SPECIATION AND DIVERSIFICATION IN THE NORTH AMERICAN TIGER BEETLES OF THE *CICINDELA SYLVATICA* GROUP: MORPHOLOGICAL VARIATION AND AN ECOPHYLOGEOGRAPHIC APPROACH

Daniel Paul Duran

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

In partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Biological Sciences

December, 2010

Nashville, Tennessee

Committee 2 (2009-2010)

Approved:

Professor David McCauley

Professor Patrick Abbot

Professor Christopher Nice

Professor David Pearson

Professor Charles Singleton

SPECIATION AND DIVERSIFICATION IN THE NORTH AMERICAN TIGER BEETLES OF THE *CICINDELA SYLVATICA* GROUP: MORPHOLOGICAL VARIATION AND AN ECOPHYLOGEOGRAPHIC APPROACH

Daniel Paul Duran

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

In partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Biological Sciences

December, 2010

Nashville, Tennessee

Committee 1 (2006-2009)

Professor Daniel J. Funk - Advisor

Professor David McCauley - Chair

Professor Patrick Abbot

Professor Christopher Nice

Professor David Pearson

To my wife, Melissa, for her love, support, and understanding,

and

To my son, Chase, for his amazing energy and curiosity about the world

ACKNOWLEDGEMENTS

I am grateful to the current and former members of my dissertation committee, including Dr. Patrick Abbot, Dr. Dan Funk, Dr. Dave McCauley, Dr. Chris Nice, Dr. David Pearson, and Dr. Charles Singleton who have provided me with helpful guidance towards my research and professional goals. From 2004-2009 Dan Funk was my graduate advisor. He funded this project via a Vanderbilt Discovery Grant. Dan also made intellectual contributions to the studies in each of the following data chapters, and provided support as a mentor. He also provided valuable input and support during the writing of the dissertation. I am thankful to all members of the Ecology and Evolution group within the Department of Biological Sciences at Vanderbilt University. Each member of this small group of faculty, postdoctoral fellows, and graduate students provided inspiration and encouragement during my dissertation. Most notable was the help of fellow graduate students Scott Egan and Chris Brown. I would also like to especially thank my good friend and colleague Rodger Gwiazdowski, for his constant moral support and for generously giving his time to thoroughly read and provide valuable feedback on each of my chapters.

I would like to thank the National Park Service for granting collection permits (Olympic National Park and Glacier National Park), and for the field assistance and advice of Jerry Freilich (Olympic N.P.). Dave Brzoska, Todd Lawton, Mike Kippenhan, Barry Knisley, Jason Schmidt, and Richard Freitag provided invaluable information on collection localities and knowledge of species

natural history. My dissertation could not have been done without the thorough and impressive work of Tim Spanton, who revised the North American *Cicindela sylvatica* group in his Master's thesis. His research provided the framework for asking questions about the group. I'd also like to thank the many tiger beetle collectors who generously contributed specimens to this project. They spent considerable effort to obtain critical populations that strengthened the results immeasurably. They are: Todd Lawton, Mike Kippenhan, Barry Knisley, Jason Schmidt, Rodger Gwiazdowski, Wayne Steffens, Reggie Webster, Richard Waldrep, Chris Nice, Dave Brzoska, Randy Dzenkiw, Melanie Bergolc, Anita Stjernberg, Richard Freitag, Bruce Rodrigues, Paul Novak, James McCann, Tim Simmons, Doug Taron, Kris Simpson, Matt Brust, Andy Cline, Daniel F. Duran, Kelly Tindall, Tony Allen, Ladislav Bocak, Giff Beaton, John Stamatov, Tom Bentley, and Jonas King. Todd Lawton donated well over a hundred specimens from twenty localities throughout Canada and the northern US. Karl Kjer and Catherine Duckett supplied lab space and use of equipment for part of the project. April Brown, Mark Mandel, Ryan Waples, Austin Ayres, and Nuri Kodaman provided lab assistance. I would also like to thank the entire Duran family, Tina Calandrino, and Rodger Gwiazdowski for help with weather updates, road conditions, and assorted valuable miscellaneous assistance. This work was funded by Vanderbilt University's Discovery Grant program.

I could not have even begun to pursue this project without the constant support and help of my entire family. I would like to thank my parents, Dan and Genevieve, for their constant and unwavering encouragement throughout my life

in everything I have ever done. I would like to thank my brothers, Andrew and Mike for being the good friends that they are and for listening to me at any time that I needed them. Most importantly, I wish to thank my loving wife, Missy, and my beautiful son, Chase, who provide me balance and unending inspiration.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Page

CHAPTER I

INTRODUCTION

The distribution of extant species and their current population structure are best interpreted within the context of the impacts of historic events on ancestral populations. Towards this end, numerous studies over the last two decades have focused on the geographical distribution of genealogical lineages (reviewed in Avise 2000). This recently developed field is termed phylogeography (Avise *et al* 1987) and is concerned with the connection between microevolutionary population structure and genealogical history in a spatial context. The principal goals of phylogeographic studies are 1) to describe geographical patterns of genetic subdivision within species, and 2) to identify the underlying processes that create these patterns. With the advent of coalescent theory (Kingman 1982) and improved mathematical models and methodologies, it is now possible to perform more rigorous tests to reject or support specific hypotheses about ancient and contemporary causes of population structure (*e.g.* Templeton *et al* 1995; Templeton 1998, 2004; Knowles and Maddison 2002). Additionally, comparative studies of phylogeographic patterns across a wide range of species have supported prior biogeographic hypotheses, such as the locations of glacial refugia (*e.g.* Remington 1968; Avise 1996; Swenson and Howard 2004, 2005). Phylogeographic data contribute to our general knowledge of the context,

causes, and timing of speciation. Furthermore, phylogeographic studies are especially useful for revealing the existence and location of deep genealogical subdivisions within species, sometimes referred to as “evolutionarily significant units” or ESUs (Ryder 1986; Waples 1991; Moritz 1994). These ESUs can qualify as entities deserving of legal protection under the US Endangered Species Act (56 FR 58612-58618), underscoring the conservation application of these principles and techniques.

Phylogeographic studies typically emphasize the importance of abiotic factors, such as limited dispersal or past vicariance events, but biotic interactions (*i.e.* predation, competition) could also play a role in shaping or maintaining the contemporary geographic distributions of lineages (Tautz 2004). With the advent of ecological niche modeling (*e.g.* Nix 1986; Stockwell and Nobel 1992; Phillips *et al* 2006; Elith *et al* 2006), it is possible to examine both the exogenous and endogenous factors which create and maintain the ranges of species, and hybrid zones between species (Kohlmann *et al* 1988; Cicero 2004; Swenson 2006). Yet surprisingly, ecological niche modeling has only recently begun to emerge as a tool for use in conjunction with phylogeographic studies (Hugall *et al* 2002; Graham *et al* 2004; Peterson *et al* 2004, Waltari *et al* 2007, Waltari and Guralnick 2009). Integration of phylogeographic and niche-modeling datasets could lead to a much clearer picture of the contributions of non-adaptive factors (*i.e.* neutral genetic drift due to population subdivision), and selection-based processes of ecological divergence in determining the distribution of species.

It has also been of great interest to biologists to understand the relationship between genetic and phenotypic variation. Long before the use of molecular genetic data, naturalists had noted the existence of geographical patterns of phenotypic variation in plant and animal species. These geographic races or subspecies were commonly construed to be populations with reduced gene flow, or even incipient species that may become completely reproductively isolated in the future (*e.g.* Mayr 1942, 1963). In more recent years, it has become possible to test for concordance between taxonomic designations and patterns of molecular genetic variation, as phylogeographic studies make use of neutrally evolving molecular genetic markers to uncover the patterns of population subdivision (Avise *et al* 1987; Avise 2000). One of the surprising outcomes of many such studies is the identification of deep genealogical splits that are not reflected in obvious differences in phenotypic characters (*e.g.* Avise *et al* 1979; Vogler and DeSalle 1993, 1994; Bernatchez 1997), but other workers have found higher levels of concordance between molecules and subspecies boundaries (review in Phillimore and Owens 2006). Moreover, some species also exhibit significant phenotypic variation within populations, a pattern that could be a result of neutral allelic variation (*e.g.* Tan 1945), developmental plasticity (*e.g.* Nice and Fordyce 2006), local adaptation (*e.g.* Dudley 1996), or a combination of these factors. Thus patterns of the geographic distribution of polymorphism could be the result of historical sundering of species, local adaptation, environmental developmental plasticity, or some combination of

these. More thorough and integrative studies are needed to test the causes of phenotypic variation.

I propose to use tiger beetles in the North American *Cicindela sylvatica* group to test for the causes of polymorphism distribution over geographic ranges. Specifically, I will concurrently test hypotheses from the following questions:

What factors promote population differentiation and speciation?

What are the determinants of species ranges?

What are the principal causes of phenotypic variation?

Through analyses of multiple molecular, environmental/ecological, and morphological datasets I will develop an integrated picture of the factors driving phenotypic and genetic divergence in this group of tiger beetles. This approach may serve as a model for future phylogeographic studies of other organisms, and provide better details for management plans and conservation biology.

Study system

The tiger beetles (Coleoptera: Carabidae: Cicindelinae) are a group of generalist predatory insects that are cosmopolitan in distribution. Both adults and larvae are predaceous, and most species are diurnally active in open habitats, such as sand dunes, open fields, alkali flats, and patches of bare soil or

rock (Pearson 1988). Larvae are ambush predators that live in burrows in soil or rarely in other substrates such as rock crevices (Kaulbars and Freitag 1993b). The larvae lie in wait at the top of their burrows with their mandibles open and their heads and pronota flush with the ground surface (Knisley and Schultz 1997). When a small invertebrate comes within reach they fling their heads in the direction of the prey, grasp it with their mandibles, and drag the prey down into the burrow to feed. Larvae and adults typically occur in the same habitats, with only a few known exceptions (Knisley and Schultz 1997). Adult tiger beetles run on the ground after prey, capturing and killing them with their mandibles. Tiger beetle adults are among the fastest animals for their size, running at speeds of up to 2.5 m/s, or approximately 125 body lengths per second (Kamoun and Hogenhout 1996). Most species require 1-3 years to complete their lifecycle (Knisley and Schultz 1997), although a small number of species are known to take four or more years to reach adulthood, especially at high latitudes (Spanton 1988).

