the niche space at occurrence points in the four distributions (Fig. 5). Variables that were strongly correlated (r > 0.70) with the first axis were annual mean temperature, mean minimum temperature for the coldest month, annual precipitation, precipitation of the wettest month, relative humidity during May, and number of frost days (Table 1). Mean maximum temperature for the warmest month and November photoperiod correlated most with the second axis (Table 1). MRPP analyses revealed environmental characteristics at occurrence points were significantly different between all pairs of continents. Tests for niche similarity revealed all native and introduced distributions were more similar than random (Table 3). However, for all comparisons, the niche for introduced distributions was not equivalent to the native niche (Table 3).

Projecting invasive distributions onto all continents revealed areas at risk of invasion and/or further spread (Figure 6). In addition to predicting high-risk areas in north Australia, the South American model predicted a broad area of high risk in Africa (Fig. 6a). The North American model predicts a high-risk area along the northeast coast of South America (Fig. 6b), and the European model predicts high-risk areas in northwestern South America where populations have not yet established (Fig. 6c).
Table 1. Climate and environmental layers used in models. PCA eigenvalues are presented for all variables. Eigenvalues for the most important variables in the PCA in bold. Additional source information and citations are provided in the text.

<table>
<thead>
<tr>
<th>Data source</th>
<th>Variable</th>
<th>Description</th>
<th>PCA Eigenvalues</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Axis 1</td>
</tr>
<tr>
<td>WORLDCLIM</td>
<td>BIO1</td>
<td>Annual mean temperature</td>
<td>-0.77</td>
</tr>
<tr>
<td></td>
<td>BIO5</td>
<td>Mean max temperature of the warmest month</td>
<td>-0.09</td>
</tr>
<tr>
<td></td>
<td>BIO6</td>
<td>Mean min temperature of the coldest month</td>
<td>-0.85</td>
</tr>
<tr>
<td></td>
<td>BIO12</td>
<td>Annual precipitation</td>
<td>-0.77</td>
</tr>
<tr>
<td></td>
<td>BIO13</td>
<td>Precipitation of the wettest month</td>
<td>-0.81</td>
</tr>
<tr>
<td></td>
<td>BIO14</td>
<td>Precipitation of the driest month</td>
<td>-0.07</td>
</tr>
<tr>
<td>IWMI</td>
<td>SUNFEB</td>
<td>% maximum sunlight hours during February</td>
<td>-0.10</td>
</tr>
<tr>
<td></td>
<td>SUNJUN</td>
<td>% maximum sunlight hours during June</td>
<td>0.50</td>
</tr>
<tr>
<td></td>
<td>SUNAUG</td>
<td>% maximum sunlight hours during August</td>
<td>0.53</td>
</tr>
<tr>
<td></td>
<td>SUNNOV</td>
<td>% maximum sunlight hours during November</td>
<td>-0.04</td>
</tr>
<tr>
<td></td>
<td>REHFEF</td>
<td>% relative humidity during February</td>
<td>-0.57</td>
</tr>
<tr>
<td></td>
<td>REHMAY</td>
<td>% relative humidity during May</td>
<td>-0.74</td>
</tr>
<tr>
<td></td>
<td>REHAUG</td>
<td>% relative humidity during August</td>
<td>-0.47</td>
</tr>
<tr>
<td></td>
<td>REHNOV</td>
<td>% relative humidity during November</td>
<td>-0.57</td>
</tr>
<tr>
<td>CRU Cl 2.0</td>
<td>FRS</td>
<td>Days with ground frost per month</td>
<td>0.77</td>
</tr>
<tr>
<td></td>
<td>WIND</td>
<td>Mean wind speed</td>
<td>0.54</td>
</tr>
<tr>
<td>GLCF</td>
<td>LANDCOV</td>
<td>Land cover</td>
<td>0.22</td>
</tr>
<tr>
<td>Eigenvalue</td>
<td></td>
<td></td>
<td>5.53</td>
</tr>
<tr>
<td>Percent variance</td>
<td></td>
<td></td>
<td>30.73</td>
</tr>
<tr>
<td>Cumulative percent variance</td>
<td></td>
<td></td>
<td>30.73</td>
</tr>
</tbody>
</table>
Table 2. Model accuracy results using area under the curve (AUC) and binary tests of omission. AUC is a global assessment of model performance and values range from 0-1, where 1 is a perfect fit. A binomial omission test evaluated the rate of failure for test point prediction for each model. A significant binomial test indicates that the predicted omission rate was less than a random prediction. P values for all omission tests were less than 0.001.

<table>
<thead>
<tr>
<th>Model</th>
<th>Omission rate (% of test points not predicted)</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Native (Southeast Asia)</td>
<td>4.9</td>
<td>0.989</td>
</tr>
<tr>
<td>South America</td>
<td>0.2</td>
<td>0.985</td>
</tr>
<tr>
<td>North America</td>
<td>0.5</td>
<td>0.993</td>
</tr>
<tr>
<td>Europe</td>
<td>0.0</td>
<td>0.998</td>
</tr>
</tbody>
</table>
Table 3. Values for niche overlap (I) and statistics evaluating niche similarity and niche equivalency between distribution models. Significant values for niche similarity indicate that the two distributions are more similar than random. For niche equivalency, significant values indicate that the two distributions are not equivalent. All comparisons reveal the niche is more similar between native and introduced distributions than expected by chance, but they are not equivalent. For niche similarity, the first I-value reported in the pair represents the comparison between the native and introduced model projected onto the introduced distribution. The second value represents the comparison between the native and introduced model projected onto the native range.

<table>
<thead>
<tr>
<th>Comparison</th>
<th>I</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Niche similarity</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Europe and Asia</td>
<td>0.50, 0.59</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>North America and Asia</td>
<td>0.66, 0.62</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>South America and Asia</td>
<td>0.72, 0.54</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td><strong>Niche equivalency</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>North America and Asia</td>
<td>0.39</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>South America and Asia</td>
<td>0.39</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Europe and Asia</td>
<td>0.48</td>
<td>&lt;0.01</td>
</tr>
</tbody>
</table>
Figure 2. Distributions based upon maximum entropy niche models using environmental data and point occurrences for *Aedes albopictus*. Color darkens on a ramp from low to high probability of occurrence (0–100%). White dots represent occurrences for *A. albopictus*. A: Native model generated with occurrences in the native range (Asia, Japan, Indonesia). B: Native model (Asia, Japan, Indonesia) projected onto South America. C: Model generated with occurrences in South America. D: South American model projected onto the native range. Arrows indicate the direction of model projections.
Figure 3. Distributions based upon maximum entropy niche models using environmental data and point occurrences for *Aedes albopictus*. Colour darkens on a ramp from low to high probability of occurrence (0–100%). White dots represent occurrences for *A. albopictus*. A: Native model generated with occurrences in the native range (Asia, Japan, Indonesia). B: Native model (Asia, Japan, Indonesia) projected onto North America. C: Model generated with occurrences in North America. D: North American model projected onto the native range. Arrows indicate the direction of model projections.
Figure 4. Distributions based upon maximum entropy niche models using environmental data and point occurrences for *Aedes albopictus*. Colour darkens on a ramp from low to high probability of occurrence (0–100%). White dots represent occurrences for *A. albopictus*. A: Native model generated with occurrences in the native range (Asia, Japan, Indonesia). B: Native model (Asia, Japan, Indonesia) projected onto Europe. C: Model generated with occurrences in Europe. D: European model projected onto the native range. Arrows indicate the direction of model projections.
Figure 5. Principal components analysis (PCA) ordination of 18 environmental characteristics at known occurrences for *Aedes albopictus*. Symbols represent continents: black triangle, South America; grey triangle, North America; hollow circle, Asia; square, Europe. The three axes shown explain 62% of the variance for the PCA (axis 1 = 31%, axis 2 = 19%, axis 3 = 12%). Multi-response permutation procedures revealed that all paired comparisons were significantly different (P << 0.001).
Figure 6. Models from non-native distributions projected across the globe: South American (A), North American (B) and European (C). The darkest shading shows areas at risk of invasion from propagules originating from non-native distributions. The ellipse in (A) shows the approximate distribution for *Aedes albopictus* in Africa. The arrow in (B) shows a population postulated to have been founded from a temperate source (Lounibos et al., 2003). Maps are projected using the Robinson projection.
Discussion

Niche shifts are apparent for three invasions (North and South American, Europe) of Ae. *albopictus*, based upon reciprocal niche-based distribution models, ordination, and measures of niche similarity and equivalency. Ordination techniques such as PCA are commonly employed to evaluate niche differences between distributions, but because this approach only considers environment at sampling points, information from other areas where populations likely exist (but aren’t sampled) is lost. In the current study, some error could have been introduced into the ordination (PCA) by extracting environmental data from the centroid of U.S. counties, but a more comprehensive comparison of the niche was possible by using newly developed permutation-based niche overlap statistics (Warren et al. 2008). These statistics revealed the niche for introduced distributions was more similar than expected by chance, but the niche for invasive distributions was not equivalent to the native niche. The niche between populations of the same species and even closely related species is expected to be similar (Wiens & Graham 2005), so the result that native and invasive populations of *Ae. albopictus* were more similar than random is not surprising. Niche divergence between native and introduced distributions revealed by niche equivalence statistics, however, leads to hypotheses explaining mechanisms for divergence during invasion.

Differences in niche characteristics among all current distributions of *Ae. albopictus* could result from a shift in the fundamental or realized niche, where the realized niche is a subset of the fundamental niche because of biotic interactions (Hutchinson 1957). By definition, occurrence points used to generate niche-based distribution models represent the realized niche
(Phillips, Anderson, et al. 2006), but sampling efforts can incorporate a broad range of environmental heterogeneity. Phillips et al. (Phillips, Anderson, et al. 2006) recommend that models be calibrated with occurrence data from a broad geographic extent so that they approximate the fundamental niche as much as possible. Because the native range for *Ae. albopictus* is quite broad, it should encompass sufficient environmental variation to predict invasive ranges well. However, patterns of biotic interactions that could constrict the niche throughout the native range are not known, although ecological experiments suggest limited effects of biotic interactions on *Ae. albopictus*. For instance, recent experiments reveal *Ae. albopictus* as a consistently dominant competitor, providing evidence that *Ae. albopictus* is not likely constrained by competition from other mosquitoes (Braks et al. 2004, Juliano and Lounibos 2005, and references therein). In addition, invasive populations of *Aedes aegypti* in the U.S. have contracted their range since the invasion of *Ae. albopictus* and now occur primarily in urban areas in the southeast U.S. (Juliano & Lounibos 2005). This pattern suggests competitive exclusion of *Ae. aegypti* by *Ae. albopictus*, though *Ae. aegypti* may exclude *Ae. albopictus* from heavily urban areas. Despite this latter possibility, exclusion from habitat patches does not constrain the extent of the U.S. distribution of *Ae. albopictus*. Therefore, enemy release seems unlikely for *Ae. albopictus*, and niche shifts appear to be in the fundamental niche rather than the realized niche, although verifying this hypothesis requires direct empirical evaluation.

Potential explanations for niche shifts are adaptive evolutionary changes or sampling of genetic material due to founder events. Several cases of adaptive evolution during species invasions have been recently documented. The cane toad invasion of Australia corresponded with greater hind leg length for toads at the invasion front, suggesting an adaptive advantage to colonizing new habitats (Phillips, Brown, et al. 2006). In England, climate change resulted in
range expansion and increased niche breadth for two butterfly species, and selected for flight in a
dimorphic cricket (Thomas et al. 2001). Other explanations for apparent niche shifts result from
founder events. Founding populations are a genetic subset of the source population, which likely
results in a loss of genetic diversity (Holgate 1966). This should result in lower fitness, but some
alleles that are lost during founder events could have previously masked the expression of other
alleles (i.e. epistasis). Thus, founder events during invasion could allow expression of beneficial
fitness-related alleles in novel environments (Blows & Hoffmann 2005). Furthermore, multiple
introductions could result in novel genetic combinations and allow invasive species to rapidly
adapt to different conditions (Kolbe et al. 2004).

Experimental and observational studies with *Ae. albopictus* support evolutionary changes
in the fundamental niche. In the U.S., northern populations of *Ae. albopictus* have refined the
ability to diapause and can withstand periods of severe cold and desiccation (Focks et al. 1994;
Hanson & Craig Jr. 1995). This is expected because evidence suggests northern Asian
populations as the source of the U.S. invasion and these populations are more likely to need
diapause to persist in the temperate climate (Hawley et al. 1987). However, populations in south
Florida have recently reduced the expression of diapause (Lounibos et al. 2003). Armbruster and
Conn (2006) detected geographic differences in larval growth rate partially due to temperature
on the northern edge of the U.S. range. In South America, genetic evidence points to non-
diapausing populations in Southeast Asia as the source of introduction (Kambhampati, Black, &
Rai 1991). Interestingly, the southernmost populations of *Ae. albopictus* in South America have
now developed diapause (Lounibos et al. 2003). It is unclear whether this is due to local
adaptive divergence or introduction of propagules from source populations that express diapause
(as discussed above).
A mechanistic understanding of apparent niche shifts observed here for *Ae. albopictus* requires empirical evaluation. For example, range margins can expand as a result of local adaptation by sink populations. Dispersal and gene flow may influence this process, though this relationship is equivocal (Garant et al. 2007). Experiments designed to detect adaptation over contemporary time scales are challenging, and understanding evolution in sink populations is difficult given the transient nature of such populations. Thus, much work in this area has involved simulation modeling (Holt et al. 2004, 2003), which provides a sound foundation for experimental tests of niche shifts.

If niche conservatism doesn’t apply to the spread of invasive species, niche conservation remains useful to predict the location of introduction. A study examining niche shifts in the spotted knapweed in the western U.S. showed models poorly predicted spread in the invasive range, but accurately predicted sites of introduction (Broennimann et al. 2007). The current study showed a similar result. The first recorded occurrence of *Ae. albopictus* in the U.S. was in Houston, Texas in 1985 (Hawley et al. 1987). This area was predicted by the native model, but the species has since spread into areas not predicted by this model. The first occurrence of *Ae. albopictus* in Europe was recorded in Albania in 1979, and was predicted by the native model (Adhami & Reiter 1998). Subsequent spread into Italy in 1990 (Sabatini et al. 1990), however, was not predicted by the native model. In South America, Rio de Janerio was likely the first location of establishment for *Ae. albopictus* (Lounibos 2002). Again, this location was predicted by the native model, but the subsequent northward spread of the invasion was not well-predicted.