Holarctic species of *Cicindela* (*sensu strictu*) tiger beetles (family Carabidae: Cicindelinae) are well suited for phylogeographic studies, as this group includes *c*. 200 species (Wiesner 1999) that occur in diverse habitats throughout most of temperate North America and Eurasia. These regions have been extensively altered by climatic changes as a result of glaciation cycles (*e.g.* Hewitt 1996; Rowe *et al* 2004; Swenson and Howard 2005; Waltari and Guralnick 2009). The five North American species of tiger beetles in the *C. sylvatica* group are especially appropriate test organisms as their combined

distributions include most of the continent north of Mexico (Pearson *et al* 1997), including hypothesized glacial refugia for many North American species (Remington 1968; Swenson and Howard 2004, 2005) as well as extensive areas that were covered by glacial ice (CLIMAP 1981; Sibrava *et al* 1986; Mix *et al* 2001). These taxa form a monophyletic group (Vogler and Welsh 1997; Vogler *et al* 2005) that includes both wide-ranging and geographically restricted species (see Chapter II, Fig 2.1 for range maps and sampling). From an ecological perspective, the members of this group are remarkable for being able to tolerate cooler climates and habitats such as montane grasslands, openings in boreal forests, and alpine meadows (*C. longilabris, C. nebraskana*) (Spanton 1988; Schultz *et al* 1992) or semi-shaded forested areas (*C. patruela, C. sexguttata, C. denikei*) (Knisley *et al* 1990, Kaulbars and Freitag 1993a, b) unlike the high temperature and open habitats of most of the other 109 Nearctic tiger beetle species (Pearson *et al* 1997; Freitag 1999; Pearson *et al* 2006). Moreover, the species *C. longilabris* is extraordinary for its occurrence at extremely high latitudes and altitudes, including areas where no other tiger beetle species exist. *Cicindela longilabris* has an extensive geographic distribution that spans a latitudinal gradient from southern Arizona to north of the Arctic Circle (Spanton 1988), and an east-west distribution from the coast of Newfoundland to western Alaska (Knisley *pers. comm.* 2006). In the southwestern extent of its range, this species is also found at high elevations up to 3800 m, indicative of its tolerance of cooler climates (Schultz *et al* 1992).

Because of these and other characteristics, *Cicindela longilabris* is ideal for studying the nature of polymorphism and underlying causes of phenotypic variation. This species exhibits an extreme amount of phenotypic variation in color pattern across its range, which has led to the description of numerous subspecies (reviewed in Spanton 1988), some of which were given full-species rank by previous authors. Spanton revised the taxonomy, synonymizing many of the previous names, but recognized three subspecies based on linear discriminant function analyses of morphometric data (Sneath and Sokal 1973). Although there appears to be a geographical component to the phenotypic variation, there also exists considerable variation within each of these subspecies and within local populations in a number of geographical areas (*e.g.* ID, MT, AZ, WA). This variation may be a result of elevated variation in heritable traits, developmental plasticity in particularly variable environments, or an interaction of these factors.

Yet another source of taxonomic confusion is the relationship between the phenotypically variable *C. longilabris* and the nearly invariant, all-black *C. nebraskana* (Figure 1.1). In areas where the two putative species overlap in the foothills of the Rockies, many individuals exhibit intergrade morphological traits that may represent hybridization. Yet, interestingly, the two remain morphologically distinct in other areas of sympatry in the Sierra Nevada and Cascade ranges. This pattern suggests these two taxa may be incipient species, recently diverged species that may hybridize occasionally, or that they may simply represent polymorphism within a single species.

The *C. sylvatica* group includes species that have been suggested as candidates for listing under the Endangered Species Act (*C. patruela, C. denikei*), and four of the five species are considered to be species of special conservation concern, in at least part of their distributions (NatureServe 2009, http://www.natureserve.org/explorer. Accessed: May 2010). *Cicindela patruela* is of the highest conservation priority within the group, and is listed as "vulnerable" to "critically imperiled" in 17 states, and "possibly extirpated" in the remaining four states assessed. Consequently, characterizing the species limits and phylogeographic subdivisions (*e.g.* ESUs) in this group will have important conservation implications.

Phylogeography and historical demography

Significant associations between alleles and geography can be the result of current demographic factors, ancient events, or some combination of these. F-statistics (Wright 1969) have often been used to assess levels of population structure, as inferred by deviations of heterozygosity from null expectations. However, the observed population genetic structure may have little or nothing to do with current patterns of gene flow (Larson 1984; Templeton *et al* 1995). Instead, the spatial distribution of allelic variance may be the product of historical fragmentation events that could create confounding patterns if these factors are not identified. In particular, Quaternary glacial cycles appear to have had profound effects on species distributions and for creating population

fragmentation within species (*e.g.* Huntley and Webb 1989; Hewitt 1996; Klicka and Zink 1997; Knowles 2001; Rowe *et al* 2004; Burg *et al* 2005; Nice *et al* 2005; Carnaval and Bates 2007; Provan and Bennett 2008; Stewart *et al* 2010). The historical inferences of phylogeographic studies can be strengthened when analyzed in conjunction with other similar studies (*e.g.* "genealogical concordance") (Avise 1996, 2000). If multiple codistributed species in a region exhibit similar phylogeographic structuring, then this would strongly suggest that ancient environmental processes had a similar impact throughout the regional fauna. In addition to identifying the factors responsible for these patterns, another result is the identification of geographic areas of high genetic endemism (Ryder 1986; Avise 1987; Waples 1991; Dizon *et al* 1992). These are areas that could be considered a high priority for conservation. Consequently, the addition of more detailed phylogeographic studies will contribute to our understanding of basic evolutionary processes and can also have real-world applications.

Mitochondrial sequence data are typically used to make phylogeographic inferences, and they are often the only type of marker employed in these studies (reviewed in Avise 2000; Ballard and Whitlock 2004). However, due to non-recombination, all loci in the mitochondrial genome are effectively a single linked unit. The history of a single gene genealogy may not be representative of a species genealogy, and consequently species may be polyphyletic with respect to a single marker. There have been concerns about the use of a single molecular marker to infer the relationships of recently split biological species (*e.g.* Funk and Omland 2003; Maddison and Knowles 2006), because single genes

are prone to patterns of paraphyly and polyphyly resulting from biological processes such as incomplete lineage sorting, and introgression as a result of interspecific hybridization. Funk and Omland (2003) suggested a 'congeneric phylogeography' sampling approach in order to identify species polyphyly and to avoid misinterpretation of the causes of genetic variation that could potentially result from inadequate sampling and unrecognized para- and polyphyly. This approach requires the sampling of multiple individuals from each of several species in order to detect the presence of species-level polyphyly and to evaluate alternative hypotheses for observed patterns.

In addition to thorough sampling of populations, the use of multiple unlinked markers should produce a more representative picture of a "true" species tree as multiple genes are less prone to exhibit the idiosyncratic history of a single marker. Amplified Length Fragment Polymorphisms (AFLPs) (Vos *et al* 1995) can be used to conduct a genome-wide scan for patterns of total genomic divergence between closely related species or diverging populations within a species. Furthermore, recent comparisons of mtDNA gene trees and AFLP datasets have indicated that the use of mtDNA alone may not be representative of total genomic divergence and species boundaries. Frequently, mtDNA gene trees misidentify species and ESUs that were subsequently confirmed with whole-genome scans (Gompert *et al* 2006; Weisrock *et al* 2006; Forister *et al* 2008).

To investigate the phylogeographic structure and historical demography of the group, I have conducted a series of investigations that combined fieldwork

and multiple phylogenetic, population genetic, and coalescent-based analyses. I used the above 'congeneric phylogeography' approach to sample intensively from all species, subspecies, and notable variants within the North American *C. sylvatica* group. Phylogenetic analyses were conducted using Maximum Parsimony and Neighbor Joining algorithms as implemented in PAUP (Swofford 1999) and Bayesian tree reconstruction as implemented in MrBayes 3.3 (Ronquist and Huelsenbeck 2003). Subsequently, I used AMOVA (Excoffier *et al* 1992) and SAMOVA (Dupanloup *et al* 2002) to investigate the relative contributions of taxonomy, geography, and population structure to explain the genetic variation. I used coalescent simulations to address alternative hypotheses about the causes of species-level polyphyly, as implemented in MESQUITE (Maddison and Maddison 2009). In addition, I was able to obtain demographic estimates of effective population size using LAMARC (Kuhner 2006) and time since divergence using MDIV (Nielsen and Wakeley 2001). To further characterize the demographic patterns, I used DNASP (Librado and Rozas 2009) to estimate multiple population genetics statistics, which taken together were used to infer the existence and location of glacial refugia and patterns of recent demographic expansion.

Ecological niche modeling and validation of Pleistocene refugia

One fundamental question in evolutionary ecology concerns the ecological and environmental barriers that delimit species distributions. These borders can

be influenced by a complex interplay of evolutionary, ecological, and physiological processes (Cicero 2004). Separating these effects has been a challenge, but theoretical and computational advances have resulted in the development of ecological niche modeling (*e.g.* Nix 1986; Stockwell and Nobel 1992; Phillips *et al* 2006; Elith *et al* 2006) to identify the principal factors responsible for the creation and continuing maintenance of range boundaries. These GIS-based methods combine data from known point locations for species with key environmental variables to generate a fundamental niche profile and expected geographic distribution. Furthermore, statistical resampling methods can be used to determine the relative contribution of each variable to model predictive accuracy.

Ecological niche-modeling methods most robustly predict the fundamental niches of species when detailed and accurate distribution data exist for the taxa of interest. Fortuitously, dependable range maps exist for all North American species of *Cicindela* (Pearson *et al* 1997) largely as a result of the incredible popularity of these insects among collectors (Knisley and Schultz 1997). To date, the factors limiting *Cicindela* distributions have been attributed to the effects of dispersal (*e.g.* Kaulbars and Freitag 1993b) and as a consequence of species' strong affinities with specific habitats and soil types (*e.g.* Spanton 1988; Schincariol and Freitag 1991). In addition to these and other potentially important abiotic factors, the ecological process of competitive exclusion could also prevent species from co-occurring at fine spatial scales or over larger geographic areas (Connell 1980; Vitt *et al* 1999). Some studies have demonstrated that

competitive interactions could be important in shaping community structure (Pearson and Mury 1979; Pearson 1980; Pearson and Juliano 1991), and one study identified food availability as a potentially limiting resource (Pearson and Knisley 1985) in a mesic grassland habitat. Consequently, the potential exists for competitive exclusion to be a factor in delimiting ranges of species or populations in the *C. sylvatica* group.

An interesting application of ecological niche models is the validation of previously hypothesized glacial refugia (Waltari *et al* 2007; Waltari and Guralnick 2009). By generating an ecological niche model for a species, it is possible to locate potential distributional areas under contemporary climatic conditions. Once this model is created it can be compared to reconstructions of Last Glacial Maximum conditions (Hijmans *et al* 2005) in order to assess where a species could have persisted at that point in time. Comparing the LGM ranges to predictions of glacial refugia based on the molecular data could be incredibly informative.

To characterize the factors limiting species distribution I collected species locality data from published records, museum specimens, and my own locality database from prior fieldwork. These presence data were then used to generate ecological niche models using MAXENT (Phillips *et al* 2006) and I then used statistical resampling methods included in the program to characterize the variables that contributed to each species ecological niche. Species distribution models were mapped and compared to known distributions in order to test hypotheses concerning the biotic and abiotic factors limiting their ranges. The

extent of niche overlap between species was quantified using ENMTools (Warren *et al* 2008) to assess the extent of potential niche differentiation and competition. Lastly, I applied the Waltari, et al (2007) method described above to validate the existence of the glacial refugia identified with molecular data.

Causes of phenotypic variation

Tiger beetle taxonomy has generally been based on genitalic morphology, setal (hair) characters, body shape, and especially color pattern at the species level. *Cicindela* often display striking variation in their color patterns within and among species (Shelford 1917). These can vary both in ground color and in the extent and shapes of dorsal unpigmented areas (maculations). All dorsal colors in tiger beetles are "structural colors" created by the optical properties of the cuticle, and not the result of different pigments (Schultz and Rankin 1983a,b). The cuticle is laminated with alternating layers of melanin pigment and translucent epicuticle, and the distance between these layers largely determines the color reflected. In addition to the wavelengths reflected through these layers, the surface microsculpturing can also create a mosaic of microscopic patches that reflect different wavelengths, which pointillistically mix to create many of the brown and olive colors in *Cicindela* (Schultz and Bernard 1989). Some studies have demonstrated that color may be important in predator avoidance as a result of crypsis by resemblance to the substrates they frequent (Willis 1967; Schultz 1986, 1991). Other studies have demonstrated that extent of maculation in

Cicindela can have significant adaptive consequences for thermoregulation by facilitating heat transfer of the integument (Schulz and Hadley 1987; Acorn 1992; Hadley *et al* 1992). Consequently, the potential exists for natural selection to be operating on different phenotypes within populations of *C. longilabris.* However, the structure of the cuticular layers and the resulting colors have been found in some cases to partly reflect phenotypic plasticity in development as a consequence of such abiotic factors as varying levels of humidity or temperature (Shelford 1917; Schultz 1983). Because these experiments were carried out on only a handful of species under a small set of possible conditions, it has not been possible to say how representative these patterns are. Despite these issues, the taxonomy of *Cicindela* subspecies (and occasionally species) is often based on subtle differences in color and pattern (*e.g.* Willis 1968; Graves *et al* 1988; Schincariol and Freitag 1991), characters that are either potentially under strong selection, or possibly the result of developmental plasticity. In both cases, these characters may be inappropriate for inferences about systematic relationships; as such traits are prone to patterns of convergence and can result in 'polytopic subspecies' (Wilson and Brown 1953) that are not representative of evolutionarily meaningful entities. An improved understanding of the underlying causes for phenotypic variation would be valuable for accurate taxonomy in this popular and conservationally important group of insects.

To investigate the causes of phenoptyic variation, I photographed dried beetle specimens from throughout the range of *C. longilabris*, using high-resolution photographic Nikon DF-Fi1 camera and SMZ 1500 microscope.

Specimens were chosen to represent populations of *C. longilabris* comprising the three subspecies, “intergrade” forms (Spanton 1988) and phenotypically distinct or variable populations. In addition, populations of *C. nebraskana* and purported phenotypic “hybrids” were included in the dataset. Color and maculation pattern were quantified using Photoshop CS5. I then used JMP 9.0 to perform multiple regression analyses to determine if environmental variables were predictive of phenotypic characters within the group.

Through the use of an integrative approach I was able to concurrently address fundamental questions of evolutionary ecology in a tiger beetle study system. Together, the above analyses of molecular, environmental, and phenotypic data allowed for a more complete understanding of the forces acting upon population differentiation, divergence, and speciation. This approach could be a model for future studies.

References

Acorn, J. A. 1992. The historical development of geographic color variation among dune *Cicindela* in western Canada. In The Biogeography of Ground Beetles of Mountains and Islands. Noonan, G. E., G. E. Ball, and N. E. Stork (Eds.). Intercept Press, Andover UK.

Avise, J.C., Giblin-Davidson, C., Laerm, J., Patton, J.C. and Lansman, R.A. 1979. Mitochondrial DNA clones and matriarchal phylogeny within and among geographic populations of the pocket gopher, Geomys pinetis. PNAS 76: 6694-6698.

Avise, J.C. 1987. Identification and interpretation of mitochondrial DNA stocks in marine species. In H. Kumpf and E.L. Nakamura (eds.) Proceeding of the Stock Identification Workshop. pp. 105-136. Pub. National Oceanographic and Atmospheric Administration. Panama City, Florida.

Avise, J.C., J. Arnold, R. M. Ball, E. Bermingham, T. Lamb, J.E. Neigel, C. A. Reeb, and N. C. Saunders. 1987. Intrspecific phylogeography: The mitochondrial DNA bridge between population genetics and systematics. Annu. Rev. Ecol. Syst. 18: 489-522.

Avise, J.C. 1996. Toward a regional conservation genetics perspective. Phylogeography of faunas in the southeastern United States. In J.C. Avise and J.L. Hamrick (eds.) Conservation Genetics: Case Histories from Nature. Pp. 431-470. Chapman and Hall. New York.

Avise, J.C. 2000. Phylogeography: The history and formation of species. Harvard University Press., Cambridge, Mass.

Ballard, J.W. and M.C. Whitlock. 2004. The incomplete natural history of mitochondria. Molecular Ecology 13: 729–744.

Bernatchez, L. 1997. Mitochondrial DNA analysis confirms the existence of two glacial races of rainbow smelt Osmerus mordax and their reproductive isolation in the St. Lawrence River estuary (Quebec, Canada) Mol. Ecol. 6:73-83.

Burg, T.M., A.J. Gaston, K. Winker, and V.L. Friesen. 2005. Rapid divergence and postglacial colonization in western North American Steller's jays (Cyanocitta stelleri). Mol. Ecol. 14(12):3745-3755.

Byers, J.A. 2006. Analysis of Insect and Plant Colors in Digital Images Using Java Software on the Internet. Ann. Entomol. Soc. Am. 99:865-874.

Carnaval, A.C. and J.M. Bates, Amphibian DNA shows marked genetic structure and tracks Pleistocene climate change in northeastern Brazil, *Evolution* 61 (2007), pp. 2942–2957.

Cicero, C. 2004. Barriers to sympatry between avian sibling species (Paridae: Baeolophus) in local secondary contact. Evolution 58:1573-1587.

CLIMAP. 1981. Seasonal reconstructions of the Earth's surface at the last glacial maximum in Map Series, Technical Report MC-36. Boulder, Colorado: Geological Society of America.

Connell, J.H. 1980. Diversity and the coevolution of competitors, or the ghost of competition past. Oikos 35, 131-138.

Dizon, A.E., C. Lockyer, W.F. Perrin, D.P. Demaster, and J. Sisson. 1992. Rethinking the stock concept: A phylogeographic approach. Conservation Biology 6: 24-36.

Dudley, S.A. 1996. The response to differing selection on plant physiological traits: Evidence for local adaptation. Evolution 50: 103-110.

Dupanloup, I., Schneider, S., Excoffier, L. (2002) A simulated annealing approach to define the genetic structure of populations. Molecular Ecology 11(12):2571-81.

Elith, J., C.H. Graham, R.P. Anderson, et. Al 2006. Novel methods improve prediction of species distributions from occurrence data. Ecography 29(2):129-151.

Excoffier, L. P.E. Smouse, J.M. Quattro. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131: 479–491.

Forister, M.L., C.C. Nice, J.A. Fordyce, Z. Gompert, and A.M. Shapiro. 2008. Considering evolutionary processes in the use of single-locus genetic data for conservation, with examples from the Lepidoptera. J. Insect Conserv. 12: 37-51.

Freitag, R. 1999. Catalogue of the Tiger Beetles of Canada and the United States. NRC Research Press, Ottawa, Ontario, Canada

Funk, D.J. 1999. Molecular systematics of cytochrome oxidase I and 16S from *Neochlamisus* leaf beetles and the importance of sampling. Molecular Biology and Evolution 16: 67-82.

Funk, D. J. and K. E. Omland. 2003. Species-level paraphyly and polyphyly: Frequency, causes, and consequences, with insights from animal mitochondrial DNA. Annual Review of Ecological Systematics 34:397-423.

Gompert, Z., C.C. Nice, J.A. Fordyce, M.L. Forister, and A.M. Shapiro. 2006. Identifying units for conservation using molecular systematics: the cautionary tale of the Karner Blue butterfly. Mol. Ecol. 15(7):1759-1768.

Graham, C.H., S.R. Ron, J.C. Santos, C.J. Schneider, and C. Moritz. 2004. Integrating phylogenetics and environmental niche models to explore speciation mechanisms in dendrobatid frogs. Evolution 58: 1781-1793.