The poor ability of niche-based distribution models to predict invasive distributions is not likely due to insufficiencies in the Maxent algorithm. This algorithm predicted the native range with high accuracy using native occurrences (Table 2), and many studies comparing this
algorithm to others consistently show Maxent is most accurate (Graham et al. 2008; Hernandez et al. 2006; Ortega-Huerta & Peterson 2008; Phillips, Anderson, et al. 2006). Rather, such reciprocal comparisons can be used to understand the process of invasion and broader questions about biogeography.

Projecting non-native models onto the native distribution provides information about the independent invasions and indicates risk for expansion of the native range. Projecting the South American model onto the native range suggests that initial propagules likely originated from Taiwan or northeast India and Burma/Myanmar. Based on model results, North American propagules likely originated from Indonesia, northern China or Japan. Japan or northern China had previously been considered sources for the initial introduction into Houston based on incidence of egg diapause in colonists (Hawley et al. 1987). Interestingly, high probability areas predicted in Indonesia by the North American model suggest that Indonesian populations may also have been a viable source of founding propagules. Projecting the European model onto Asia reveals a wide, high-probability band on the northern edge of the native distribution, suggesting that dispersal from Europe into Asia could cause a northward and westward expansion of the native range.

Projecting invasive models onto all continents reveals areas at risk of further expansion and alternate explanations for sources of invasions. High-risk areas predicted in Australia suggest that propagules from South America could establish in Australia. The South American model also predicts a large area of suitability from the center of the African continent southward. The current known extent of African populations of *Ae. albopictus* is rather narrow (Fig. 6a, circle), so it appears that propagules originating from South America could contribute to a southern expansion in Africa. To evaluate invasion into South America, Lounibos *et al.* (2003)
examined diapause incidence in South American populations of *Ae. albopictus*. They found that populations in São Luis, Brazil on the northeast coast had an unusually high diapause response compared to other South American populations, and pointed to an independent invasion from a temperate source. Interestingly, the North American model predicts a high-probability area near São Luis, suggesting that North America could have been the source of this local invasion (Fig. 6b, arrow). Finally, the high-risk area predicted in northwestern South America by the European model is currently unoccupied by *Ae. albopictus* (Fig. 6c). These comparisons collectively reveal that dispersal from native to non-native areas, subsequent niche shifts and dispersal from invasive distributions could serve as an “invasion ratchet” that broadens the overall niche and the potential for propagules to invade new areas globally. In support of this hypothesis, recent work has shown that models using both native and invasive occurrence points together to predict invasive species spread are generally more accurate at predicting the extent of spread and the pattern of risk for invasive distributions (Broennimann & Guisan 2008; Mau-crimmins, Schussman, & Geiger 2006).

These results show that successful invasions might best be explained by a combination of ecology (initial establishment) and evolutionary changes (spread) allowing species to occupy novel habitats and spread into new regions. Furthermore, these models illustrate the importance of monitoring both export of propagules from the native range and then from introduced ranges to prevent further expansion in the native range and globally.
Conclusions

The analyses herein resulted in two important conclusions: the niche for invasive populations of *Ae. albopictus* has shifted from its native state during invasion of North and South America and Europe, and propagules dispersing from invasive populations have the potential to increase the extent of global spread. These results are important in the broader context of predicting the spread of invasive species because niche-based distribution models may not be an appropriate tool for predicting patterns of spread. However, because the locations of introduction appear to be predicted accurately, niche-based models can be used to focus eradication efforts at ports of entry. Furthermore, these results highlight the importance of controlling transport of invasive species from non-native areas into other areas, as niche shifts in invasive areas and subsequent spread may lead to “invasion ratcheting”, a process that increases the global invasive potential for the species.
CHAPTER 3: LANDSCAPE GENETICS REVEAL UNEXPECTED POST-INVASION DISPERSAL PATTERNS FOR THE ASIAN TIGER MOSQUITO

Introduction

Species invasions can have negative effects on native communities, cause severe economic damage, and have important effects on human health (Mack et al. 2000; Vitousek et al. 1996). Understanding dispersal patterns for invasions is important to mitigate spread and to focus eradication efforts. Many contemporary introductions result from long-distance, human-aided propagule movement from exotic localities (Suarez, Holway, & Case 2000; Wilson et al. 2009). Altered long-distance dispersal can affect phylogeographic patterns (Lajbner, Linhart, & Kotlík 2011), invasion speed (Neubert & Caswell 2000), and invasion success (Wilson et al. 2009). Dispersal patterns for invasive species within non-native regions are well studied through theoretical models and empiricism (Hastings et al. 2005; Shigesada & Kawasaki 2001), but the role of anthropogenic factors affecting dispersal for invasive species is not often quantified in the new range. Moreover, because dispersal is linked to gene flow, human-altered dispersal patterns during invasion could have consequences to genetic structure and evolutionary trajectories (Slatkin 1985; Suarez & Tsutsui 2008).

Landscape genetics provides a framework to test multiple dispersal hypotheses and to evaluate the relative role of natural versus human-altered dispersal. Landscape genetics is an integrative approach that combines landscape ecology, population genetics, and geo-spatial statistics to identify spatial patterns of genetic structure (Manel et al. 2003; Storfer et al. 2007). Common questions pursued using this framework (Storfer et al. 2010) include identifying
barriers to gene flow (Guillot, Mortier, & Estoup 2005; Latch et al. 2008; Wofford & Gresswell 2005; Zalewski et al. 2009), inferring landscape effects on genetic diversity (Johansson et al. 2005), detecting dispersal corridors (Braunisch, Segelbacher, & Hirzel 2010), and identifying landscape factors influencing gene flow/connectivity (Holderegger & Wagner 2008; Murphy et al. 2010). For many studies, the intent is to evaluate fragmentation effects on genetic structure, particularly for conservation purposes (Hether & Hoffman 2012; Neel 2008; Radespiel, Rakotondravony, & Chikhi 2008; Spear & Storfer 2010). Few studies to date have used this framework to explicitly study human effects (Lada et al. 2008; Radespiel et al. 2008; Spear & Storfer 2010) or species invasions (Lecis et al. 2007; Zalewski et al. 2009), and none have done both. In this paper, we use landscape genetics to assess the relative explanatory power of human-aided (direct and indirect) and natural (e.g. flight) dispersal pathways for a species invasion.

Recent progress in landscape genetics incorporates emerging geo-spatial and statistical techniques to improve the robustness and broaden the applicability of model results. For instance, GIS-based methods to measure geographic distance have shifted from least-cost to resistance distances. Least-cost paths are single routes between two points that incur the least cumulative cost to travel, and their utility has been demonstrated in a number of studies (Wang et al. 2009). However, genetic distance between two localities is the result of movement by many individuals, and not all dispersing individuals follow identical paths. Genetic analysis based upon circuit theory was developed to incorporate multiple dispersal pathways between two points by calculating resistance distances (McRae 2006; McRae et al. 2008). This approach is theoretically a more realistic measure of geographic distance and is effective at improving the explanatory power of linear models between genetic and geographic distance relative to least-cost modeling (McRae & Beier 2007). Both least-cost and resistance methods rely on cost surfaces, where a
cost of traversing different land-cover types is assigned based upon field data (Adriaensen et al. 2003) or expert opinion (Broquet et al. 2006). Both approaches for cost assignment have received considerable criticism because they lack objectivity (Spear et al. 2010). However, multiple surfaces can be created representing competing dispersal hypotheses, and information-theoretic model-selection procedures can then compare models objectively (Burnham & Anderson 2001; Garroway, Bowman, & Wilson 2011; Shirk et al. 2010). This approach can be used to not only select the most informative resistance surface, but to compare the performance of geographic distance measures, including traditional isolation by distance (Wright 1943), and to compare the relative performance of groups of models together based upon broader hypotheses (e.g. human-aided versus natural dispersal). Commonly used tools for comparing genetic and landscape distances (i.e. Mantel and partial Mantel tests, Mantel 1967), do not lend themselves to model selection because they lack measures of model fit. Moreover, they have been shown to have inflated Type I error rates compared to other methods (Balkenhol, Waits, & Dezzani 2009). Multiple Regression on Distance Matrices (MRDM, Legendre and Legendre 1998, Lichstein 2006) is a good approach for modeling the relationship between multiple matrices, and the output lends itself to model selection without inflated Type I error.

Here, I present landscape genetics analyses for a highly mobile, flying species. Taxonomic foci of landscape genetics studies have been generally biased toward vertebrates (Storfer et al. 2010), and mostly amphibians (Murphy et al. 2010; Spear & Storfer 2010). This is for good reason; many amphibians are of conservation concern and have relatively low vagility, leading to marked population genetic structure and strong effects of landscape (Hether & Hoffman 2012; May et al. 2011; Murphy et al. 2010). Highly vagile species (e.g. flying insects) are more difficult to study in this context because genetic structure is often less pronounced. The
A relationship between landscape features and flight have not been well-studied to date (Rasic & Keyghobadi 2011), but is an area with great potential. For instance, three-dimensional habitat can affect flight ability between localities by providing refuge from predators and wind. Wind-mediated dispersal can also be detected in this context, as open areas are more likely to correlate with genetic distance for wind-dispersed species. Moreover, stepping stone dispersal between proximal populations is more likely to occur where habitat is suitable, and correlations between suitable habitat and genetic distance should reveal such behavior.

I compared 60 landscape-genetic models representing natural (i.e. flight) and human-aided (i.e. hitchhiking) dispersal for the invasive Asian tiger mosquito (*Aedes albopictus*). Native to Southeast Asia, this species was introduced to the US in 1985 (Hawley et al. 1987), and rapidly spread to its current distribution across much of the eastern and Midwest US (Fig. 1). Because of its limited natural dispersal capacity (800-1000 m for adults; Liew & Curtis, 2004), human mechanisms were implicated in its rapid spread. In recent years, its US range has expanded northward, so I also evaluated the 60 models in three spatial regions: the entire US range, the range core, and the range edge. Overall, I predicted highways and forests to facilitate dispersal, reflecting a combination of human-aided and natural dispersal. I expected forests to facilitate dispersal because *Ae. albopictus* adults spend their days along forest edges near oviposition sites. I also expected dispersal along the northern range edge to be most influenced by natural routes because of the slow northward expansion, but I expected dispersal patterns among range-core sites to remain a combination of natural and human-aided routes.
Methods

Study Species and Sampling

The Asian tiger mosquito (*Ae. albopictus*, Diptera: Culicidae), or forest-edge/forest-day mosquito, is a container mosquito whose females lay eggs singly in treeholes and artificial containers just above water line. After rains, eggs become inundated and hatch into larvae that undergo 4 instars prior to a pupal stage. In urban areas, *Ae. albopictus* colonizes artificial containers (e.g. cemetery vases, discarded tires) and readily out-competes other larvae in most environments (Bevins 2007; Braks et al. 2004; Costanzo et al. 2005). Females lay eggs in multiple containers and multiple females lay eggs in the same container (Hawley 1988). Thus, a single container can potentially contain offspring from a wide selection of individuals from the local population. I collected larvae and pupae from ca. 20 abandoned flower vases (> 100 individuals per site) in cemeteries from 26 localities throughout its US range (Table 4, Fig. 7). The extent of the study area was broad to capture predicted long-distance dispersal patterns from human-aided dispersal.

Lab Husbandry

I reared immature stages to adults in black plastic cups (~300 ml) at a density of ~ 50 individuals per cup. Cups were placed in 30 x 30 cm mesh cages under a long-day photoperiod (18:6 L/D) and fed a combination of ground fish food and cultured infusion from leaves (*Quercus* spp., *Platanus* spp.) incubated in water for 7 days. Culture cups were filled with infusion water and leaves. After emergence from pupae, I killed adults by exposing them to freezing temperatures for ca. 30 minutes, and preserved them in 99% ethanol.
Genotyping

I extracted DNA from 1-2 legs per adult using a chelex extraction protocol (Floyd et al. 2005) and genotyped 739 individuals at nine polymorphic microsatellite loci. Four loci were specific to *Ae. albopictus* (AealbA9, AealbB52, AealbD2, and AealbF3; Porretta et al. 2006), four were developed for *Ae. polynesiensis* (Ap1, Ap2, Ap3, Ap5; Behbahani et al. 2003), and one was developed for *Ae. aegypti* (AC2, Slotman et al. 2007). Primers for non-specific loci were optimized to amplify DNA for *Ae. albopictus*. I conducted polymerase chain reaction (PCR) amplifications in a total volume of 20 µl using a Mastercycler gradient thermal cycler (Eppendorf AG, Hamburg, Germany) and a MyCycler thermal cycler (Bio-RAD, Hercules, CA, USA). I amplified samples under the following conditions: 94°C for 3 minutes, 30 cycles of 30s at 94°C, 45s at a locus-specific annealing temperature ranging from 48°C to 60°C, 45s at 72°C, and a final extension for 5 min at 72°C. Annealing temperatures were as follows: 48°C for AealbB52; 50°C for AealbA9, Ap2, Ap3, and Ap5; 52°C for Ap1; 53°C for AealbF3; and 55°C for AealbD2 and AC2. PCR product was visualized on a 2% agarose gel to verify amplification and alleles were scored with capillary electrophoresis on a CEQ8000 (Beckman-Coulter, Fullerton, CA) following the manufacturer’s protocol. I used Microchecker v2.23 (Van Oosterhout et al. 2004) to screen for high-frequency null alleles (>0.08), allelic dropout, and scoring errors. I tested for pairwise linkage disequilibria and significant deviations from HWE with Fisher exact tests with Genepop v4.0.7 (Raymond & Rousset 1995).
Range Expansion

To detect a signature for recent colonization along the northern range edge, I calculated allelic richness and heterozygosity and compared mean values between 11 range-core localities and 15 range-edge localities using ANOVA. Low metric values in edge compared to core localities should reflect recent colonization of edge relative to core localities.

Genetic Distance

I measured two types of genetic distance shown to perform well in landscape genetic studies (Garroway et al. 2011; Hether & Hoffman 2012): Dps’ and cGD. I measured Dps’ (hereafter Dps, Bowcock et al. 1994) using MICROSAT v 1.5b (Minch et al. 1996). Dps differs from heterozygosity-based distance measures (e.g. Fst, Rst) because it is based upon allele frequency distributions. Such distance measures are expected to reveal relatively recent genetic differences among populations, particularly those resulting from landscape characteristics, and are also useful when highly polymorphic markers such as microsatellites are used (Murphy et al. 2008).