Graves, R.C. 1988. Geographicl distribution of the North Ameican tiger beetle *Cicindela hirticollis* Say. Cicindela 20(1):1-21.

Hewitt, G. M. 1996. Post-glacial re-colonization of European biota. Biological Journal of the Linnean Society 68: 87-112.

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. 2005. Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology 25: 1965-1978.

Hugall, A., C. Moritz, A. Mousalli and J. Stanisic. 2002. Reconciling paleodistribution models and comparative phylogeography in the Wet Tropics rainforest land snail *Gnarosophia bellendenkerensis* (Brazier 1875). Proc. Natl. Acad. Sci. 99:6112-6117.

Huntley, B. and T. Webb. 1989. Migration: species' responses to climatic variations caused by changes in the earth's orbit. Journal of Biogeography 16: 5-19.

Kaulbars, M.M., and R. Freitag 1993a. Foraging behaviour of the tiger beetle *Cicindela denikei* Brown (Coleoptera: Cicindelidae). Can. Field-Nat. 107:53-58.

Kaulbars, M.M., and R. Freitag 1993b. Geographic variation, classification, reconstructed phylogeny, and geographic history of the *Cicindela sexguttata* group (Coleoptera: Cicindelidae). Can. Entomol. 125:267-316.

Kingman, J.F.C. 1982. The coalescent. Stoch. Process. Appl. 13:235-248.

Klicka, J. and R.M. Zink. 1997. The importance of recent Ice Ages in speciation: A failed paradigm. Science 277: 1666-1669.

Knisley, C. B., T. D. Schultz and T. R. Hasewinkel. 1990. Seasonal activity and thermoregulatory behavior of *Cicindela patruela*. Ann. Entomol. Soc. Amer. 83: 911-915.

Knisley, C. B. and T. D. Schultz. 1997. The Biology of Tiger Beetles and a Guide to the Species of the South Atlantic States. Virginia Museum of Natural History, Martinsville, VA. 210 pp.

Knowles, L.L. 2001. Did the Pleistocene glaciations promote divergence? Tests of explicit refugia models in montane grasshoppers. Mol. Ecol. 10: 691-701.

Knowles, L.L. and W.P. Maddison. 2002. Statistical phylogeography. Mol. Ecol. 11:2623-2635.

Kamoun, S, and S.A. Hogenhout. 1996. Flightlessness and rapid terrestrial locomotion in tiger beetles of the *Cicindela* L. subgenus *Rivacindela* van Nidek from saline habitats of Australia (Coleoptera: Cicindelidae). Coleopterists Bulletin 50: 221-230

Kohlmann, B., H. Nix, and D.D. Shaw. 1988. Environmental predictions and distributional limits of chromosomal taxa in the Australian grasshopper Caledia captiva (F.). Oecologia 75: 483-493.

Kuhner, M.K. 2006. LAMARC 2.0: Maximum likelihood and Bayesian estimation of population parameters. Bioinformatics 22(6): 768-770.

Larson, A. 1984. Neontological inferences of evolutionary pattern and process in the salamander family Plethodontidae. Evol. Biol. 17:119-217.

Librado, P. and Rozas, J. 2009. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25:** 1451-1452 | doi: 10.1093/bioinformatics/btp187.

Maddison, W.P., and L.L. Knowles 2006. Inferring phylogeny despite incomplete lineage sorting. Syst. Biol. 55(1):21-30.

Maddison, W. P. and D.R. Maddison. 2009. Mesquite: a modular system for evolutionary analysis. Version 2.72 http://mesquiteproject.org

Mayr, E. 1942. Systematics and the origin of species. Columbia Univ. Press, New York, NY.

Mayr, E. 1963. Animal species and evolution. Harvard University Press. Cambridge, Mass.

Mix AC, Bard E, Schneider R. 2001. Environmental processes of the ice age: land, oceans, glaciers (EPILOG). Quaternary Science Reviews 20: 627-657.

Moritz, C. 1994. Defining “evolutionarily significant units” for conservation. Trends Ecol. Evol. 9: 373-375.

NatureServe. 2009. NatureServe Explorer: An online encyclopedia of life [web application]. Version 7.1. NatureServe, Arlington, Virginia. Available http://www.natureserve.org/explorer. (Accessed: May 2010)

Nice, C.C., N. Anthony, G. Gelembiuk, D. Raterman, and R. Ffrench-Constant. 2005. The history and geography of diversification within the butterfly genus Lycaeides in North America. Mol. Ecol. 14:1741-1754.

Nice, C. C. and J. A. Fordyce. 2006. How caterpillars avoid overheating: behavioral and phenotypic plasticity of pipevine swallowtail larvae. *Oecologia* 146:541-548

Nielsen, R. and J. W. Wakeley. 2001. Distinguishing Migration from Isolation: an MCMC Approach. Genetics 158: 885-896.

Nix, H.A. 1986. A biogeographic analysis of Australian elapid snakes. In: Atlas of elapid snakes of Australia. Australian Flora and Fauna Series No. 7 (R. Longmore, Ed.), pp.4-15. Australian Government Publishing Service, Canberra.

Pearson, D.L., and E.J. Mury 1979. Character divergence and convergence among tiger beetles (Coleoptera: Cicindelidae). Ecology 60:557-566.

Pearson, D.L. 1980. Patterns of limiting similarity in tropical forest tiger beetles (Coleoptera: Cicindelidae). Biotopica 12:195-204.

Pearson, D.L., and C.B. Knisley. 1985. Evidence for food as a limiting resource in the life cycle of tiger beetles (Coleoptera: Cicindelidae). Oikos 45:161-168.

Pearson, D. L. 1988. Biology of tiger beetles. Annual Review of Entomology 33: 123-147.

Pearson, D.L. and S.A. Juliano. 1991. Mandible length ratios as a mechanism for co-occurrence: Evidence from a world-wide comparison of tiger beetle assemblages (Cicindelidae). Oikos 61:223-233.

Pearson, D. L., T. G. Barraclough, and A. P. Vogler. 1997. Distibutional maps for North American species of tiger beetles (Coleoptera: Cicindelidae). Cicindela 20:33-84.

Pearson, D. L., C. B. Knisley and C. J. Kazilek. 2006. A field guide to the tiger beetles of the United States and Canada: Identification, natural history, and distribution of the Cicindelidae. Oxford Univ. Press, NY, 288 pp.

Peterson, A.T., E. Martinez-Meyer, and C. Gonzalez-Salazar. 2004. Reconstructing the Pleistocene geography of the Aphelocoma jays (Corvidae). Divers. Distrib. 10:237-246.

Phillimore and Owens. 2006. Are subspecies useful in evolutionary and conservation biology? Proc Royal Soc B 273:1049-1053.

Phillips, S.J., R.P. Anderson, and R.E. Schapire. 2006. Maximum entropy modeling of species geographic distributions. Ecological Modelling 190: 231-259.

Pons, J., T.G. Barraclough, J. Gomez-Zurita, A. Cardoso, D. Duran, S. Hazell, S. Kamoun, W.D. Sumlin, A.P. Vogler. 2006. Evolutionary species delineation for the DNA taxonomy of undescribed insects. Systematic Biology 55(4): 595-609.

Provan J., and K.D. Bennett. 2008. Phylogeographic insights into cryptic glacial refugia. Trends Ecol. Evol. 23: 564–571.

Remington, C.L. 1968. Suture-zones of hybrid interaction between recently joined biotas. pp. 321–428. In Evolutionary Biology (Eds.) T. Dobzhansky, M.K. Hecht, and W.C. Steere, Plenum Press, New York.

Rogers, A.R., and H. Harpending. 1992. Population growth makes waves in the distribution of pairwise genetic differences. Mol. Biol. Evol. 9:552–569.

Ronquist, F. and J. P. Huelsenbeck. 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19: 1572-1574.

Rowe, K.C., E.J. Heske, P.W. Brown, and K.N. Paige. 2004. Surviving the ice: Northern refugia and postglacial colonization./ Proc. Nat. Acad. Sciences. 101(28):10355-10359.

Ryder, O.A. 1986. Species conservation and the dilemma of subspecies. Trends Ecol. Evol. 1:9-10.

Schincariol, L.A., and R. Freitag. 1991. Biological character analysis, classification, and history of the North American *Cicindela splendida* group taxa (Coleoptera: Cicindelidae). Can. Entomol. 123:1327-1353.

Schultz, T.D. and M.A. Rankin. 1983a. The ultrastructure of the epicuticular interference reflectors of tiger beetles (*Cicindela*). J. Exp. Biol. 117: 87-110.

Schultz, T.D. and M.A. Rankin. 1983b. Development changes in the interference reflectors and coloration of tiger beetles (*Cicindela*). J. Exp. Biol. 117:111-117.

Schultz, T.D. 1986. Role of sctructural colors in predator avoidance by tiger beetles of the genus Cicindela (Coleoptera: Cicindelidae). Bull. Entomol. Soc. Am. 32:142-146.

Schultz, T.D. and G.D. Bernard 1989. Pointillistic mixing of interference colours in crypyic tiger beetles. Nature 337:72-73.

Schultz, T.D. 1991. Tiger Hunt. Natural History 100: 38-44.

Schultz, T. D. and N.F. Hadley. 1987. Structural colors of tiger beetles and their role in heat transfer through the integument. Physiological Zoology 60:737-745.

Schultz, T.D., M.C. Quinlan, and N.F. Hadley. 1992. Preferred body temperature, metabolic physiology, and water balance of adult *Cicindela longilabris*: A comparison of populations from boreal habitats and climatic refugia. Physiological Zoology 65(1): 226-242

Schultz, T.D. 1998. The utilization of patchy thermal microhabitats by the ectothermic insect predator, Cicindela sexguttata. Ecological Entomology 23: 444-450.

Shelford, V.E. 1917. Color and color pattern mechanism of tiger beetles. Illinois Biological Monographs 3:395-532.

Sibrava, V., Bowen, D.Q, and Richmond, G.M., 1986, Quaternary Glaciations in the Northern Hemisphere, Quaternary Science Reviews. vol. 5, pp. 1-514.