I also measured conditional genetic distance (cGD, Dyer et al. 2010), or graph distance (Dyer & Nason 2004). Dyer et al. (2010) recently showed this approach was better than Fst and chord distance for estimating genetic distance in a landscape genetics context. Briefly, cGD is based upon graph theory (Dyer & Nason 2004), where localities (nodes) are connected to all other localities through a network of gene flow. The shortest distance between two localities is the sum of all edges (lines connecting localities) along the shortest path between them. Typical pairwise estimates of genetic distance limit comparisons to two localities, whereas cGD considers genetic distance between two localities to be influenced by direct gene flow between
the pair in addition to gene flow occurring through other populations within the network. Thus, I calculated cGD as the shortest graph distance between locality pairs using the Graph function in GeneticStudio (Dyer 2009).

**Least-cost Paths and Resistance Analyses**

I measured GIS-based least-cost path distances and used circuit theory (McRae et al. 2008) to calculate resistance distances between localities.

**Least-cost Paths**

I calculated least-cost path distances between localities to identify habitat types that were most permeable to Asian tiger mosquito dispersal. First, I created 24 cost surfaces representing hypothesized costs of traversing different cover types using a 30-m resolution land cover dataset (http://www.mrlc.gov/nlcd2006.php) in ArcGIS 10 (Table 5). Land-cover data were re-sampled to 1 km$^2$ to improve computational efficiency at the broad study extent. I chose 1 km$^2$ as grain size because it is equal to or smaller than the average natural dispersal distance for *Ae. albopictus* (Liew & Curtis 2004), and the grain should be smaller than the average dispersal distance for the study organism (Fortin & Dale 2005). Fifteen surfaces assigned a low cost to one land cover type and a high cost to all others. For instance, for the cost surface representing woody wetlands, I assigned a low cost to cells within the land cover raster that were dominated by woody wetlands and I assigned a very high cost to all other cells. The remaining nine surfaces assigned low costs to multiple land cover types: all water, all urban, all forest, and all farm assigned a low cost to all land-cover types within that category (e.g. all urban: low cost to low-intensity urban, medium-intensity urban, high-intensity urban, and developed open area). In this way, I was able to test the relative importance of individual land cover types (e.g. high-intensity urban) and general land
cover categories (e.g. urban areas) by comparing the performance of each distance measure calculated from them in statistical models (see: Model selection and parameter estimation, below). Because individuals that disperse between localities must traverse multiple land-cover types by necessity, I also created four geographic models (Models 1-4, Table 5) by assigning relative costs to multiple land-cover types based upon knowledge of habitat use and likely dispersal behavior of *Ae. albopictus*. Model 1 assigned low cost to all urban areas, wetlands, and forests and high costs to all other land cover; Model 2 assigned low cost to low-intensity urban areas, higher cost to medium- and high-intensity urban areas, and much higher costs to all other land-cover types, and Model 3 assigned low cost to wetlands, open water, and forests (equally) and higher cost across urban areas, while all other land cover was assigned much higher cost. For Model 4, I assigned costs to all land cover types based upon model output from a suite of statistical models comparing least-cost path distances calculated from each of the single land-cover types to genetic distance. I assigned costs as model weights from each of the models in a model selection process (see Model selection and parameter estimation, below). Finally, I created a highway cost surface by assigning costs based upon highway capacity; for instance, interstate highways were assigned low cost while rural routes were assigned high cost.

Using the Spatial Analyst extension in ArcGIS 10, I conducted least-cost path analyses for each cost surface. Cost-distance surfaces were created from each cost surface (above) by assigning a value to each raster cell that was the cumulative cost of arriving at it from a given source. Thus, a cost-distance surface was created for each locality (source) to calculate cost-distances to all other localities (26 localities x 24 cost surfaces = 624 cost-distance surfaces). The least-cost path was identified as the path from the source to a locality that incurred the least
cumulative cost across all raster cells. In total, I calculated 203,125 pairwise least-cost path distances (325 pair-wise distances over 624 cost-distance surfaces).

**Resistance Distances**

Circuit theory (McRae et al. 2008) is based upon the function of electrical circuits, and estimates pairwise genetic distances between points using a graph-theoretic framework. Raster cells are nodes in the network, and are connected to adjacent cells by resistors. The resistors are edges in the network, and their value is the mean of resistance values assigned to each raster cell in the pair. Raster cells are assigned a resistance value based upon the predicted resistance to movement provided by the land cover type within that cell.

I calculated resistance distances using Circuitscape 3.5.4 (Shah & McRae 2008). Circuitscape calculates a matrix of resistance distances from user-supplied resistance surfaces and locality information. I created resistance surfaces by assigning resistance values to all cells in 24 surfaces in the same fashion as costs were assigned to cost surfaces (see Least-cost paths, above). Conductance values, the inverse of resistance values, can also be assigned to cells, but I restricted my assignment to resistances to keep my analyses aligned with the least-cost path analysis. In Circuitscape, I selected pairwise mode to iterate across all focal nodes (localities), the option for focal points containing a single cell (rather than a region), and selected an 8-neighbor cell-connection scheme.

**Model Selection and Parameter Estimation**

I modeled landscape effects on genetic distance using an information-theoretic model-selection framework to select the most informative model(s) explaining genetic distance (Dps and cGD, Burnham and Anderson 2001). Landscape genetics often uses Mantel tests to correlate
pair-wise genetic distance with geographic distance (Storfer et al. 2010), which are generally limited to three matrix comparisons (one response matrix and two predictor matrices). Multiple regression on distance matrices (MRDM; Legendre et al. 1994) provides a framework to evaluate effects of multiple predictors on pair-wise genetic distance, to calculate the coefficient of determination ($R^2$) for each predictor, and to evaluate the relative importance of each predictor variable on genetic distance. Moreover, this approach has high power and little Type I error compared to commonly used Mantel and partial-Mantel tests (Balkenhol, Waits, et al. 2009). Thus, the relative role of multiple landscape characteristics can be objectively assessed using this approach.

I used multiple regression on distance matrices (MRDM, Legendre et al. 1994, Legendre and Legendre 1998, Lichstein 2006, Garroway et al. 2011) to fit 60 models predicting genetic distance. Models included single distance matrices (e.g. all water least-cost distance) and additive models with multiple distance matrices. I also included a model representing simple isolation by distance (IBD) as a matrix of straight-line geographic distances, a model representing a random cost surface, and 1 as a statistical null model to compare the performance of real models. In addition to pair-wise distances, I included two non-landscape variables into statistical models to incorporate human density and invasion history: 2010 census data and number of years-since-invasion. I evaluated collinearity among variables by creating a correlation matrix and calculating variance inflation factors (VIF). Variables with a correlation coefficient < 0.5 and a VIF < 1.1 were not considered collinear and were included within additive models. I calculated model significance using 10,000 permutations of Dps. I considered the model with the lowest AIC value and the highest model weight as the most informative model in the set. To compare subordinate models, I evaluated the evidence ratio as the weight of
the most informative model divided by the weight of the model in the comparison. The evidence ratio for the most informative model is always one. If the evidence ratio for a lesser model is, e.g., 3, the top model is three times more informative than the model in the comparison. I used the evidence ratio to compare the top model to the next best model and to the model for simple IBD. All statistical analyses were conducted in R (R Development Core Team 2011).

Results

Genotypic Data

I detected one locus (Ap2) with high-frequency null alleles and excluded it from subsequent analyses. Five comparisons out of 208 were significantly out of HWE but no locality had more than one locus out of HWE. Allelic richness ranged 27.81 – 36.33 (rarefied to 20 individuals), and was significantly lower in edge than core localities ($F_{1,24} = 5.20, p = 0.03$). Average heterozygosity ranged 0.46 – 0.58 and was not significantly different between edge and core localities ($F_{1,24}= 2.45, p = 0.13$). Pair-wise Dps ranged 0.11 – 0.28 (Table 6) and pair-wise cGD ranged 1.8 – 11.0 (Table 6).

Landscape Modeling

Conditional genetic distance performed poorly with a maximum $R^2 = 0.04$. Thus, I report modeling results for Dps only. The most informative model for all locality pairs was all-water cost-distance, and had a 0.77 probability of being the most informative model (Table 7). This model was more than 9 times better than the second best model and over 15,000 times better than simple IBD (Table 7). The linear relationship was positive: as cost distance increased
(availability of water bodies declined), genetic distance increased significantly (Table 8, Fig. 8D).

Among range-core localities only, the additive model of cost distance across all water bodies, model 1 resistance, and 2010 census size was the best model (Table 7) with a model weight of 0.86. This was 25.3 times better than the next best model and nearly 670 times better than simple IBD (Table 7). In general, genetic distance increased with an increase in all water cost distance (decrease in the amount of water bodies between localities) and with increased differences in census size between localities (Table 8, Fig. 8E).

Among range-edge localities only, the best model was all forest resistance distance, with a model weight of 0.24 (Table 7). This was nearly 163 times better than simple IBD, but a second model representing deciduous forest resistance distance was nearly equally supported with a model weight of 0.20 (Table 7). However, in this region of the US, deciduous forests dominate forested landscapes, so these models are likely an effect of deciduous forests. Genetic distance and all forest resistance distance had a significant negative relationship (Table 8, Fig. 8F).

Parameter estimates for all top models were significantly different from those achieved randomly (Table 8). Coefficients of determination ($R^2$) for top models were highest when evaluating core localities only (0.33), followed by edge localities (0.10) and all localities (0.09; Table 8, Fig. 8).
Table 4. Geographic coordinates (decimal degrees), range location (Region), and sample size for all sampling localities.

<table>
<thead>
<tr>
<th>Locality</th>
<th>Code</th>
<th>Region</th>
<th>Latitude</th>
<th>Longitude</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atlanta, GA</td>
<td>ATL</td>
<td>Core</td>
<td>33.75</td>
<td>-84.45</td>
<td>29</td>
</tr>
<tr>
<td>Baton Rouge, LA</td>
<td>BATON</td>
<td>Core</td>
<td>30.46</td>
<td>-91.14</td>
<td>31</td>
</tr>
<tr>
<td>Birmingham, AL</td>
<td>BIRM</td>
<td>Core</td>
<td>33.55</td>
<td>-86.75</td>
<td>31</td>
</tr>
<tr>
<td>Charlotte, NC</td>
<td>CHAR</td>
<td>Core</td>
<td>35.23</td>
<td>-80.84</td>
<td>31</td>
</tr>
<tr>
<td>Columbia, SC</td>
<td>COLSC</td>
<td>Core</td>
<td>33.97</td>
<td>-80.95</td>
<td>31</td>
</tr>
<tr>
<td>Jacksonville, FL</td>
<td>JACK</td>
<td>Core</td>
<td>30.37</td>
<td>-81.65</td>
<td>31</td>
</tr>
<tr>
<td>Memphis, TN</td>
<td>MEM</td>
<td>Core</td>
<td>35.12</td>
<td>-90.03</td>
<td>30</td>
</tr>
<tr>
<td>Nashville, TN</td>
<td>NASH</td>
<td>Core</td>
<td>36.15</td>
<td>-86.73</td>
<td>31</td>
</tr>
<tr>
<td>Raleigh, NC</td>
<td>RALNC</td>
<td>Core</td>
<td>35.79</td>
<td>-78.63</td>
<td>31</td>
</tr>
<tr>
<td>Richmond, VA</td>
<td>RICH</td>
<td>Core</td>
<td>37.54</td>
<td>-77.40</td>
<td>28</td>
</tr>
<tr>
<td>Springerfield, MO</td>
<td>SGF</td>
<td>Core</td>
<td>37.19</td>
<td>-93.28</td>
<td>30</td>
</tr>
<tr>
<td>Blacksburg, VA</td>
<td>BLACK</td>
<td>Edge</td>
<td>38.35</td>
<td>-81.61</td>
<td>22</td>
</tr>
<tr>
<td>Charleston, WV</td>
<td>CWV</td>
<td>Edge</td>
<td>38.35</td>
<td>-81.64</td>
<td>31</td>
</tr>
<tr>
<td>Chillicothe, OH</td>
<td>CHIL</td>
<td>Edge</td>
<td>39.32</td>
<td>-82.99</td>
<td>31</td>
</tr>
<tr>
<td>Columbia, MO</td>
<td>COLMO</td>
<td>Edge</td>
<td>38.95</td>
<td>-92.34</td>
<td>31</td>
</tr>
<tr>
<td>Doylestown, PA</td>
<td>DOYLE</td>
<td>Edge</td>
<td>40.32</td>
<td>-75.13</td>
<td>29</td>
</tr>
<tr>
<td>Hagerstown, MD</td>
<td>HAG</td>
<td>Edge</td>
<td>39.63</td>
<td>-77.72</td>
<td>31</td>
</tr>
<tr>
<td>Harrisburg, PA</td>
<td>HARPA</td>
<td>Edge</td>
<td>40.27</td>
<td>-76.87</td>
<td>31</td>
</tr>
<tr>
<td>Harrisonburg, VA</td>
<td>HARVA</td>
<td>Edge</td>
<td>38.43</td>
<td>-78.85</td>
<td>22</td>
</tr>
<tr>
<td>Indianapolis, IN</td>
<td>INDY</td>
<td>Edge</td>
<td>39.77</td>
<td>-86.15</td>
<td>31</td>
</tr>
<tr>
<td>Liberty, MO</td>
<td>LIB</td>
<td>Edge</td>
<td>39.23</td>
<td>-94.42</td>
<td>31</td>
</tr>
<tr>
<td>Litchfield, IL</td>
<td>MAC</td>
<td>Edge</td>
<td>39.17</td>
<td>-89.63</td>
<td>31</td>
</tr>
<tr>
<td>Marshall, MO</td>
<td>MARM</td>
<td>Edge</td>
<td>39.10</td>
<td>-93.17</td>
<td>31</td>
</tr>
<tr>
<td>Portsmouth, OH</td>
<td>PORT</td>
<td>Edge</td>
<td>38.73</td>
<td>-82.97</td>
<td>30</td>
</tr>
<tr>
<td>St. Louis, MO</td>
<td>STL</td>
<td>Edge</td>
<td>38.64</td>
<td>-90.29</td>
<td>22</td>
</tr>
<tr>
<td>West Peoria, IL</td>
<td>PEO</td>
<td>Edge</td>
<td>40.70</td>
<td>-89.62</td>
<td>31</td>
</tr>
</tbody>
</table>
Table 5. Cost surfaces for calculating least-cost path and resistance distances representing individual land cover types and combinations within categories. Combined surfaces (all water, all urban, all forest, and all farm) were created by assigning low costs or low resistances to all individual land-cover types within that category (e.g. woody wetlands, herbaceous wetlands, and open water for the all water surface). Additional combined surfaces assigned relative costs to land cover types. Model 1 assigned low cost to all urban areas, wetlands, and forests and high costs to all other land cover; Model 2 assigned low cost to low-intensity urban areas, higher cost to medium- and high-intensity urban areas, and much higher costs to all other land-cover types; Model 3 assigned low cost to wetlands, open water, and forests (equally) and higher cost across urban areas, while all other land cover was assigned much higher cost; Model 4 assigned costs based upon model output from a suite of statistical models comparing least-cost path distances calculated from each of the single land-cover types to genetic distance.