Sneath, P.H.A. and R.R. Sokal 1973. Numerical taxonomy. The principles and practice of numerical classification. W.H. Freeman and Company. San Francisco. 573 pp.

Spanton, T.G. 1988. The *Cicindela sylvatica* group: Geographical variation and classification of the Nearctic taxa and reconstructed phylogeny and geographical history of the species (Coleoptera: Cicindelidae). Quaestiones Entomologicae 24: 51-161.

Spellman, G.M., and J. Klicka. 2007. Phylogeography of the white-breasted nuthatch (Sitta carolinensis): diversification in North American pine and oak woodlands. Molecular Ecology 16(8): 1729-1740.

Stewart, J.R., A.M. Lister, I. Barnes, L. Dalen. 2010. Refugia revisited: individualistic responses of species in space and time. Proc R Soc B. 277: 661-671.

Stockwell, D.R.B., and I.R. Nobel 1992. Induction of sets of rules from animal distribution data: A robust and informative method of data analysis. Math. Comp. Simul. 32:249-254.

Swenson, N.G. and D.J. Howard. 2004. Do suture zones exist? Evolution 58(11): 2391-2397.

Swenson, N.G. and D.J. Howard. 2005. Clustering of contact zones, hybrid zones, and phylogeographic breaks in North America. Am. Nat. 166: 581-591.

Swenson, N.G. 2006. GIS-based niche models reveal unifying climatic mechanisms that maintain the location of avian hybrid zones in a North American suture zone. Journal of Evolutionary Biology19: 717-725.

Swofford, D. L. 1999. PAUP: phylogenetic analysis using parsimony, version 4:0b2a. Sinauer, Sunderland, MA.

Tan, C.C. 1945. Mosaic dominance in the inheritance of color patterns in the lady-bird beetle, *Harmonia axyridis*. Genetics 31:195-210.

Tautz, D. 2004. Phylogeography and patterns of incipient speciation. In Adaptive Speciation. Dieckmann, U, M. Doebeli, J.A.J. Metz, and D. Tautz (Eds.). Cambridge University Press, Cambridge, UK.

Templeton, A. R., E. Routman, and C. A. Phillips. 1995. Separating population structure from population history: a cladistic analysis of the geographical distribution of mitochondrial DNA haplotypes in the tiger salamander, *Ambystoma tigrinum*. Genetics 140: 767–782.

Templeton, A. R. 1998. Nested clade analyses of phylogeographic data: testing hypotheses about gene flow and population history. Mol. Ecol. 7:381–397.

Templeton, A. R. 2004. Statistical phylogeography: methods of evalutating and minimizing inference errors. Mol. Ecol. 13: 789-809.

Vitt, L.J., P.A. Zani, and M.C. Esposito. 1999. Historical ecology of Amazonian lizards: implications for community ecology. Oikos 87: 286–294

Vogler, A. P., and R. DeSalle. 1993. Phylogeographic patterns in coastal North American tiger beetles (*Cicindela dorsalis* Say) inferred from mitochondrial DNA sequences. Evolution 47: 1192-1202.

Vogler, A.P. and R. DeSalle. 1994. Evolution and phylogenetic information content of the ITS-1 region in the tiger beetle, *Cicindela dorsalis*. Mol. Biol. Evol. 11(3): 393-405.

Vogler, A.P. and A.Welsh 1997. Phylogeny of North American *Cicindela* tiger beetles inferred from multiple mitochondrial DNA sequences. Mol. Phylogenet. Evol. 6:321-338.

Vogler, A. P., A. Cardoso, T. G. Barraclough. 2005. Exploring rate variation among and within a densely sampled tree: Species level phylogenetics of North American tiger beetles (Genus *Cicindela*). Syst. Biol. 54(1): 4-20.

Vos,, P., R. Hogers, M. Bleeker, et al 1995. AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res. 23:4407-4414.

Waltari, E., R. Hijmans , A. Peterson , Á. Nyári , S. Perkins, and R. Guralnick. 2007. Locating Pleistocene Refugia: Comparing Phylogeographic and Ecological Niche Model Predictions. PLoS ONE 2(7): e563.

Waltari, E. and R. Guralnick. 2009. Ecological niche modeling of Great Basin montane mammals: examining past and present connectivity of species across basin and ranges. Journal of Biogeography 36(1): 148-161

Waples, R.S. 1991. "Definition of 'Species' Under the Endangered Species Act: Application to Pacific Salmon." U.S. Department of Commerce. NOAA Technical Memorandum NMFS F/NWC-194.

Warren, D.L., R. E. Glor, and M. Turelli. 2008. Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. Evolution 62:2868-2883.

Weisrock, D.W., H.B. Shaffer, B.L. Storz, S.R. Storz, S.R. Voss. 2006. Multiple nuclear gene sequences identify phylogenetic species boundaries in the rapidly radiating clade of Mexican ambystomatid salamanders.

Wiesner, J. 1999. Verzeichnis der Sandlaufkäfer der Welt, Checklist of the Tiger Beetles of the World,. Verlag Erna Bauer, Keltern, pp.364.

Willis, H.L. 1967. Bionomics and zoogeography of tiger beetles of saline habitats in the central United States (Coleoptera: Cicindelidae). Univer. Kans. Sci. Bull. 47:145-313.

Willis, H.L. 1968. Artificial key to the species of Cicindela of North America north of Mexico (Coleoptera: Cicindelidae). J. Kansas Entomol. Soc. 41:303-317.

Wilson, E. O., and W. L. Brown. 1953. The subspecies concept and its taxonomic application. Syst. Zool. 2(3):97-111.

Wright , S. 1969. Evolution and the Genetics of Populations, Vol 2. The Theory of Gene Frequencies. University of Chicago Press, Chicago.

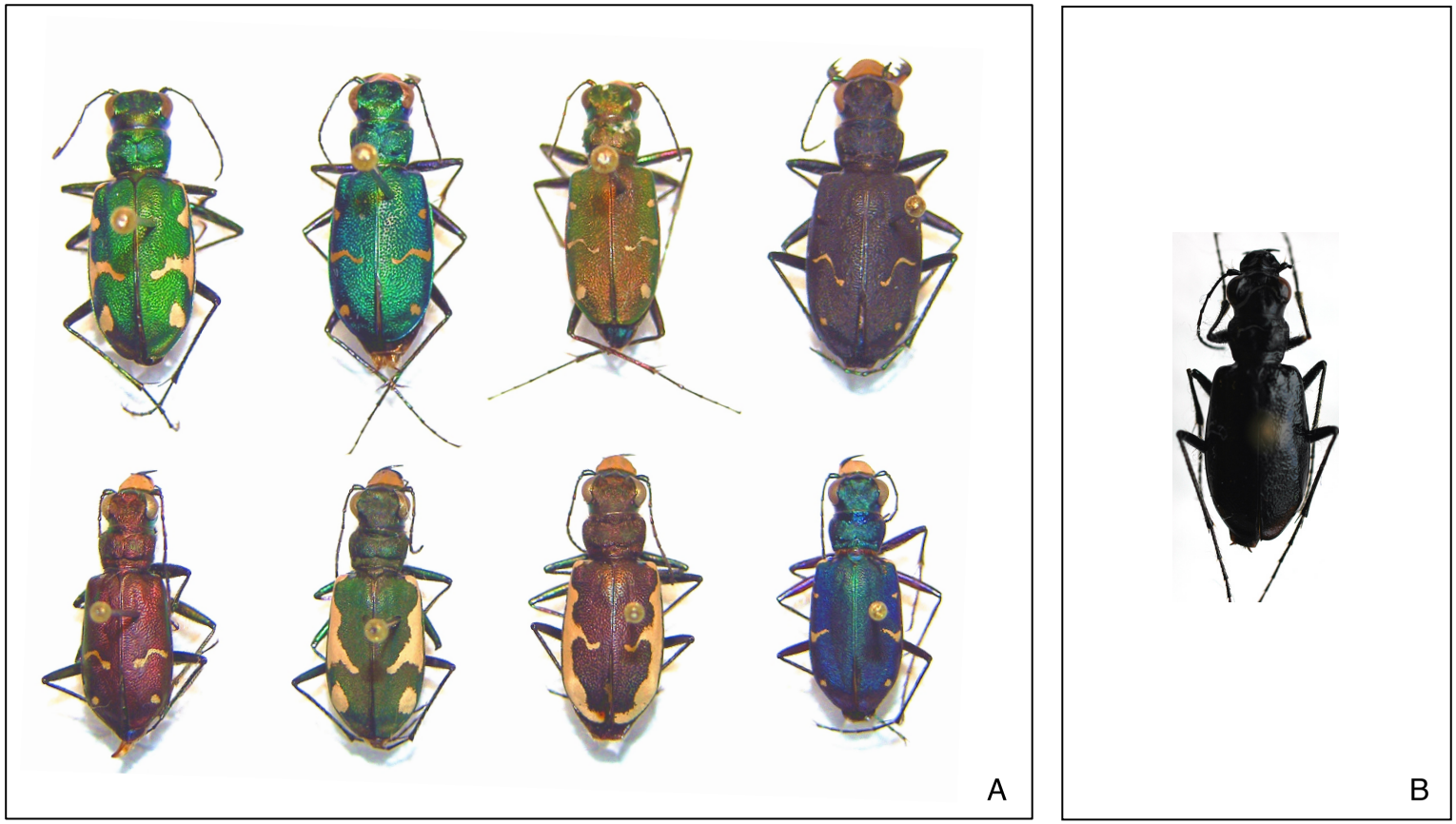


Figure 1.1. A) Examples of phenotypic variation within *Cicindela longilabris*. Localities are as follows: Top row, left to right: 1) MT: Ravalli Co.; 2,3) WA: Chelan Co., 4) Canada: QC: Duparquet; Bottom row, left to right: 5) MT: Granite Co., 6) CA: Mono Co., 7) UT: Iron Co., 8) CA: Nevada Co. **B)** Representative specimen of *C. nebraskana* from WA: Okanagon Co.

CHAPTER II

INFERRING SPECIES LIMITS USING A 'CONGENERIC PHYLOGEOGRAPHY' APPROACH: INSIGHTS FROM THE NORTH AMERICAN *CICINDELA SYLVATICA* GROUP TIGER BEETLES

Introduction

One of the primary goals of modern phylogeography is evaluating alternate hypotheses that can account for patterns of genetic variation among closely related populations and species. Understanding phylogeographic patterns within species can best be achieved through the broad sampling of closely related taxa that might share alleles for various reasons. However, phylogenetic and phylogeographic studies often implicitly assume that nominal species are biological species, or at least in practice many phylogeographic studies sample exclusively from a taxon of interest (*e.g.* review in Avise 2000; Peters *et al* 2005; Smith and Farrell 2005; Spellman and Klicka 2007). If the assumption of species monophyly is inaccurate then evolutionary interpretations of the data can be highly misleading (Funk 1999). In their review of the literature, Funk and Omland (2003) demonstrated that this assumption is often incorrect and documented that para- and polyphyly occurred in at least 23% of 2319

assayed species. Moreover, it would be impossible to empirically observe a pattern of polyphyly unless multiple individuals of at least two species are examined. To this point, Funk and Omland suggested the use of a 'congeneric phylogeography' approach to evaluate alternative hypotheses for observed patterns. This technique would also help avoid misinterpretation of the causes of genetic variation that could potentially result from inadequate sampling and unrecognized para- and polyphyly. Only by sampling multiple individuals from each of several species can the hypothesis of species-level monophyly actually be tested. In the event that species para- or polyphyly is revealed, additional hypothesis can be tested to determine the underlying cause of the incongruence between the nominal species and the gene tree. Ultimately, this approach could potentially result in a greater understanding of the biology of the organisms being examined (*e.g.* identification of a hybrid zone between species), or raise awareness about the need for re-evaluation of the taxon as a valid species.

The causes of species-level polyphyly (meaning any form of non-monophyly, including paraphyly) were reviewed by Funk and Omland (2003) and will be briefly described here. First, polyphyly can be an artifact of non-biological causes, such as inaccurate taxonomic species concepts that do not reflect actual biological species. This can be the result of "oversplitting", a situation where variants within a single species have been defined erroneously as separate nominal species. The converse of this situation is "lumping", the result of unrecognized distinct evolutionary lineages that are treated as a single taxonomic entity. Yet even when taxonomic concepts accurately reflect species

boundaries, polyphyly can occur as a consequence of the misidentification of particular specimens used in a study. Besides taxonomic issues, other methodological problems can result in observed polyphyly, such as inadequate phylogenetic information (*i.e.* weak statistical support for nodes) or the unrecognized paralogy of gene copies that are assumed to be orthologs.

Alternatively, species-level polyphyly can stem from biological causes reflecting the true allelic history of the taxa being studied. Recently diverged species are expected to share alleles as a result of retained polymorphisms present in the ancestral species prior to splitting. Consequently, nearly all sister species are likely to exhibit polyphyly for at least some portion of their history, and it is expected to take on the order of 4Ne generations before species are reciprocally monophyletic with respect to mitochondrial loci (Tajima 1983; Neigel and Avise 1986; Avise 1989; Takahata 1989; Avise and Ball 1990; Harrison 1991; Ballard and Whitlock 2004). This is a predicted amount of evolutionary time required in for a genealogy to accurately reflect the species tree (*i.e*. "lineage sorting"). However, the actual time necessary for complete lineage sorting may be affected by selection on the loci being examined. Positive selection can shorten the time to allelic fixation between species, whereas balancing selection can lengthen this time period or prolong it indefinitely. Another general cause of species polyphyly is interspecific hybridization where occasional interbreeding results in the transfer of alleles to hybrid offspring. When hybrids subsequently backcross into the parental populations the alleles from one species may introgress in the gene pool of another species, resulting in

observed polyphyly (Arnold and Bennett 1993; Shaw 1999, 2002; Chan and Levin 2005). Importantly, hybridization and incomplete lineage sorting can create similar polyphyletic phylogenetic patterns (Avise and Ball 1990; Holder *et al* 2001) and various methods have been developed that use information from reconstructed gene trees and coalescent simulations to tease apart the effects of each (Sang and Zhong 2000; Holder *et al* 2001; Kubatko 2009) although no single universally applied method has yet emerged. In some cases it may be possible to distinguish the effects of lineage sorting and hybridization based on other biological, temporal, or geographical information (*e.g.* Russell *et al* 2005; Buckley *et al* 2006; Joly *et al* 2006; McGuire *et al* 2007; Zinner *et al* 2009). To complicate matters, the two phenomena are not mutually exclusive and may be concurrently contributing to the observed polyphyly.

Undetected species-level polyphyly has several consequences for evolutionary inference and application to real world problems in taxonomy and conservation. First, if species are polyphyletic then the historical and demographic interpretations of phylogenetic patterns can be affected by the specific individuals chosen to represent species. Consequences include drastically inaccurate assessment of phylogenetic relationships, reconstruction of character evolution, and estimates of genetic divergence within and between species (Funk 1999). Second, species delineation and species identification are increasingly based on molecular data (*e.g.* Hebert *et al* 2003a, b; Tautz *et al* 2002; Pons *et al* 2006), and species-level polyphyly is especially a relevant issue as it relates to the increasingly popular method of “DNA barcoding” (Hebert *et al*

2003a, b). DNA barcoding aims to identify specimens solely through the use of a 600bp fragment of mitochondrial DNA. The methodology is founded on the assumption that individuals within species will form monophyletic groups to the exclusion of individuals from all other species. If this assumption is incorrect for a particular taxon then one possible result is a polyphyletic pattern, resulting in misidentification of the specimens of interest. Furthermore, phylogeographic studies are often used to identify the existence of major genetic subdivisions within species (reviewed in Avise 2000). Taken together, these issues underscore the importance of comprehensive sampling. A researcher is much more likely to identify the occurrence of species polyphyly by adopting Funk and Omland's (2003) proposed 'congeneric phylogeography' approach: intensively sampling populations of all closely related species of interest throughout the geographic ranges of all nominal species that have the potential to share genes.

Study system

The tiger beetles (Coleoptera: Carabidae: Cicindelinae) are a group of generalist predatory insects that are cosmopolitan in distribution. Both adults and larvae are predaceous, and most species are diurnally active in open habitats, such as sand dunes, open fields, alkali flats, and patches of bare soil or rock (Pearson 1988). Larvae are ambush predators that live in burrows in soil or rarely in other substrates such as rock crevices (Kaulbars and Freitag 1993b). The larvae lie in wait at the top of their burrows with their mandibles open and

their heads and pronota flush with the ground surface (Knisley and Schultz 1997). When a small invertebrate comes within reach they fling their heads in the direction of the prey, grasp it with their mandibles, and drag the prey down into the burrow to feed. Adult tiger beetles run on the ground after prey, capturing and killing them with their mandibles. Larvae and adults typically occur in the same habitats, with only a few known exceptions (Knisley and Schultz 1997). Most species require 1-3 years to complete their lifecycle (Knisley and Schultz 1997), although a small number of species are known to take four or more years to reach adulthood, especially at high latitudes (Spanton 1988).

Tiger beetles in the North American *Cicindela sylvatica* L. group are ideal for an assessment of the congeneric phylogeographic approach. First, previous phylogenetic work has demonstrated that these species comprise a well-supported monophyletic clade to the exclusion of all other North American species (Vogler and Welsh 1997; Vogler *et al* 2005). This result indicates that thorough sampling should be extended to all members of this clade, as there is potential for gene flow or sharing of ancestral alleles. Moreover, reliable geographic range distributions are available for all species (Spanton 1988; Kaulbars and Freitag 1993b; Pearson *et al* 1997) making it possible to intensively sample the full geographic range of each nominal species (Figure 1). Furthermore, at present four of the five species are considered to be species of special conservation concern, in at least part of their distributions (NatureServe 2009, http://www.natureserve.org/explorer. Accessed: May 2010), especially the species, *C. patruela*, which is listed as “vulnerable” to “critically imperiled” in 17

states, and “possibly extirpated” in the remaining four states assessed. Consequently, characterizing the genetic boundaries and distinctiveness of these taxa will have important conservation implications.

The systematic relationships of these species have been dealt with in multiple revisions (Spanton 1988; Kaulbars and Freitag 1993b) and these revisions suggest hypotheses about the boundaries of nominal species and the potential for gene flow. It is generally accepted that the species *C. longilabris* Say and *C. nebraskana* Casey are distinct species based on differences in color pattern, morphometrics, and subtle ecological differences (Leffler and Pearson 1978; Spanton 1988). However, they appear to hybridize in some areas where the two ranges overlap in the foothills of the Rocky Mountains and parts of the northern Great Plains (Figure 1a, b) while remaining morphologically distinct elsewhere. Originally these two species were historically considered to be the only North American members of the Holarctic *C. sylvatica* species group (Rivalier 1950, 1954) and no other species were included in Spanton’s (1988) revision of the North America taxa. More recently, molecular data were used to create species level phylogenies for the genus (Vogler and Welch 1997; Vogler *et al* 2005) that placed *C. longilabris* and *C. nebraskana* as sister to the “*C. sexguttata* group” (*sensu* Kaulbars and Freitag 1993b), which includes the eastern species *C. sexguttata* F., *C. patruela* Dejean, and *C. denikei* Brown. The taxonomy of *C. patruela* and *C. sexguttata* has been generally less contentious and all modern workers have treated them as distinct species (Boyd 1982; Knisley and Schultz 1997; Freitag 1999; Wiesner 1999; Pearson *et al* 2006)

based on multiple distinct differences in morphology, behavior, and ecology, throughout all areas of sympatry. Less clear however, is the distinction between *C. sexguttata* and *C. denikei.* The latter was once considered a form of *C. sexguttata* (Brown 1934; Wallis 1961) but more recently recognized as a distinct species (Kaulbars and Freitag 1993a, b; Freitag 1999) based on subtle differences in morphology and adult and larval behavioral ecology. Notably, *C. denikei* occurs only in a restricted geographic area of Ontario, Manitoba, and Minnesota (Fig. 1c) and is "peripatric" (Mayr 1942) and non-overlapping in distribution with the more widespread *C. sexguttata*. Based on the prior body of work, alternative hypotheses have been erected to explain the systematic relationships of the group:

H Ia) *C. longilabris* and *C. nebraskana* are separate species that hybridize, at least in some areas of sympatry, explaining the existence of morphologically intermediate populations and individuals.