<table>
<thead>
<tr>
<th>Land-cover type</th>
<th>Land-cover type</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Water:</strong></td>
<td></td>
</tr>
<tr>
<td>woody wetlands</td>
<td>pasture</td>
</tr>
<tr>
<td>herbaceous wetlands</td>
<td>crops</td>
</tr>
<tr>
<td>open water</td>
<td>all farm</td>
</tr>
<tr>
<td>all water</td>
<td>Scrub</td>
</tr>
<tr>
<td><strong>Urban areas:</strong></td>
<td></td>
</tr>
<tr>
<td>low</td>
<td>Grassland</td>
</tr>
<tr>
<td>medium</td>
<td>Non-urban</td>
</tr>
<tr>
<td>high urban</td>
<td>Model 1</td>
</tr>
<tr>
<td>developed open area</td>
<td>Model 2</td>
</tr>
<tr>
<td>all urban</td>
<td>Model 3*</td>
</tr>
<tr>
<td><strong>Forest:</strong></td>
<td></td>
</tr>
<tr>
<td>deciduous*</td>
<td>Model 4</td>
</tr>
<tr>
<td>evergreen</td>
<td></td>
</tr>
<tr>
<td>mixed</td>
<td>Highways</td>
</tr>
<tr>
<td>all forest</td>
<td></td>
</tr>
</tbody>
</table>
Table 6. Pairwise cGD (bottom left triangular matrix) and Dps (top right triangular matrix) among localities within the core and along the edge of the US range for *Ae. albopictus*. Comparisons within regions (core or edge) are shaded.

<table>
<thead>
<tr>
<th>Range core</th>
<th>Range edge</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>ATL</td>
<td></td>
</tr>
<tr>
<td>RATON</td>
<td>0.05</td>
</tr>
<tr>
<td>BIRM</td>
<td>0.03</td>
</tr>
<tr>
<td>CHAR</td>
<td>0.07</td>
</tr>
<tr>
<td>COLSC</td>
<td>0.06</td>
</tr>
<tr>
<td>JACK</td>
<td>0.07</td>
</tr>
<tr>
<td>MEM</td>
<td>0.05</td>
</tr>
<tr>
<td>NASH</td>
<td>0.07</td>
</tr>
<tr>
<td>RALNC</td>
<td>0.05</td>
</tr>
<tr>
<td>RICH</td>
<td>0.04</td>
</tr>
<tr>
<td>SGF</td>
<td>0.05</td>
</tr>
<tr>
<td>BLACK</td>
<td>0.05</td>
</tr>
<tr>
<td>BUCKS</td>
<td>0.03</td>
</tr>
<tr>
<td>CHIL</td>
<td>0.06</td>
</tr>
<tr>
<td>COLMO</td>
<td>0.03</td>
</tr>
<tr>
<td>CWV</td>
<td>0.05</td>
</tr>
<tr>
<td>HAG</td>
<td>0.05</td>
</tr>
<tr>
<td>HARPA</td>
<td>0.03</td>
</tr>
<tr>
<td>HARVA</td>
<td>0.10</td>
</tr>
<tr>
<td>INDY</td>
<td>0.05</td>
</tr>
<tr>
<td>LIB</td>
<td>0.05</td>
</tr>
<tr>
<td>MAC</td>
<td>0.02</td>
</tr>
<tr>
<td>MARM</td>
<td>0.05</td>
</tr>
<tr>
<td>PEI</td>
<td>0.05</td>
</tr>
<tr>
<td>PORT</td>
<td>0.06</td>
</tr>
<tr>
<td>STL</td>
<td>0.06</td>
</tr>
</tbody>
</table>
Table 7. Two most informative models and the geographic distance model (representing isolation by distance) for genetic distance (Dps) between all locality pairs, core pairs, and edge pairs. Numerator (ANOVA) degrees of freedom, K = number of parameters, log (£) = log likelihood, AIC: Akaike’s Information Criterion, \(w_i\) = model weight, evidence = evidence ratio. The most informative model was chosen as the model with the lowest AIC and the highest model weight. Model parameters provided in Table 8. CD: cost distance along a least-cost path; R: resistance distance.

<table>
<thead>
<tr>
<th>Model</th>
<th>(d)_{num}</th>
<th>K</th>
<th>log (£)</th>
<th>AIC</th>
<th>(w_i)</th>
<th>evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>All pairs (n = 325)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>all water CD</td>
<td>323</td>
<td>2</td>
<td>695.48</td>
<td>-1384.90</td>
<td>0.77</td>
<td>1</td>
</tr>
<tr>
<td>woody wetland CD + open water CD</td>
<td>322</td>
<td>3</td>
<td>694.27</td>
<td>-1384.96</td>
<td>0.08</td>
<td>9.3</td>
</tr>
<tr>
<td>geographic distance (log)</td>
<td>323</td>
<td>2</td>
<td>685.86</td>
<td>-1365.65</td>
<td>0.00</td>
<td>15069</td>
</tr>
<tr>
<td><strong>Core pairs (n = 48)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>all water CD + model 1 R + census 2010</td>
<td>44</td>
<td>4</td>
<td>123.15</td>
<td>-235.37</td>
<td>0.86</td>
<td>1</td>
</tr>
<tr>
<td>all water CD</td>
<td>46</td>
<td>2</td>
<td>117.59</td>
<td>-229.17</td>
<td>0.03</td>
<td>25.3</td>
</tr>
<tr>
<td>geographic distance</td>
<td>46</td>
<td>2</td>
<td>114.31</td>
<td>-222.63</td>
<td>0.001</td>
<td>666.6</td>
</tr>
<tr>
<td><strong>Edge pairs (n = 117)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>all forest R</td>
<td>115</td>
<td>2</td>
<td>249.53</td>
<td>-492.96</td>
<td>0.24</td>
<td>1</td>
</tr>
<tr>
<td>deciduous forest R</td>
<td>115</td>
<td>2</td>
<td>249.34</td>
<td>-492.59</td>
<td>0.20</td>
<td>1.2</td>
</tr>
<tr>
<td>geographic distance</td>
<td>115</td>
<td>2</td>
<td>244.44</td>
<td>-482.77</td>
<td>0.001</td>
<td>162.6</td>
</tr>
</tbody>
</table>
Table 8. ANOVA table and parameter estimates for most informative model for genetic distance (Dps) between all pairs of localities, core localities only, and edge localities only. Source of variance, df: degrees of freedom, SS: sum of squares, MS: mean square, F-ratio, $\beta_0$: intercept, $\beta_1$: slope, $R^2$: coefficient of variation, $p_{slope}$: probability of an equal or greater slope of the regression line (10,000 permutations), CD: cost distance along least-cost path, R: resistance distance.

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>$\beta_0$</th>
<th>$\beta_1$</th>
<th>$R^2$</th>
<th>$p_{slope}$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>All pairs (n = 325)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>all water CD</td>
<td>1</td>
<td>0.03</td>
<td>0.03</td>
<td>31.20</td>
<td>0.05</td>
<td>0.009</td>
<td>0.09</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>error</td>
<td>323</td>
<td>0.26</td>
<td>0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Core pairs (n = 48)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>all water CD</td>
<td>1</td>
<td>0.004</td>
<td>0.004</td>
<td>10.00</td>
<td>0.09</td>
<td>0.007</td>
<td>0.33</td>
<td>0.0004</td>
</tr>
<tr>
<td>model 1 R</td>
<td>1</td>
<td>0.001</td>
<td>0.001</td>
<td>3.63</td>
<td>-0.0003</td>
<td></td>
<td></td>
<td>0.04</td>
</tr>
<tr>
<td>census 2010</td>
<td>1</td>
<td>0.003</td>
<td>0.003</td>
<td>7.85</td>
<td>4.25 x 10^{-8}</td>
<td></td>
<td></td>
<td>0.004</td>
</tr>
<tr>
<td>error</td>
<td>44</td>
<td>0.01</td>
<td>0.0004</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Edge pairs (n = 117)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>all forest R</td>
<td>1</td>
<td>0.01</td>
<td>0.01</td>
<td>12.2</td>
<td>0.23</td>
<td>-0.00014</td>
<td>0.10</td>
<td>0.0004</td>
</tr>
<tr>
<td>error</td>
<td>115</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 7. Sample localities in 26 US cities within the US range for *Ae. albopictus*. Range-edge localities represented by circles; range core localities represented by squares. Inset: location of study area. Darker shading indicates higher habitat suitability (Medley 2010). Inset: location of study area.
Figure 8. Linear relationship for geographic model (IBD) and for most informative models for genetic distance (Dps) for all locality pairs (n = 325), range-core localities (n = 48), and range-edge localities (n = 117). Dashed lines represent 95% C.I.
Discussion

My results suggest that gene flow is facilitated by water bodies across the US range for *Ae. albopictus*. This is contrary to my expectations that highways are most important for facilitating gene flow due to human-aided propagule transport. The US introduction of *Ae. albopictus* into Houston, TX in 1985 implicated the used tire trade for transporting propagules from SE Asia (Craven et al. 1988; Hawley et al. 1987), and subsequent spread across the US was assumed to occur via similar routes (e.g. commercial or private vehicles; (Moore & Mitchell 1997). My results don’t negate this notion; rather, I hypothesize that the US invasion had an initial spreading phase including long-distance hops and a range-filling phase via natural dispersal after initial sites were established. The spreading phase was a likely result of human-aided movement either within the range or via additional introductions rather than natural flight because of the distances traveled over the short time frame and the limited dispersal ability of this species (~1km; Niebylski and Craig 1994, Braks et al. 2004, Liew and Curtis 2004). Because populations throughout most of the US range have become well-established and abundant, any overall signature of current human-aided gene flow could be swamped by natural gene flow.

It is not likely that gene flow is facilitated by water bodies *per se*, but due to another environmental factor correlated with water body density. Mosquitoes are rarely found in permanent waters because of predation by fishes and predaceous insects, and container mosquitoes are rarely found outside of natural or artificial containers. However, wetlands and natural lakes are generally more numerous in mesic areas with high relative rainfall and humidity, and containers are reliant on frequent rainfall events to maintain habitat for developing
mosquitoes. Thus, connectivity may be facilitated where a higher density of container habitats is maintained (via climate) and stepping-stone dispersal is facilitated by reduced distances between local populations.

Range-wide trends could mask variation in potentially meaningful pair-wise genetic distances. For instance, genetic distance between some core-edge pairs is nearly as low as the minimum genetic distance overall (Dps = 0.11). Richmond, Virginia has a Dps value of 0.12 with three edge cities: Columbia, Missouri; Charleston, West Virginia; and Litchfield, Illinois (Table 4). Low genetic distance between distant localities could be explained by: 1) similar ancestry from recent introductions or 2) high pair-wise gene flow. If introductions are ongoing, Richmond (and surrounding areas) could share alleles with distant US localities because of shared source population(s). Richmond is ~120 km from Chesapeake Bay, which is a major shipping port and could be receiving new propagules. However, the genetically similar localities along the US range edge are far inland, and the genetic similarity with Richmond is unlikely a result of introductions from the same international source because inland sites lack international ports. This suggests that Richmond could be highly connected to these geographically distant edge cities, possibly through human-facilitated dispersal.

Model explanatory power is reduced as edge localities are excluded from analyses, potentially a result of pair-wise outliers or invasion history. Ongoing monitoring efforts and reduced allelic richness in edge versus core localities suggests northern range-edge populations of *Ae. albopictus* were established more recently than 1985. Thus, relationships between genetic distance and landscape distance could have broken down along the range edge because edge populations are more recently affected by founder effects and non-equilibrium conditions. Moreover, pair-wise outliers in genetic-landscape relationships at the range edge could be
attributed to metapopulation dynamics. *Ae. albopictus* overwinters as diapausing eggs, and its success in the temperate US has been attributed to this strategy (Hawley et al. 1987). However, if winter conditions in northern US populations exceed the lower lethal temperature of diapausing eggs (Hanson & Craig Jr. 1995), local extinctions followed by re-colonizations from highly connected core localities could contribute to some pair-wise genetic similarities. Annual genetic analyses could test this hypothesis; consistent pair-wise genetic similarities between core and edge localities over time suggest annual re-colonization rather than overwintering success.

Chamber experiments evaluating over-wintering capabilities of core vs. edge populations could further elucidate the ecology of northern US populations and quantify the potential for northern populations becoming sources for further range expansion.

Forest resistance distance was the best model for Dps among edge localities, but the negative slope of the relationship suggests forests impede gene flow. Because *Ae. albopictus* inhabits forest edges both as larvae in treeholes and as adults resting near areas with suitable oviposition sites, I expected forests to facilitate gene flow. However, previous work examining treehole occupancy in undisturbed wooded areas showed little evidence that *Ae. albopictus* has been able to displace native *Ae. triseriatus* from treehole habitats (Lounibos et al. 2001). Thus, *Ae. albopictus* may occupy so few treeholes in forests that populations are very patchy or absent, effectively reducing gene flow across this habitat type. In addition, a high density of forests along the northern range edge is overlain with higher elevation in the Appalachian Mountains so that the effective dispersal hindrance may be an interactive effect of climate and elevation. Moreover, forests may inhibit wind-mediated dispersal.

Conducting landscape genetics analyses at three spatial scales shows that a priori predictions about genetic structure based upon invasion history or hypothesized range structure
can improve model fit and inform geographically contingent management strategies. A common approach for landscape genetic studies is to first define genetic structure and to analyze landscape/genetic correlations within genetic clusters (Zalewski et al. 2009), particularly when hierarchical structure is hypothesized (Perrier et al. 2011; Rasic & Keyghobadi 2011). For this study, no genetic sub-structure was detected in a preliminary study (K. Medley unpublished), but I suspected different trends in different geographic regions based upon invasion history (i.e. recent northward range expansion) and range-limit theory. Thus, I propose future landscape genetics studies structure hypotheses based upon genetic sub-units, but also consider ecological history and dispersal behavior (e.g. long-distance and natural in this case). This approach is relevant for species expanding their range due to invasion or climate change and for range-contractions due to habitat loss or climate change. For this study, defining regions allowed a better understanding of this invasion and highlighted potential strategies for mitigating spread.

The differential performance of Dps and cGD in this study showed different genetic distance measures are appropriate for different study designs. Conditional genetic distance (cGD) was developed as a more realistic measure of gene flow, particularly when correlating gene flow with landscape features (Dyer et al. 2010; Garroway et al. 2011). The poor performance of cGD relative to Dps for this study may be a result of study scale relative to sampling density. The role of unsampled “ghost populations” (Beerli 2004) is poorly understood, but simulated node removal within the genetic network can drastically affect graph distance, particularly if removed nodes are not highly connected (Garroway et al. 2008). I predict that ghost populations could have larger effects on cGD as the putative number of unsampled populations increases relative to number of sampling localities. The large spatial scale of this study could result in larger effects
of this phenomenon, but the effects of ghost populations at increasing spatial scales should be studied further.

My approach avoids the problems associated with assigning resistance values or costs to habitat types (Spear et al. 2010) because I set up models \textit{a priori} and used an objective information-theoretic model-selection framework. I included cost surfaces representing individual land-cover types in my landscape models. Thus, the potential for cost inaccuracy was eliminated because costs were essentially relative and I tested individual land-cover types. Landscape models including multiple land-cover types were assigned costs based upon my hypotheses, and the 60 statistical hypotheses were compared objectively. This approach also allowed me to objectively compare least-cost path distances with resistance distances. Resistance distances have proven very useful for landscape genetics studies because they capture multiple dispersal pathways between localities. Population-level genetic distance is a result of dispersal by multiple individuals, and individuals are unlikely to use the same path between localities. However, the least-cost path across water bodies was best for this study at the extent of all localities and among range-core localities. It’s unclear why this was so, excepting perhaps differences in landscape configuration by land cover type. However, had I not included least-cost paths in the model selection framework, the importance of water bodies (or associated climatic characteristics) would not have been detected.

A major challenge ahead for landscape genetics and species invasions is understanding landscape effects on the distribution and spread of adaptive variation (Balkenhol, Gugerli, et al. 2009). For my study system, the relationship between landscape features and variations in traits affecting over-wintering success (i.e. diapause) could improve the understanding of US range expansion into more northern localities than predicted by native climate (Medley 2010). Altered
gene flow resulting from different landscape permeability in non-native versus native ranges in addition to human-aided dispersal can also change adaptive potential. If background rates of gene flow are low, facilitated gene flow could provide the genetic material necessary for adaptation to occur, particularly at range edges where populations are sometimes genetically and demographically depauperate, fragmented, and subject to genetic drift (Garant et al. 2007; Holt 2003; Vucetich & Waite 2003). Moreover, genetic admixture from multiple introductions could create hybrids that outperform founders (Geiger et al. 2011), or increase genetic variation sufficiently for adaptation to proceed (Gillis et al. 2009; Kolbe et al. 2004; Lavergne & Molofsky 2007). Indeed, many invasions have involved traits increasing invasive potential (Phillips, Brown, et al. 2006; Thomas 2001), but whether these are due to sampling effects from native sources populations, genetic admixture, or adaptation is unclear. Understanding the genetic consequences of landscape patterns on gene flow and adaptation are promising avenues of continued research.

Managing invasive species is a difficult task, but I show here how landscape genetics can be used to inform mitigation strategies. In the range core, homeowners must empty accumulated water from artificial containers. This recommendation is not new, but my results emphasize the importance of this practice not only for reducing local mosquito abundance but for eliminating sources of recolonization. Municipalities often spray larvicide or adulticide to manage populations, so focusing efforts where sources may contribute propagules to re-colonization events is important. For instance, managers may be able to streamline eradication efforts by evaluating local rainfall and limiting spraying to areas where rainfall events occurred ~1-week prior (typical development time to pupation). In northern populations, implications of timber
harvest or forest conversion to other land uses go beyond local effects; reduced forested habitat could facilitate dispersal and colonization events where the range is already expanding.
CHAPTER 4: RAPID LOCAL ADAPTATION DURING INVASION DESPITE ASYMMETRIC IMMIGRATION AT AN EXPANDING RANGE EDGE

Introduction

A growing body of evidence for rapid adaptation increases the importance of understanding evolutionary processes during species invasions (Lambrinos 2004; Lee 2002; Novak 2007; Sax et al. 2007; Suarez & Tsutsui 2008; Whitney & Gabler 2008). At the same time, species invasions can serve as natural experiments to test hypotheses generated from range-limit theory (Holt 2003; Sax et al. 2007). Species distributions theoretically reach limits through a combination of demographic and evolutionary mechanisms related to underlying environmental gradients. As the environment toward the range edge approaches a species’ physiological limits, populations become smaller and more fragmented than in the range core where environmental conditions are more hospitable (abundant center model; Andrewartha and Birch 1954, Gaston 2003). Increased fragmentation and small population size exacerbate Allee effects, and population size can continue to decline in the absence of constant immigration.

Demographic declines theoretically inhibit evolution at the range edge because abundant range-core populations produce more immigrants into range-edge populations than vice versa. Asymmetric migration can lead to gene swamping at the range edge, where genotypes adapted to range-core conditions swamp edge genotypes, and edge populations are maintained in a constant state of maladaptation (Kawecki & Holt 2002). However, immigration can also relieve demographic sinks, boosting population size long enough for adaptation to occur if sufficient
adaptive genetic variation is present (Alleaume-Benharira, Pen, & Ronce 2006; Gomulkiewicz, Holt, & Barfield 1999). Edge populations are expected to have reduced genetic variation compared to the range core because of founder events, small population size, and fragmentation (central-peripheral hypothesis; Hoffmann and Blows 1994). Recent work based upon broader evolutionary theory predicts that intermediate immigration rates into sub-optimal environments can improve genetic variation locally so that adaptive potential is improved (Garant et al. 2007). Most research on range limits relies on theoretical models, but such models have led to numerous hypotheses that can be tested empirically (Holt et al. 2003; Holt 2003; Keitt, Lewis, & Holt 2001). My study aims to empirically model the relationship between gene flow and local adaptation at a range edge during a species invasion.

A first step to empirically test the gene flow-adaptation relationship is to measure local adaptation. A sub-population (i.e. deme) that is locally adapted has higher relative fitness in its own habitat than genotypes originating from other habitats (Williams 1966). The extent of local adaptation is determined by measuring deme fitness in two or more habitats. As an example, consider two demes (hereafter termed genotypes), A and B, and two habitats, 1 and 2 (Fig. 9). If genotype A is always found in habitat 2 and genotype B is always found in habitat 1, this is a genotype by environment correlation. If A is more fit than B, the difference could be due to differences in habitat or differences in genotype. Fitness differences due to genotype suggest local adaptation, and can be detected only by rearing both genotypes in both habitats (i.e. reciprocal transplants) to determine if there is a genotype by environment interaction.

Kawecki and Ebert (2004) define two comparisons of fitness commonly used to assess local adaptation: “local vs. foreign” and “home vs. away” (Fig. 9). The “local vs. foreign” comparison assesses fitness in one habitat for a genotype that lives in that habitat (the “local”)
and another genotype originating from a different habitat (the "foreigner", Kawecki and Ebert 2004). For Fig. 9, this is any vertical comparison between genotypes A and B. In this case, local adaptation is indicated if the local genotype has higher fitness relative to the foreigner in its home habitat (Fig. 9A and 9B). The “home vs. away” comparison assesses fitness for one genotype reared under home and away conditions. Local adaptation in this comparison is indicated if the genotype performs better at home than it does away (Fig. 9A and 9C), but the “home vs. away” comparison is confounded by differences in habitat. If a genotype is optimally adapted to conditions in a poor habitat, it may experience an increase in fitness when placed in a resource-rich habitat, but may still be favored over other genotypes in its home habitat. Thus, fitness differences detected in the “local vs. foreign” comparison result from selective differences in habitats rather than differences in habitat per se.

Multiple tests of local adaptation for fitness-related traits have been conducted, but the role of gene flow in adaptive divergence is still poorly understood (Brown, Kann, & Rand 2001; Mopper et al. 2000). Moreover, frameworks for empirical studies don’t mirror hypotheses generated from theory. For instance, Dionne, et al. (2008) detected a correlation between hierarchical genetic structure and coastal distance for Atlantic salmon. They assumed this correlation was a signature of local adaptation, and concluded that gene flow was likely constrained by local adaptation. This approach has three flaws common in the literature: 1) local adaptation wasn’t measured adequately (i.e. reciprocally), 2) genetic distance was assumed to correlate directly with gene flow (i.e. omitting asymmetries), and 3) gene flow and local adaptation weren’t statistically modeled, but qualitatively compared. In another example, Wittkopp et al. (2010), identified a geographic cline in Drosophila pigmentation (an adaptive trait), but found no population genetic structure using putatively neutral loci. Their conclusion
was that pigmentation was a locally adapted trait despite high gene flow. Here, the pigmentation pattern was not evaluated using reciprocal transplants. Geographic clines are confounded by differences in habitat, and any differences in phenotype matching the environmental cline could be purely a result of phenotypic expression within those habitats. Moreover, the comparison between gene flow and the cline was qualitative.

Two goals of this study were to evaluate the role of gene flow in local adaptation for the invasive Asian tiger mosquito (*Aedes albopictus* Skuse) in the US and to test existing range-limit theory. The Asian tiger mosquito was first established in the continental US in Houston, TX in 1985, likely via used tires imported from Japan (Hawley et al. 1987). Initially established populations were capable of egg diapause, which implicated a temperate source of introduction and allowed the species to succeed in the temperate US (Hawley et al. 1987). Native populations are found in both temperate and tropical habitats, but tropical populations are incapable of egg diapause (Hawley 1988). Spread across the US was rapid, and the current range extends from Texas in the southwest to New Jersey in the northeast of its US range (Fig. 10). *Ae. albopictus* is naturally a poor disperser, travelling < 800 m from natal sites (Liew & Curtis 2004), so its rapid spread pointed to human-aided movement through, e.g., used tires and lucky bamboo as immature stages and as hitchhikers in automobiles for adults. Within the overall goals of this study, my objectives were to:

1) Determine whether US populations of *Ae. albopictus* were locally adapted by conducting reciprocal transplants throughout the US range.

2) Measure the extent of local adaptation using both local/foreigner and home/away comparisons.
3) Measure gene flow asymmetries and compare genetic diversity between core and edge populations.

4) Model the relationship between local adaptation and asymmetric gene flow.

Shortened photoperiods stimulate adult females to lay diapausing eggs (Bradshaw & Holzapfel 2007; Hawley 1988). Diapause expression is evident in temperate populations of Ae. albopictus in its native range, and was key to the survival of introduced propagules in the US (Hawley et al. 1987; Lounibos 2002). Moreover, diapause has evolved a cline in US and South American populations; southern US populations have reduced their diapause response, while southern Brazilian population have acquired egg diapause (Lounibos et al. 2003). Thus, I used diapause as a key component in this study; I exposed diapausing eggs to range-core and range-edge winters, hatched eggs and reared larvae under standardized conditions so that the combination of diapause ability and overwintering success affected results. After winter treatments, I measured pupal mass as a proxy for fitness to detect local adaptation because “energy reserves expended during diapause can have profound effects on post-diapause success” (Hahn & Denlinger 2007). I expected to detect local adaptation for edge and core populations in both “local vs. foreigner” and “home vs. away” comparisons. I also expected to detect a wide range of gene flow resulting from the combination of natural and human-aided dispersal. Thus, I hypothesized that intermediate rates of gene flow would facilitate local adaptation in sub-optimal range edges by counter-acting genetic drift and providing the genetic variation necessary for adaptive evolution. Conversely, I expected gene flow to have little effect on local adaptation for larger core populations that are unlikely subject to drift.
Methods

Study species

The Asian tiger mosquito (*Ae. albopictus* Diptera: Culicidae), or forest-edge/forest-day mosquito, is a container mosquito whose females lay eggs singly in treeholes and artificial containers just above water line. After rains, eggs become inundated and hatch into larvae that undergo 4 instars prior to a pupal stage. In urban areas, *Ae. albopictus* colonizes artificial containers (e.g. cemetery vases, discarded tires) and readily out-competes other larvae (Bevins 2007; Braks et al. 2004; Costanzo et al. 2005). Females lay eggs in multiple containers and multiple females lay eggs in the same container (Hawley 1988). Thus, a single container can potentially contain offspring from a wide selection of individuals from the local population.

Field collections

I conducted field collections in 14 cities: five range-core and nine range edge (Table 9, Fig. 10). Larvae and pupae were collected during summer 2010 from abandoned flower vases in cemeteries; ca. 20 vases were sampled per cemetery for >100 individuals per site. Cities were chosen to complement genetics analyses from Chapter 3: core localities were large cities and both small and large cities were sampled along the range edge to capture a wide range of potentially human-aided transport to edge localities. Vases were emptied into a mesh sieve, and mosquitoes were back-washed into plastic bottles filled with water from the vases. Larval and pupal density was kept low by splitting high-density collections into multiple bottles (< ca.100 per bottle). Larvae were provided ground fish food *ad libitum*. All bottles for each cemetery were
placed inside a mesh cage and the bottle lid was removed so that eclosed adults could escape into the cage; adults were provided organic white raisins as food. Bottles and cages were transported to UCF for continued culture.