H Ib) *C. longilabris* and *C. nebraskana* are not distinct species, but instead are the result of intraspecific polymorphism.

H IIa) *C. patruela* and *C. sexguttata* are separate species that remain distinct in sympatry.

H IIb) *C. patruela* and *C. sexguttata* are not distinct species, but instead are the result of intraspecific polymorphism.

H IIIa) *C. sexguttata* and *C. denikei* are recently diverged species resulting from a peripatric speciation event.

H IIIb) *C. denikei* is an isolated population and phenotypic variant of *C. sexguttata* but not a distinct species.

I utilized a congeneric phylogeography approach to test these alternative systematic hypotheses, which to date have never been addressed using molecular data and modern evolutionary analyses. Using phylogenetic, population genetic, and coalescent-based methodologies. I examined the genetic relationships amongst the putative species and compared the resultant predictions for species polyphyly from each hypothesis. I evaluated both the biological and taxonomic interpretations of observed genetic patterns. Due to the fact that much of the species collective ranges encompass areas that were recently glaciated at the Last Glacial Maximum (CLIMAP 1981; Sibrava *et al* 1986*;* Mix *et al* 2001) it is plausible that these taxa could have diverged during the Quaternary ice ages, or at least been impacted as a result of population isolation into multiple refugia. Species that have diverged recently ($\leq 4N_e$ generations) are more likely to share alleles due to the effects of incomplete lineage sorting. Given these issues, I explicitly evaluated the effects of stochastic lineage sorting in phylogeographical hypothesis testing through the integrated use of multiple phylogenetic and coalescent methods. The results emphasize the importance of thorough geographic sampling of species,

especially in areas of geographic proximity or sympatry between taxa, and demonstrate the utility of this method in phylogeographic hypothesis testing.

Materials and Methods

Sampling

Specimens were sampled from populations representing a considerable portion of the ranges of all nominal species (Figure 1), including all recognized subspecies (Freitag 1999) and as many notable variants as possible. The sampling localities (Table 1) were based on published records (Leffler 1979; Spanton 1988; Kippenhan 1994), localities provided by other North American tiger beetle researchers and amateur collectors, and additional populations located through exploratory collecting of appropriate habitat by the author. Other populations and individuals were contributed by collectors. Beetles were captured using aerial nets or insecticidal sprays and preserved using $\geq$ 95% ethyl alcohol. Species identifications were based on morphological characters used in Willis's (1968) key and Spanton's (1988) revision. Specimens that could not be unambiguously assigned to either *C. longilabris* or *C. nebraskana* were referred to as "hybrids". These individuals often occurred in intermediate habitat types alongside other individuals that could be confidently assigned to one or both of

the nominal species. No putative hybrids were ever observed between *C. sexguttata* and *C. patruela*. Outgroup taxa included two species in closely related subgenera (Vogler *et al* 2005), *C. (Cicindelidia) punctulata* Olivier, *C. (Sopiodela) chinensis japonica* Thunberg, and three *Cicindela (sensu strictu)*, including the hypothesized closest relative of the North American taxa, the Palearctic *C. sylvatica* L. (Rivalier 1950, 1954), plus the Palearctic *C. sachalinensis* Morawitz, and the North American *C. tranquebarica* Herbst.

DNA extraction and sequence analysis

Genomic DNA was extracted from most specimens by separating the head + prothorax from the rest of the body then removing flight muscles with sterilized forceps. DNA was isolated from the muscle tissue using the protocol of the DNeasy DNA isolation kit (Qiagen Corp.). Upon removal of tissue for extraction, the voucher specimens were stored in 70% ethyl alcohol for eventual pinning and vouchering. A small fraction of samples were obtained from dried specimens and these were extracted by perforating the abdomens using small "minuten" pins; following this the entire specimens were placed in 1.5ml microcentrifuge tubes containing lysis buffer and soaked overnight in a 55^{o}C water bath. No specimens were destroyed in the DNA extraction process and the reassembled vouchers exist for each. Using combinations of several standard insect mtDNA primers (Simon *et al* 1994) initial sequence data was obtained and the following degenerate primers were designed with the program

Primer3 (http://www.bioinformatics.nl/primer3plus) for PCR amplification and sequencing: CicF1 5'-AAA GGA AAC ATT TGG TTC ATT (A/G)GG-3', and CicR2 5'-AGT CGA AGA GAT GGA AG(C/T) GC-3'. These primers were used to amplify a 1.1 kilobase fragment of the mitochondrial genes *cytochrome oxidase c* subunit I + tRNA + coII corresponding to positions 2212-3342 of the *Drosophila yakuba* sequence. PCR conditions consisted of an initial denaturation at 96^{o}C for 2 mins, then 10 cycles at 96^{o}C for 30 s, 46^{o}C for 30 s, 72^{o}C for 1 min, and an additional 30 cycles at 96^{o}C for 30 s, 48^{o}C for 30 s, 72^{o}C for 1 min. Negative controls were used in all PCR reactions and no amplifications were ever observed in these. The target fragments were sequenced using the Dyenamic Terminator Sequencing Kit (Amersham Biosciences) run out on a polyacrylamide gel using an MJ Basestation and analyzed with the program Cartographer (MJ Research, Waltham, MA). When available, at least 4 individuals were extracted and sequenced per population for use in the analyses. A maximum of up to 10 individuals were extracted and sequenced per population.

Sequences were edited using Sequencher 4.2 (GeneCorp) and unambiguously aligned, as no indels were present in the fragment. The alignment was verified by eye and trimmed to 972bp for all 523 taxa. MODELTEST version 3.06 was used to determine that the unpartitioned data best fit a GTR + I + Γ model of substitution. Subsequent population genetic analyses were carried out using the assumptions of this model and parameter values specified.

Phylogenetic and population genetic analyses

Due to the large size of the mtDNA dataset (523 individuals and over 500,000 bp of sequence data), I used DNASP 4.0 to identify unique haplotypes and remove identical sequences. Phylogenetic reconstruction was carried out on this haplotype dataset, utilizing both Bayesian likelihood and maximum parsimony analyses. The Bayesian analysis was performed with MRBAYES version 3.1 (Ronquist and Huelsenbeck 2003) implementing a GTR + I + Γ model of evolution. I ran four Markov chains for 20 million generations each, with sampling every 1000 generations, resulting in a sample of 20,000 trees. The first 25% of the trees were discarded as a burn-in and the remaining 15,000 trees were used to construct a 50% majority consensus tree. After the burn-in was discarded, plots of the generation vs. the log likelihood values resulted in a random distribution of points, indicating that the chains had reached stationarity (Ronquist and Huelsenbeck 2003) which was further corroborated by convergence diagnostics (Gelman and Rubin 1992). The maximum parsimony analysis was implemented in PAUP* (Swofford 1999) with 1000 bootstrap replicates used to assess statistical support for recovered clades.

Analysis of molecular variance (AMOVA; Excoffier *et al* 1992) was employed to examine the degree of congruence between species boundaries and genetic data. Following this, the program SAMOVA 1.0 (Spatial Analysis of Molecular Variance; Dupanloup *et al* 2002) was used to identify populations that are geographically continuous and maximally genetically differentiated from each

other by maximizing Φ_{CT}. By comparing the percent of genetic variation that was consistent with nominal species boundaries to that explained by SAMOVA identified populations, it was possible to quantify the relative contributions of taxonomy and geography towards explaining genetic variation in the group.

Demographic analyses and coalescent-based hypothesis testing

Stochastic lineage sorting can result in the sharing of alleles between species; even those that are no longer exchanging genes (Avise *et al* 1983; Pamilo and Nei 1988). Historically this has been problematic for evolutionary inferences, as the effects of lineage sorting are difficult to separate from ongoing gene flow. However, coalescent-based simulations can potentially distinguish these causes by modeling the effects of retained ancestral polymorphisms on gene trees. Because lineage sorting is a stochastic process that proceeds as a function of time and population size (Neigel and Avise 1986), it is highly probable that closely related species with large effective population sizes will continue to share alleles even after gene flow has stopped. Prior to running coalescent simulations of lineage sorting, it was necessary to obtain estimates of the evolutionary effective population size for these simulations. Since $\theta = 2Ne\mu G$, the effective population size (Ne) can be calculated by estimating the parameter θ from genetic data, when mutation rate (μ) and generation time (G) are known. Maximum likelihood estimates of θ were obtained using LAMARC (Kuhner 2006) starting with an initial value of θ based on the number of segregating sites

(Watterson 1975), a point estimate obtained from DNASP 4.0. Published ecological data suggest that all species in this group typically have 2-year generations (Shelford 1917; Spanton 1988; Kaulbars and Freitag 1993b) throughout most of their ranges, although *C. longilabris* was demonstrated to have generation times of up 3 years or possibly longer (Spanton 1988), and *C. nebraskana* is presumed to have similar phenology. Mutation rates for insect mtDNA have been estimated across taxa, with the average rate of 2% per million years per gene (Brower 1994) and studies of tiger beetle molecular evolution have suggested that similar rates exist in the subfamily (Barroclough and Vogler 2002).

Coalescent simulations were performed using the program MESQUITE 2.72 (Maddison and Maddison 2009) to assess the amount of time required for complete lineage sorting. First, population trees were constructed for each pair of species as described in Hypotheses I-III with differing branch lengths as a function of population size (ranging from 0.01Ne to 1Ne). I simulated 1000 gene trees constrained within each population tree using a coalescent process based on values of Ne as obtained above. Slatkin and Maddison's (1989) *s* statistic is a measure of the discordance between a gene tree and population tree and can be used to estimate gene flow or the time to lineage sorting under the assumption of no gene flow. I calculated *s* for each simulation, and graphed the distribution of *s* values for each population tree. This was compared to *s* from the observed tree to estimate the number of generations of lineage sorting with which the observed data is consistent under the no gene flow model. The expected time to complete

lineage sorting in the absence of gene flow is given by the distribution that is consistent with an *s* of 2.