**Lab husbandry**

Field-collected individuals were cultured in mesh cages in an environmental chamber maintained at 26.5 ± 1 °C, a 14L:10D photoperiod, and variable humidity ≥70%. Humidity was maintained by placing trays of water below shelving inside the environmental chamber (i.e. the chamber was not capable of manipulating humidity independently). I reared immature stages to adults in black plastic cups (~300 ml) at a density of ~ 50 individuals per cup. Cups were placed in 30 x 30 cm mesh cages and fed a combination of ground fish food and cultured infusion from leaves (*Quercus* spp., *Platanus* spp.) incubated in water for 7 days. Culture cups were filled with infusion water and leaves. Adults were provided 5-7 fresh organic golden raisins twice weekly. Adult females were provided restrained Sprague-Dawley mice for blood meals three days/week (UCF IACUC project # 10-30) for one hour/feeding. Black cups were lined with brown paper towels (“egg papers”) for oviposition. After egg papers were removed from cups, eggs papers were air-dried and eggs were allowed to embryonate for 7 days prior to storage or inundation with water for subsequent hatching (*sensu* Lounibos et al. 2003). To reduce non-genetic maternal effects (Mousseau & Fox 1998), experiments (below) were conducted on F2 or F3 offspring from field-collected individuals.

**Experiment**

Egg papers from F2 or F3 offspring of field-collected individuals were submerged in tap water in black cups. Hatched larvae were reared to adults and cultured as above (“Lab
husbandry”) at 21.5° C, which is the optimal temperature at which photoperiodically induced traits are expressed (Pumpuni et al. 1992). Pupae and adults were exposed to a shortened photoperiod (10L:14D) to induce diapause. Diapause is induced in pupal and adult stages: females exposed to a shortened photoperiod produce eggs in diapause (Hawley 1988). Ten hours of daylight is a sufficiently short day length to induce diapause for populations of Ae. albopictus from multiple latitudes in the US (Lounibos et al. 2003). Egg papers collected from matings were dried slowly, allowed to embryonate for 7 days, and placed in ziploc bags with a moist paper towel and stored at 4° C until exposed to winter treatments.

Diapause-induced eggs were randomly split into over-wintering groups (by city) and placed in individual Ziploc bags for each home city X overwintering city combination (14 X 12 = 168). All eggs per winter treatment (in individual bags) were placed in a plastic tub and shipped to each over-wintering site. Tubs were placed outside and allowed to withstand natural winter conditions from December 31-March 1. Egg papers were returned to the lab, submerged in infusion water (“Lab husbandry”, above), and provided 0.50 mg ground fish food. Larvae were counted every other day and provided ground fish food ad libitum until pupation. Pupae were removed during larval counts and fixed in 90% ethanol to preserve samples for potential genetic analysis (90% ethanol is best for preserving arthropod DNA, e.g. King & Porter 2004), and stored for ~three months. This approach certainly led to loss of mass, but the proportion of lost mass should be equal across treatments. Moreover, studies show loss rates decline and masses stabilize at ~10-20 days post submersion (Wetzel, Leuchs, & Koop 2005). Lipid loss is an important concern for pupae in this study, but lipid leaching as a result of ethanol fixation is expected to occur over longer time periods (Edwards et al. 2009). Thus, because most mass is lost in the first few days of submersion, water loss (not lipids) may explain most of the lost mass.
Therefore, my approach should approximate dry mass for individual pupae, while allowing for specimen archiving for future analyses.

To measure pupal mass, I removed pupae from the ethanol, blotted them, and allowed them to air dry for ≥ 5 minutes (until pupal mass was stable), and measured mass to the nearest 0.01 mg on a Fisher Scientific accu-124D microbalance. The blotting technique has proven successful and consistent for measuring pupal mass (Armbruster & Hutchinson 2002).

Statistical analyses

I analyzed mean pupal mass and calculated relative pupal mass in two ways. First, I calculated relative pupal mass between local and foreign genotypes as the average pupal mass for individuals over-wintered in their home city divided by the average pupal mass for all foreign genotypes (i.e., those from the other region) over-wintered in that city. Because this required averaging values for foreign genotypes, sample size was reduced to 41 (from 198 for mean pupal mass). Second, I calculated relative pupal mass for the home-away comparison by dividing the average pupal mass for a genotype in its home region by the average pupal mass for individuals of that genotype in a city in the away region. This also reduced the sample size to 41 because pupal mass in home cities was averaged across the home region to compare to pupal mass in cities in the away region. Thus, factors in the analyses are regions (core, edge) rather than cities, and cities were treated as replicates within regions. Moreover, 1:1 relationships for local/foreign or home/away values are zero and indicate no advantage or disadvantage, positive values represent a relative advantage, and negative values represent a relative disadvantage. Finally, I tested the assumption that regions differ in overwintering temperatures by comparing daily
minimum and mean temperatures using ANOVA, where sample units were cities within the two regions. I obtained temperatures for the over-wintering period from Weather Underground (www.wunderground.com).

I created statistical models to test multiple hypotheses predicting raw pupal mass, local/foreigner relative pupal mass, and home/away relative pupal mass and used information-theoretic model selection (AIC, Akaike 1974, Burnham and Anderson 2001) to select the best model(s). For raw pupal mass, I included home region and winter region as factors, larval density, 1/immigration, and 1/net gene flow (see Gene Flow, below) as covariates, and the interaction between 1/immigration and home region in the full model. The quadratic (1/immigration + (1/immigration)^2) tests for peak adaptation at intermediate immigration rates. Additional models for pupal mass removed individual factors or covariates. Models for local/foreigner relative pupal mass excluded winter region but included average daily minimum temperature between home and winter sites as a continuous covariate, and models for home/away relative pupal mass omitted winter region. Raw and relative pupal mass values were ln-transformed to better approach parametric assumptions.

Gene Flow and Diversity

I calculated gene flow from microsatellite data gathered using techniques outlined in Chapter 3. I calculated immigration between each pair of cities using BayesAss 3.0 (Rannala and Mountain 1997, Wilson and Rannala 2003, http://rannala.org). BayesAss measures recent migration rates (over several generations) using a Bayesian approach, together with Markov Chain Monte Carlo (MCMC) resampling. Migration rates (Bayesian posterior probabilities) are
estimated with few assumptions and results can be applied to nonstationary populations far from
genetic equilibrium. Moreover, BayesAss can distinguish between genetic similarity due to
common ancestry versus that due to recent migration. I ran Monte Carlo analyses with 5,000,000
iterations, sampled every 500 generations, with the first million iterations discarded as “burn-in”.
I used a random number seed for iterations with a mixing parameter of 0.6 for allele frequencies
and inbreeding coefficients and a mixing parameter of 0.5 for migration rates. I conducted two
independent runs to check for convergence, and I averaged the final two results.

I also calculated net gene flow to test whether characterization of cities as genetic sources
or sinks was a better predictor of local adaptation than pairwise immigration. BayesAss
calculates asymmetric immigration rates, and so provides an estimate in two directions between
each pair of cities. To calculate net gene flow, I first subtracted emigration from immigration for
each pair-wise comparison for a given city and summed the values for each city. A negative
value indicated net emigration and a positive value indicated net immigration from cities.

Finally, I tested hypotheses generated from range theory in two ways. I tested the
abundant center model by comparing pair-wise gene flow asymmetries between core and edge
localities using ANOVA. Significantly higher immigration into edge versus core localities was
predicted to support the abundant center model. I compared allelic richness and heterozygosity
between core and edge localities using ANOVA to test the central-peripheral hypothesis.
Significantly reduced allelic richness and heterozygosity in edge versus core populations was
predicted to support the central-peripheral hypothesis.
Results

Gene Flow and Diversity

Pair-wise immigration ranged from 0.005 to 0.095 individuals/generation. Immigration was significantly higher from core into edge localities than from edge into core localities (Table 10, Fig. 11). Net gene flow for localities ranged from -0.17 to 0.23 individuals/generation; there were no significant differences between core and edge localities; thus, pairwise immigration was asymmetric from core to edge localities, but a simple characterization of core localities as sources and edge localities as sinks was inadequate.

Allelic richness ranged 27.81 – 36.33 (rarefied to 20 individuals), and was significantly lower in edge than core localities (F1,24 = 5.20, p = 0.03). Average heterozygosity was high, ranging 0.46 – 0.58 but was not significantly different between edge and core localities (F1,24 = 2.45, p = 0.13).

Winter Temperatures

Daily minimum and mean temperatures during winter exposure were significantly lower in edge than in core over-wintering sites (min: F1,4 = 34.11, p = 0.004; mean: F1,4 = 44.96, p = 0.003; Figs. 12-13). Thus, cities within core and edge regions serve as replicates of overwintering conditions within each regional category.
Pupal Mass

Raw Pupal Mass

Mean pupal mass ranged widely (0.09 – 2.8 mg) and females were significantly larger than males ($F_{1,196} = 42.45$, $p << 0.001$). All further results refer to only female pupal mass because it is most critical for fitness (Armbruster & Hutchinson 2002). In the most informative model for raw female pupal mass, larval density and winter region were the only significant factors, but eggs that overwintered in the range edge developed pupae with significantly greater mass (ln-transformed) than eggs overwintering in the range core (Tables 11-12; Fig. 14).

Local/foreigner and home/away comparisons of female absolute pupal mass revealed signatures of local adaptation. Local genotypes had a higher pupal mass than foreigners across all localities (paired $t = 94.91$, df = 40, $p < 0.01$; Fig. 15). Genotypes over-wintered in their home region had a significantly greater pupal mass than those over-wintered away (paired $t = -112.77$, df = 40, $p < 0.01$; Fig. 15). Thus, local adaptation was indicated by both local/foreigner and home/away comparisons, and surprisingly, pupae that had over-wintered at the range edge as eggs were larger than those over-wintered at the range core.

Relative Pupal Mass: Local versus Foreign and Home versus Away Fitness Comparisons

Local vs. Foreign

Edge genotypes had a positive ln(relative female pupal mass) in 28 of 32 instances (87.5%), indicating local (edge) genotypes typically had greater pupal mass than foreign (core) genotypes in edge winters (Fig. 16). However, relative pupal mass for local (core) genotypes was always less than zero, indicating the foreign (edge) genotypes always had greater pupal mass.
than the local (core) genotypes in core winters (9 of 9 instances, Fig. 16). A Fisher exact test revealed this association was significant (p <0.001). Thus, edge genotypes are locally adapted to edge winters, but also out-perform core genotypes in core winters by attaining a larger female pupal mass.

In the most informative model for local/foreign relative female pupal mass, larval density and home region (core or edge) were the only variables significantly related to relative pupal mass (p<<0.001 for each). Interestingly, I did not detect an effect of immigration (Tables 11 & 13, Fig. 16).

**Home vs. Away**

Edge genotypes had a positive ln(relative female pupal mass) in 17 out of 18 instances (94.4%), indicating edge genotypes had greater pupal mass at home than away in the majority of comparisons (Fig. 17). In contrast, core genotypes typically had greater pupal mass when in edge winters than in their own home winters (18 out of 22 instances, 81.8%). Collectively (35 of 40 instances), genotypes performed better after overwintering in edge habitats than they did in core habitats. A Fisher exact test revealed this association was significant (p = 0.003).

In the most informative model for the home/away analysis, larval density, climatic conditions and gene flow were significantly related to relative female pupal mass (Tables 11 & 14, Fig. 17). Most of the variation was due to larval density, but there was a significant effect of immigration after this variation was accounted for. Moreover, the relationship was hump-shaped, indicating a peak in relative pupal mass at an intermediate rate of net immigration (Fig. 17).
Table 9. Geographic coordinates (decimal degrees) and range location (Region) for all sampling localities and overwintering sites.

<table>
<thead>
<tr>
<th>Locality</th>
<th>Region</th>
<th>Latitude</th>
<th>Longitude</th>
<th>Winter site (Y/N)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atlanta, GA</td>
<td>Core</td>
<td>33.75</td>
<td>-84.45</td>
<td>Y</td>
</tr>
<tr>
<td>Charlotte, NC</td>
<td>Core</td>
<td>35.23</td>
<td>-80.84</td>
<td>N</td>
</tr>
<tr>
<td>Columbia, SC</td>
<td>Core</td>
<td>33.97</td>
<td>-80.95</td>
<td>N</td>
</tr>
<tr>
<td>Memphis, TN</td>
<td>Core</td>
<td>35.12</td>
<td>-90.03</td>
<td>N</td>
</tr>
<tr>
<td>Raleigh, NC</td>
<td>Core</td>
<td>35.79</td>
<td>-78.63</td>
<td>Y</td>
</tr>
<tr>
<td>Columbia, MO</td>
<td>Edge</td>
<td>38.95</td>
<td>-92.34</td>
<td>Y</td>
</tr>
<tr>
<td>Doylestown, PA</td>
<td>Edge</td>
<td>40.32</td>
<td>-75.13</td>
<td>Y</td>
</tr>
<tr>
<td>Harrisburg, PA</td>
<td>Edge</td>
<td>40.27</td>
<td>-76.87</td>
<td>Y</td>
</tr>
<tr>
<td>Harrisonburg, VA</td>
<td>Edge</td>
<td>38.43</td>
<td>-78.85</td>
<td>Y</td>
</tr>
<tr>
<td>Indianapolis, IN</td>
<td>Edge</td>
<td>39.77</td>
<td>-86.15</td>
<td>Y</td>
</tr>
<tr>
<td>Marshall, MO</td>
<td>Edge</td>
<td>39.10</td>
<td>-93.17</td>
<td>Y</td>
</tr>
<tr>
<td>Portsmouth, OH</td>
<td>Edge</td>
<td>38.73</td>
<td>-82.97</td>
<td>Y</td>
</tr>
<tr>
<td>St. Louis, MO</td>
<td>Edge</td>
<td>38.64</td>
<td>-90.29</td>
<td>Y</td>
</tr>
<tr>
<td>Springfield, IL</td>
<td>Edge</td>
<td>39.78</td>
<td>-89.68</td>
<td>Y</td>
</tr>
<tr>
<td>West Peoria, IL</td>
<td>Edge</td>
<td>40.70</td>
<td>-89.62</td>
<td>Y</td>
</tr>
</tbody>
</table>

Table 10. ANOVA table for 1/immigration. Significance indicates gene flow is asymmetric; immigration is higher from core to edge than from edge to core localities. $R^2 = 0.16$.

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Region (core/edge)</td>
<td>1</td>
<td>16845</td>
<td>16845</td>
<td>13.48</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>error</td>
<td>65</td>
<td>81257</td>
<td>1250.1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 11. Three most informative models for each measure of pupal mass (Response). Numerator (ANOVA) degrees of freedom, K = number of parameters, log (£) = log likelihood, AIC: Akaike’s Information Criterion, \( w_i \) = model weight, evidence = evidence ratio. Most informative model was chosen as the model with the lowest AIC and the highest model weight. Model parameters provided in Tables 12-14.

<table>
<thead>
<tr>
<th>Response</th>
<th>Model</th>
<th>df</th>
<th>K</th>
<th>log (£)</th>
<th>AICc</th>
<th>( w_i )</th>
<th>evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>pupal mass (ln)</td>
<td>larval density + home region + winter region + home region*winter region</td>
<td>91</td>
<td>5</td>
<td>-76.73</td>
<td>166.14</td>
<td>0.58</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>larval density + winter region*1/immigration</td>
<td>92</td>
<td>4</td>
<td>-79.10</td>
<td>168.64</td>
<td>0.17</td>
<td>3.49</td>
</tr>
<tr>
<td></td>
<td>larval density + 1/immigration + (1/immigration)^2 + home region + winter region + home region*winter region</td>
<td>89</td>
<td>7</td>
<td>-76.11</td>
<td>169.50</td>
<td>0.11</td>
<td>5.36</td>
</tr>
<tr>
<td>relative pupal mass (ln local/foreign)</td>
<td>larval density + 1/immigration + home region + home region:1/immigration</td>
<td>35</td>
<td>5</td>
<td>-24.66</td>
<td>63.08</td>
<td>0.31</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>home region</td>
<td>38</td>
<td>2</td>
<td>-28.38</td>
<td>63.08</td>
<td>0.31</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>larval density + net gene flow + home region + home region*net gene flow</td>
<td>35</td>
<td>5</td>
<td>-25.36</td>
<td>64.49</td>
<td>0.15</td>
<td>2.02</td>
</tr>
<tr>
<td>relative pupal mass (ln home/away)</td>
<td>larval density + net gene flow + (net gene flow)^2 + minimum daily temperature + home region + home region*netgf</td>
<td>33</td>
<td>7</td>
<td>-27.14</td>
<td>73.79</td>
<td>0.69</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>larval density + net gene flow + (net gene flow)^2 + min daily temp</td>
<td>35</td>
<td>5</td>
<td>-31.18</td>
<td>76.12</td>
<td>0.21</td>
<td>3.21</td>
</tr>
<tr>
<td></td>
<td>larval density + 1/immigration + net gene flow + minimum daily temperature + home region + home region*1/immigration</td>
<td>33</td>
<td>7</td>
<td>-29.78</td>
<td>79.06</td>
<td>0.05</td>
<td>13.93</td>
</tr>
</tbody>
</table>
Table 12. Most informative model for female pupal mass (ln-transformed). Multiple $R^2_{\text{adj}} = 0.39$, $F_{4,91} = 16.35$, $p << 0.001$.

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>larval density (ln)</td>
<td>1</td>
<td>16.20</td>
<td>16.20</td>
<td>53.01</td>
<td>$&lt;&lt;0.001$</td>
</tr>
<tr>
<td>home region</td>
<td>1</td>
<td>0.32</td>
<td>0.32</td>
<td>1.04</td>
<td>0.31</td>
</tr>
<tr>
<td>winter region</td>
<td>1</td>
<td>3.29</td>
<td>3.29</td>
<td>10.78</td>
<td>$0.001$</td>
</tr>
<tr>
<td>home region*winter region</td>
<td>1</td>
<td>0.18</td>
<td>0.18</td>
<td>0.58</td>
<td>0.45</td>
</tr>
<tr>
<td>error</td>
<td>91</td>
<td>27.80</td>
<td>0.31</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 13. Most informative model for relative pupal mass for the local/foreigner comparison. Response = ln(female pupal mass in home city/average female pupal mass in that city for all genotypes from the foreign region). Multiple $R^2_{\text{adj}} = 0.63$, $F_{4,35} = 17.8$, $p << 0.001$.

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>larval density (ln)</td>
<td>1</td>
<td>8.86</td>
<td>8.86</td>
<td>38.61</td>
<td>$&lt;&lt;0.001$</td>
</tr>
<tr>
<td>1/immigration</td>
<td>1</td>
<td>0.04</td>
<td>0.04</td>
<td>0.18</td>
<td>0.67</td>
</tr>
<tr>
<td>home region</td>
<td>1</td>
<td>7.44</td>
<td>7.44</td>
<td>32.40</td>
<td>$&lt;&lt;0.001$</td>
</tr>
<tr>
<td>1/immigration*home region</td>
<td>1</td>
<td>0.01</td>
<td>0.01</td>
<td>0.02</td>
<td>0.88</td>
</tr>
<tr>
<td>error</td>
<td>35</td>
<td>8.04</td>
<td>0.23</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 14. Most informative model for relative female pupal mass for the home/away comparison. Response = ln(average female pupal mass for a genotype in its home region/mass for that genotype in the away region). Multiple $R^2_{adj} = 0.69$, $F_{6,33} = 12.06$, $p <<0.001$. min daily: difference in average minimum daily temperature over the experimental winter between a genotype’s home city and away city.

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>larval density (ln)</td>
<td>1</td>
<td>8.77</td>
<td>8.77</td>
<td>31.79</td>
<td>$&lt;&lt;0.001$</td>
</tr>
<tr>
<td>net gene flow</td>
<td>1</td>
<td>3.76</td>
<td>3.76</td>
<td>13.64</td>
<td>0.001</td>
</tr>
<tr>
<td>(net gene flow)$^2$</td>
<td>1</td>
<td>2.65</td>
<td>2.65</td>
<td>9.62</td>
<td>0.004</td>
</tr>
<tr>
<td>min daily</td>
<td>1</td>
<td>2.74</td>
<td>2.74</td>
<td>9.95</td>
<td>0.003</td>
</tr>
<tr>
<td>home region</td>
<td>1</td>
<td>1.80</td>
<td>1.80</td>
<td>6.53</td>
<td>0.02</td>
</tr>
<tr>
<td>net gene flow*home region</td>
<td>1</td>
<td>0.23</td>
<td>0.23</td>
<td>0.84</td>
<td>0.37</td>
</tr>
<tr>
<td>error</td>
<td>33</td>
<td>9.10</td>
<td>0.28</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 9. Hypothetical patterns of genotype by environment interactions on fitness. Lines represent genotypes (A, B) originating from different habitats (2, 1, respectively). In panel A, local genotypes fare better than foreign genotypes, and genotypes fare better at home than away. In panel B, genotype A fares worse at home (habitat 2) than it does away, but still better than the foreign genotype. In Panel C, genotype A fares better at home than away, but genotype B still fares better under both conditions. In Panel D, genotype B fares better than the foreign genotype at home, but both genotypes fare worse in habitat 2; differences in fitness are due mostly to environment. Figure modified from Kawecki and Ebert (2004).
Figure 10. Sampling localities for range-edge (circles) and range-core (squares) populations. Red sites were not included as overwintering sites, but genotypes from those areas were overwintered elsewhere. Darker shading indicates higher habitat suitability (Medley 2010). Inset: location of study area.
Figure 11. Immigration from core to edge and from edge to core populations. Genes flow asymmetrically from core to edge localities. Error bars represent standard deviation; ANOVA was performed using 1/immigration as the response.
Figure 12. Average daily mean (top) and minimum (bottom) temperatures for over-wintering sites. Core sites: grey dashed lines; edge sites: black solid lines. Average values across all winter cities in core or edge regions shown with diamonds.
Figure 13. Average daily mean (left) and minimum (right) temperatures for core and edge winter sites. Both mean and minimum temperatures were significantly lower for edge sites (mean: $F_{1,9} = 44.96$, $p = 0.003$; min: $F_{1,9} = 34.11$, $p = 0.004$). Error bars represent standard deviation.
Figure 14. Female mean pupal mass by region. Home region shown inside bars. Ln (pupal mass) was significantly higher in edge winters than in core winters, and edge populations had a significantly higher ln(pupal mass) in both over-wintering regions. Error bars represent standard deviation.
Figure 15. Mean pupal mass for local vs. foreign and home vs. away comparisons. Local genotypes had significantly higher (ln) pupal mass than foreigners in their home region. Genotypes had significantly higher (ln) pupal mass in their home region compared to the away region. Paired t-test on ln(pupal mass): local/foreign $t = 94.91$, home/away $t = -112.77$. Error bars represent standard deviation.
Figure 16. Immigration versus (ln) relative pupal mass for the local/foreigner comparison. Relative mass = pupal mass for female mosquitoes over-wintering in their home city/average female pupal mass for foreigners over-wintering in an away region. Values in the grey box indicate the local genotype has a higher pupal mass than foreign genotypes. Edge localities: hollow circles; core localities: filled circles.
Figure 17. Net gene flow versus (ln) relative pupal mass for the home/away comparison. Relative mass = average female pupal mass for a genotype over-wintering in their home region/female pupal mass for the same genotype in a locality in the away region. Values in the grey box indicate the genotype has a higher pupal mass at home than away. Edge localities: hollow circles; core localities: filled circles. Model for net gene flow vs. (ln) relative pupal mass: $y = 0.79 + 2.02x - 22.25x^2$; $F = 5.59$, $p = 0.008$. 
Discussion

My results indicate populations of *Ae. albopictus* near the northern range edge in the US are well adapted to northern US winters and have differentiated in a fitness-related trait from populations in the southeastern core of the range. Core genotypes in the southeastern US are themselves capable of producing high-fecundity offspring even when eggs are exposed to extreme winters. However, core-edge differences did not correspond to immigration according to range-limit theory. Rather, an intermediate rate of net immigration was related to relative pupal mass for genotypes overwintered in the away region. I conclude that genotypes near the edge of the range are locally adapted just 25 years after introduction into the US, and that range-limit theory for native species may not clearly apply to a recently established and expanding invasive species.

Contrary to expectations, both edge and core genotypes generally had higher pupal mass when eggs were overwintered in edge versus core winters. Two scenarios provide plausible explanations for this pattern. First, smaller embryos inside eggs may have a higher lethal temperature than large embryos, leading to reduced survivorship for small versus large embryos during colder winters. If this is so, larger pupal mass for core genotypes in edge versus core winters is a statistical artifact: average pupal mass shifts upward as small values are omitted from summary statistics where small embryo survivorship is reduced. The relationship between survivorship and embryo mass is unclear, but survivorship improves with cold acclimation by lowering the lethal temperature for diapausing eggs (Hanson & Craig Jr. 1995; Romi, Severini, & Toma 2006). Because winter temperatures were unseasonably low when eggs were prepared to undergo winter treatments, I exposed eggs to cool temperatures (4°C in refrigerator) to allow
eggs to cold-acclimate and to prevent complete mortality upon exposure to real winters. However, cold-acclimation was standardized across all genotypes, so any effect of cold acclimation was equal for all eggs. Moreover, the relationship between embryo size and pupal mass has not been evaluated for this species, although there was no relationship between the two for the western treehole mosquito, *Ae. sierrensis* (Hawley 1985). However, whether differential survivorship by embryo size skewed embryo mass upwards does not affect the relationship between local/foreign comparisons and the detection of local adaptation. Small-embryo mortality also implies strong selection for large embryo size, and could also result in rapid adaptation to local (northern) conditions.

A second plausible hypothesis explaining higher pupal mass in edge winters is that edge-adapted genotypes have infiltrated core populations through time, improving the core response to cold winters. Relative pupal mass increases as immigration from edge into core localities increases, suggesting edge-winter-adapted alleles may be improving fitness for core genotypes under edge winters. Indeed, core and edge genotypes converge towards a similar pupal mass when exposed to core winters, although pupal mass for edge genotypes remains slightly higher. That edge genotypes have higher pupal mass in core winters than core genotypes suggests that 1) core populations have not fully converged on the edge phenotype, or 2) there is a trade-off between producing large pupae under extreme winters and living in a region with more benign winters (i.e. range core). Eggs that diapause have high fatty acid content resulting from an up-regulation of acetyl-CoA elongases (Urbanski et al. 2010). Producing more fatty acids for all eggs is energetically costly (Hahn & Denlinger 2007), so that selection against over-producing fatty acids where it is not necessary is likely strong. Large pupae develop into large adults (Armbruster & Hutchinson 2002), but the tradeoff between increased individual fecundity by
latitude is unclear (Leisnham et al. 2011). However, higher pupal mass for edge genotypes than core genotypes in core winters remains.

Patterns of asymmetric gene flow persist despite local adaptation along the range edge. When local adaptation is strong, local genotypes are expected to out-compete immigrants and prevent substantial genetic exchange (Hendry & Taylor 2004; Hendry, Taylor, & McPhail 2002). For this invasion, asymmetric immigration may be a signature of colonization of edge localities from core migrants rather than substantial current genetic exchange. In addition, adaptation may have occurred recently so that genetic structure between edge and core populations may increase and genetic exchange may decrease over time. Moreover, immigration rates in this study were quite low (< 0.025 migrants/generation) overall. Thus, the statistical model may not encompass the full range of gene flow necessary to capture a peak in local adaptation at intermediate immigration rates. Alternatively, asymmetric gene flow may be consistent with dispersal into edge localities, but selection may be strong enough to overwhelm gene swamping.

These results contribute to a growing body of evidence for adaptation during invasion. Other examples come from multiple taxonomic groups with a wide range of life histories. The invasive Drosophila subobscura developed a cline in wing length and body size within two decades of introduction into the US (Gilchrist, Huey, & Serra 2001). Body size clines were recently evaluated for Ae. albopictus in the US, but were not detected. However, there is no evidence for body size clines in the native range (Urbanski et al. 2012). Selection favors dispersal at expanding range edges, as individuals colonizing open habitats have an advantage. This is shown in the cane toad invasion in Australia, where toads at the expanding range edge have evolved longer legs that those deeper in the range (Phillips, Brown, et al. 2006). Multiple global cichlid invasions throughout the 20th century have resulted in morphological divergence
(Firmat et al. 2012), and plant invasions have shown evidence of adaptation of reproduction traits over contemporary time scales (Barrett, Colauatti, & Eckert 2008). These and numerous other studies highlight the importance of understanding evolutionary mechanisms and consequences of invasions (Whitney & Gabler 2008).

Range expansion resulting from adaptation for the Asian tiger mosquito has important implications to ecology, economics, and human health. Previous work showed that US populations of *Ae. albopictus* have spread into novel climate compared to those encountered in its native range (Chapter 2, Medley 2010). This study provides evidence that adaptation has occurred and could explain this shift, and highlights that adaptation could promote further range expansion. *Ae. albopictus* readily outcompetes native treehole mosquitoes in artificial containers (Lounibos et al. 2001), and is often the only mosquito species found in many containers, so further expansion could shift mosquito community structure. The Asian tiger mosquito is active throughout the day and can be a nuisance for humans during outdoor activities. It is also a lab-competent vector for numerous diseases, and has recently been responsible for outbreaks of dengue and chikungunya in other countries (Gratz 2004; Reiter, Fontenille, & Paupy 2006; Rezza, Nicoletti, Angelini, Romi, Finarelli, Panning, Cordioli, Fortuna, Boros, Magurano, Silvi, et al. 2007; Rezza 2012; Vazeille et al. 2007). A dengue outbreak occurred recently in Florida, and expansion of *Ae. albopictus* into new areas in the US could expose more people and facilitate transmission in the event of dengue spread. Chikungunya is an emerging virus that has recently developed a single mutation that improved its transmissibility by *Ae. albopictus* beyond that of it’s better-known vector, *Ae. aegypti* (DeLamballerie et al. 2008; Delatte et al. 2008; Schuffenecker et al. 2006). Although chikungunya has not arrived in the US, recent outbreaks have been reported in India and Italy and the concern for introduction to the US is certainly real
(Charrel et al. 2007; Reiter et al. 2006; Rezza, Nicoletti, Angelini, Romi, Finarelli, Panning, Cordioli, Fortuna, Boros, Magurano, Silvi, et al. 2007). Finally, combating *Ae. albopictus* involves larvicide, adulticide, and other techniques that require regular application, and are costly to local governments.

The results from this study provide mixed support for range-limit theory but point to intermediate immigration as a driver for one measure of adaptation. I detected asymmetric immigration from core to edge localities, which is a pattern consistent with the abundant core model (Gaston 2003). Support for the central-peripheral hypothesis was equivocal: allelic richness was reduced in edge populations compared to core populations, but heterozygosity (gene diversity) was equally high in core and edge localities. According to evolutionary theory, I predicted immigration to improve local adaptation for edge populations, assuming they had been recently colonized and were subject to drift. Indeed, intermediate gene flow improved relative pupal mass for home/away relative pupal mass, although no effect of gene flow was detected for local/foreigner relative pupal mass. Moreover, adaptation at the range edge has occurred despite asymmetric immigration. High heterozygosity across all populations (with or without effects of immigration) suggests a signature of invasions facilitated by multiple human-aided introductions and subsequent genetic admixture. Increased genetic diversity during species invasions has been increasingly detected, the first notable example being the brown anole invasion from Caribbean Islands into the US (Kolbe et al. 2004). Multiple similar studies subsequently ensued, and many found increased genetic diversity in invasive compared to native populations, often attribute to novel genotypes created by genetic admixture in invasive populations (Consuegra et al. 2011; Gillis et al. 2009; Hufbauer 2008; Kang, Buckley, & Lowe 2007; Keller & Taylor 2010; Marrs et al. 2008). A similar pattern could hold for *Ae. albopictus*, although genetic comparisons between
native and non-native populations is required to test this hypothesis. Continued immigration from non-US sources could also boost genetic diversity throughout the range and contribute to adaptive potential for many traits.

In sum, this study provides evidence for rapid adaptation for an invasive mosquito, and contributes to theory on gene flow and adaptation at range edges. At the levels of immigration I measured, local adaptation can occur despite asymmetric immigration. Strong selection may contribute to adaptation in the face of immigration. Moreover, genetic admixture during invasion provides the genetic variation necessary for adaptation. Rapid adaptation during invasions may be more common than previously thought, and highlights the importance of considering effects of adaptation during range expansions in non-native regions. For *Ae. albopictus*, understanding the role of adaptation for range expansion has important consequences for regional mosquito ecology, local economies, and human health. Overall, inconsistent support for range-limit theory, effects of intermediate gene flow on adaptation, high genetic diversity and adaptation to northern winters highlight a need for a new synthetic framework for species range structure, particularly in the face of continued transformation of natural systems (Ellis et al. 2010).
CHAPTER 5: SYNTHESIS

The overarching goal of the research presented herein was to investigate the extent that human-aided dispersal influences range expansions for invasive species within non-native regions by facilitating adaptation to novel environments. This goal emerged by observing characteristics of the Asian tiger mosquito invasion into the US together with developments in range theory. *Ae. albopictus* was introduced into Houston, TX in 1985, and its rapid spread across the eastern US implicated human-aided dispersal (Hawley et al. 1987; Hawley 1988). Recently, the US range has crept northward, warranting a detailed examination of the mechanisms behind this spread and an evaluation of potential for further expansion. Range theory hypothesizes that intermediate rates of gene flow can facilitate adaptation to sub-optimal environments, such as those found at range edges (Garant et al. 2007). Historic evidence of human-aided gene flow, together with limited natural dispersal (for a winged species), led to the hypothesis that intermediate gene flow rates in the US may have facilitated adaptation to northern climate. I evaluated this overall hypothesis with five objectives:

1) To test if US populations of *Ae. albopictus* occupied significantly different climatic and habitat conditions than those in the native range,

2) To evaluate the extent to which *Ae. albopictus*’ US dispersal is currently influenced by human-aided versus natural movement,

3) To determine whether local adaptation has occurred in the US,

4) To model the relationship between gene flow and local adaptation in the US range, and

5) To test whether the abundant core model and the central-peripheral hypotheses of traditional range-limit theory was supported during this invasion.
Objective 1 was addressed with reciprocal distribution models (Chapter 2), which revealed the US range for *Ae. albopictus* has spread further north than predicted from its native environment. Permutation tests revealed the same climate and habitat utilized in the US are available in the native range, but aren’t occupied by *Ae. albopictus* in SE Asia. Objective 2 was addressed using landscape genetics and microsatellite markers analyzed across the northern range margin and in the core of the distribution. Contrary to expectations, current gene flow patterns were consistent with natural dispersal rather than excessive human-aided movement, although some distant populations were more connected than expected by natural dispersal alone (Chapter 3). I tested Objective 3 (local adaptation) by conducting reciprocal transplants with populations sampled in Objective 2. Range-core and range-edge populations performed well along the northern edge of the US range, and edge populations were locally adapted (Chapter 4). Whether US populations have evolved beyond the abilities of native populations was not evaluated, but incongruent distribution models combined with evidence for local adaptation together provide strong support that US populations have adapted to northern US winters since their introduction in 1985. The relationship between gene flow and local adaptation (Objective 4) was unexpected: I detected no effect of gene flow on adaptation when comparing local vs. foreigner pupal mass, but relative pupal mass peaked at intermediate immigration rates for the home vs. away comparison of local adaptation (Chapter 4). Asymmetric migration rates from core to edge populations support the abundant core model, but equally high genetic diversity in core and edge populations are contrary to expectations of the central-peripheral hypothesis (Chapter 4). Moreover, high levels of genetic variation detected in edge and core populations (Chapter 3) support the hypothesis that multiple introductions (rather than a spread from a single
Houston introduction) have improved genetic variation in range-edge populations, likely facilitating adaptation together with the effects of intermediate gene flow.

Rapid adaptation during invasion has important implications for the ecology of invaded ecosystems, local economies, and human health. Invasive species are introduced, non-native species that gain a competitive advantage after natural obstacles to population growth and spread have been overcome (Valéry et al. 2008), and have been the subject of much ecological inquiry and debate (Bradley, Wilcove, & Oppenheimer 2009; Gurevitch & Padilla 2004; Kanarek & Webb 2010; Lounibos 2002; Mack et al. 2000; Richardson et al. 2000; Sakai et al. 2001; Sax et al. 2005; Simberloff & Von Holle 1999; Valéry et al. 2009; Willis et al. 2000). Continued spread facilitated by adaptation affects additional native species and ecosystems and can have lasting effects on ecological dynamics (Strayer et al. 2006). Some invasive species carry human disease (e.g., the Asian tiger mosquito, Aedes albopictus), and spread resulting from adaptation has the potential to increase vulnerability to disease for vectors carrying exotic diseases (Juliano & Lounibos 2005). The Asian tiger mosquito is a potential vector for numerous potentially severe and life-threatening diseases (Charrel et al. 2007a, Morens and Fauci 2008), including dengue and chikungunya virus. Moreover, vector competence for chikungunya improves for Ae. albopictus when immature stages are reared under cooler temperatures (Westbrook et al. 2010) so northward spread coupled with increased affinity of chikungunya at lower temperatures could be an unfortunate combination. Continued spread of Ae. albopictus via adaptation to novel climate and habitat increases global health risk for this emerging disease. For the Asian tiger mosquito, re-assessing risk or combating spread where risk is unexpected can be costly to local mosquito control and health agencies who apply insecticides and larvicides for prevention and
control. This study points to a need to develop better predictive models of invasion and disease risk *Ae. albopictus* and other invasive species.

Niche-based distribution models have grown in popularity to predict range expansions for invasive species, but are conservatively based upon the assumption that all aspects of the niche are conserved during invasion. Distribution models have proven useful for modeling geographic ranges for native species and can predict suitable site of introduction (Broennimann et al. 2007; Medley 2010), both of which can be useful to plan mitigation for invasive species in a new range. However, niche-conservative models fail for species that rapidly adapt to new environments (Medley 2010), although models created using data from both native and non-native occurrences improve model performance (Broennimann & Guisan 2008). When an invasive species adapts in a new range, mitigating its spread requires an iterative, trial and error approach to predict range expansion as the invasion proceeds and traits evolve. This strategy is markedly different from that for a niche-conserved invasive species, in which traits are fixed and strategies can be static. Alternatively, rapid range shifts may actually promote range stability as the environmental gradient between source and destination localities steepens (Phillips 2012), and models may be more accurate. However, preparation for further range expansion is a more appropriate approach than assuming rapid range stability when mitigating invasions.

Overall, this study provided mixed support for hypotheses generated by classic range-limit theory. Asymmetric immigration rates from core to edge populations were consistent with the abundant center model (Chapter 4), but this pattern does not clearly indicate larger populations within the range core relative to the range edge. Instead, genetic signatures for *Ae. albopictus* in the US are likely influenced by a history of multiple invasions rather than by a cline in population abundance from range core to range edge. However, landscape genetics
analyses (Chapter 3) provide indirect support for increased fragmentation along the range edge, a pattern expected by the abundant center model. Forests were a barrier to gene flow along the range edge, while water bodies facilitated gene flow within the range core; populations may be more continuous in the range core and fragmented by forests along the range edge. Further evaluation of the abundant center model requires direct measures of population abundance throughout the US range of *Ae. albopictus*. Measures of genetic diversity (Chapter 3) contrasted with the central-peripheral hypothesis. Allelic richness was lower in edge than core populations, but heterozygosity (gene diversity) was equally high throughout. This pattern may reflect recent range expansion, and neither supports nor rejects the central-peripheral hypothesis as a general rule for range structure. Finally, adaptation peaked at intermediate gene flow rates when traits were compared between home and away regions, but local/foreigner comparisons did not reveal an effect of gene flow. Theory predicts gene swamping can lead to range edges, but immigration can improve genetic variation enough to relieve populations from drift when population size is small. For this study, genetic variation was uniformly high. Moreover, pupal mass for edge populations was improved by intermediate immigration rates, but pupal mass was high for all populations in edge winters. Thus, immigration within the range may not explain overall adaptation to edge winter. Based upon previous genetic analyses (Black IV et al. 1988; Kambhampati et al. 1990) together with those generated within this dissertation, multiple introductions provide the most plausible explanation for the high performance of edge and core populations in northern (range-edge) winters. If multiple introductions and concomitant increased genetic diversity become the rule for species invasions, invasions cannot serve as tests of range-limit theory until invasions reach equilibrium. However, continued examination of
hypotheses generated by range-limit theory (and alternatives) provide very important clues to the success of invasive species.

This dissertation highlights the need for incorporating modern processes into evolutionary range theory, particularly in the face of an increasingly human-dominated landscape (Ellis et al. 2010) where human-aided, long-distance species invasions have become the rule rather than the exception (Olden et al. 2004). Evolutionary theory does not adequately represent Homogenee conditions (Olden et al. 2004), where invasive species are likely to be introduced multiple times, among multiple source populations, providing positive feedback to local adaptation over broad regions (an invasion ratchet; Medley 2010). Genetic admixture during invasions is a unique consequence of multiple, repetitive, human-aided introductions across long distances, and evidence for increased genetic diversity resulting from these processes is rapidly growing (Consuegra et al. 2011; Gillis et al. 2009; Hufbauer 2008; Kang et al. 2007; Keller & Taylor 2010; Knouft et al. 2006; Kolbe et al. 2004, 2008; Marrs et al. 2008). This trend vastly shifts expectation of range limits based upon classic theory, because theory was built upon study within natural systems and often assume equilibrium conditions. To work towards a more realistic range-limit theory, research would benefit from a shift from natural systems to human-dominated landscapes, because evolutionary theory does not adequately represent adaptation in modern, human-dominated landscapes (Vitousek et al. 1997).

In summary, this dissertation used a suite of analytical tools to provide evidence for rapid adaptive evolution during invasion. Distribution models provided evidence of niche shifts during invasion while highlighting the need for caution when generating distribution models to predict invasive species spread (Chapter 2). Analyses within a landscape genetics framework provided useful insight into dispersal pathways and mechanisms during this invasion, and provide
important information for mitigating spread (Chapter 3). My work contributes to the rapidly evolving discipline of landscape genetics, and contributed to the growing body of evidence for evolution during invasion. Research presented in Chapters 2-4 combines to suggest that range expansion for *Ae. albopictus* in the US could continue, and emphasizes the importance of continued diligence among urban and sub-urban communities to prevent spread. Finally, my work highlights the need for a better understanding of contemporary ecological and evolutionary processes leading to range-limits (or expansion) to more accurately reflect processes occurring in a human-dominated world.
REFERENCES