The program MDIV uses a coalescent framework to calculate maximum likelihood estimates of the divergence time between populations under the assumption of no recombination. This was used to estimate the time of divergence parameter *T* between the principal mtDNA clades. Each estimate of divergence was based on three independent runs with different random seeds to ensure that simulations were run sufficiently long enough to approach convergence. All runs were based on 2,000,000 cycles with a burn-in of 200,000 cycles.

Multilocus AFLP analyses

Ninety-six specimens were selected for a comparative multilocus phylogeny based on amplified fragment length polymorphisms (AFLPs) and were chosen to include all species, subspecies, and all major clades and subclades recovered in the mtDNA phylogeny. In addition, this set included two positive controls and an outgroup (*C. sylvatica*). Genomic DNAs were quantified using NanoDrop 1000 (Thermo Scientific, Wilmington, DE) and aliquots were made to bring the total DNA to 150ng per 25 ul AFLP reaction. To generate AFLP data following Vos *et al* (1995), I used AFLP Core Reagent Kits (Invitrogen, Carlsbad, CA). After the pre-selective amplification, four selective amplification primer combinations were used to generate PCR products (Table 2) that were purified

using Sephadex (GE Healthcare, Piscataway, NJ). Each run plate contained samples arranged in a randomized design, so as to preclude any possible block effects. In addition, two samples were run twice on each plate, as a form of positive control. Products were sent to University of Arizona Genetics Core (Tucson, AZ) for initial fragment analysis. This raw data was analyzed using the program RawGeno (Arrigo *et al* 2009), as this allows for an objective and repeatable method for analyzing particularly large AFLP datasets. Most individuals contained at least 200-300 bands present (out of a total of 1252 loci), and those samples that had fewer were than 200 bands were removed from the final dataset. Diagnostics in the RawGeno program also indicated these same samples were of poor quality, so they were not used in any subsequent analyses. Fragments of length 100-500 bp were included in the analysis and the default settings and thresholds were used. Additional analyses were performed using different thresholds for band size ranges (150-500bp) and larger bin widths (2bp instead of 1.5bp) and the results were not significantly different than those obtained with default values. Presence versus absence of peaks for these anonymous loci was called automatically within RawGeno according to the noise thresholds set by the program.

AFLP data were analysed using the program STRUCTURE 2.3 (Pritchard *et al* 2000; Falush *et al* 2003, 2007; Hubisz *et al* 2009) to investigate population genetic structure. The model assumes there are *K* populations and samples are assigned to one of the populations by identifying clusters that minimize deviations from Hardy-Weinberg equilibrium and linkage disequilibrium. The

newest version program was updated to handle dominant data, such as AFLPs (Hubisz *et al* 2009). I used the model parameter that assumed admixture, as it is biologically defensible and shown to be the most flexible and robust assumption (Pritchard *et al* 2000; Hubisz *et al* 2009). Each run was performed using a burnin of 50,000 steps and 500,000 additional steps. Runs were repeated five times for each *K* and results were consistent, therefore additional runs were not necessary. There are multiple methods to estimate *K,* and Pritchard *et al* (2007) advise against simply choosing the value of K that maximizes log *P [data / model]*, and to instead "aim for the smallest value of *K* that captures the major structure in the data". I followed the Evanno *et al* (2005) method to estimate the point at which the improvement in log *P [D/M]* levels off with increasing *K.* Even this method is still an *ad hoc* approximation and the authors conclude that a reasonable biological interpretation should be a key factor in selecting the "correct" *K*.

In addition to the STRUCTURE analyses of the AFLP data, I also ran phylogenetic reconstructions of the same data. AFLP data is inherently distance-based (presence vs. absence) and consequently many reconstruction methods cannot be defensibly used to reconstruct such a tree. Although no model of evolution for AFLP data currently exists, the data does lend itself to a distance-based tree-reconstruction, such as Neighbor Joining (Satui and Nei 1987). NJ trees often approximate trees used by other reconstruction methods (Felsenstein 2004; Mihaescu *et al* 2009) and taken in conjunction with the STRUCTURE results, these may be informative as to the history of the focal taxa.

Results

Overall mtDNA and AFLP results

A total of 523 individuals were sequenced for a 973 bp fragment of mtDNA, yielding over 508,000 bp of sequence data. Phylogenetic analyses were run using the total dataset and using a reduced dataset that included only unique haplotypes (196). Topologies were entirely consistent for each tree and the strongly supported clades reported below were contained in either dataset. As a result, the more compact haplotype tree was used to illustrate results (Figure 2). The Bayesian and Maximum Parsimony trees recovered a monophyletic ingroup, and indicated that the Palearctic *C. sylvatica* and *C. sachalinensis* were the closest relatives of the North American *C. sylvatica* group. Within the ingroup, the first major phylogenetic split was between the members of the traditionally recognized North American *C. sylvatica* group (comprised of individuals of *C. longilabris* and *C. nebraskana*, hereafter referred to as the "Meadow Group" for their ecological association with alpine meadows and grasslands, and Kaulbars and Freitag's (1993b) *C. sexguttata* group, hereafter termed the "Forest Group" for their association with primarily forested habitats. The two groups are separated by a net genetic distance of 3.26% using the method of Edwards (1997) that corrects the absolute distance by accounting for the genetic variation present in the ancestral population, as implemented in DNASP. These groups

were estimated to have diverged on the order of 2.51 million years ago from the maximum likelihood estimate obtained with MDIV (Table 3). This is assuming an average generation time of 2 years for the ancestor, typical of most Holarctic *Cicindela* (*sensu strictu*) (reviewed in Knisley and Schultz 1997).

The final AFLP dataset contained 61 unique individuals from the Meadow Group, and 26 from the Forest Group, in addition to the outgroup, *C. sylvatica*. Positive controls (samples run twice on each plate) were recovered as sister to their same-sample counterparts in each phylogenetic tree, and similarly they were found to cluster together in STRUCTURE analyses. At $K = 3$, STRUCTURE identified populations corresponding to 1) the outgroup, 2) the Meadow Group, and 3) the Forest Group. This was consistent with the results of the mtDNA phylogeny. Evolutionary inferences in the two major groups were distinct and results will be described separately below.

Results for Meadow Group

Phylogenetic and population genetic analyses

Phylogenetic analyses revealed that the nominal species, *C. longilabris* and *C. nebraskana,* were extensively polyphyletic, and haplotypes were shared between the two species and the morphological “hybrids” (Figure 2.2). Despite lack of species-level monophyly, there was significant structure present in the tree, and both analyses revealed the existence of three strongly supported

mtDNA clades each separated by 1.46 – 1.97% corrected net distance in pairwise comparisons (Table 2.3). These clades diverged from each other between 1.3 and 2.4 million years ago, based on MDIV analyses. Multiple individuals of *C. longilabris, C. nebraskana*, and "hybrids" were present in each clade. The clades were each distributed throughout a specific geographic area, and as such they are named for their location: Continental, Northwest, and Southwest Clades (Figure 2.3). Furthermore, these clades are non-overlapping in distribution, even in areas of close geographic proximity. Only one sampling locality was comprised of more than one clade, the population at Fremont Co, Idaho. At this site, four of the five individuals fell within the Southwest Clade, and the remaining individual was recovered within the Northwest Clade. This population occurs in a location where all three mtDNA clades appear to converge geographically.

AMOVA revealed that only an exceptionally small fraction (0.04%) of the genetic variation was explained by nominal species boundaries and this was not significantly different from zero (Table 2.4A). In contrast, 90.27% of the genetic variation was partitioned among sampling localities demonstrating very high levels of population structure, with the remaining 9.69% attributable to within population variation. SAMOVA was used to identify populations that are spatially contiguous and genetically differentiated from each other by maximizing $\Phi_{CT.}$ I compared results from $K = 2$ onwards to determine the point at which there was a plateau in the increase of Φ_{CT} with the addition of more K populations. That point was reached at $K = 4$ while little increase in Φ_{CT} occurred with additional increase

in *K*. At *K* = 3, SAMOVA identified three populations that were also identical to the major mtDNA haplotype clades. At *K* = 4 another geographic clade emerged, which I termed the SW Continental Clade, a divergent subgroup nested within the Continental Clade. These regional groups explained 72.25% of the genetic variation. Taken together, the results indicate that taxonomy is a very poor predictor of genetic variation compared to geography and population structure.

Coalescent simulations

The observed species polyphyly may be the result of ongoing gene flow, some level of hybridization/introgression, or may be attributable to the presence of retained ancestral polymorphisms. Using coalescent simulations, I modeled the potential effects of lineage sorting on the data given an estimated evolutionary effective population size (Ne). LAMARC was used to estimate the parameter θ and from that it was calculated that Ne for the group was on the order of 4.1×10^5. Coalescent simulations implemented in MESQUITE showed that when branch lengths were scaled to 0.75Ne the distribution was consistent with an *s* of 2, meaning that in the absence of gene flow complete lineage sorting would require approximately 300,000 generations or 900,000 years, given a three year average generation time (Spanton 1988). Slatkin and Maddison's *s* was originally developed to estimate levels of gene flow based on discordance between gene trees and species trees. Here, Slatkin and Maddison's *s*, given the

observed tree, indicates 46 gene flow events between *C. longilabris* and *C. nebraskana* (Table 2.5).

AFLP analyses

A total of 61 individuals were included in the final AFLP analyses for the group, and the four primer pairs yielded 1252 anonymous loci (Table 2.2). The Neighbor Joining tree recovered two major clades where both contain a mix of *C. longilabris*, *nebraskana*, and "hybrid" individuals (Figure 2.4). Although the clades were not associated with taxonomic species, they did match a geographic break between the mtDNA clades (Figure 2.5). STRUCTURE analyses of $K = 2 – 10$ were run and $K = 2$ was most strongly supported as the "correct" number of populations, based on multiple lines of evidence. First, the Evanno *et al* (2005) method revealed a sharp decrease in the improvement in log *P [D/M]* after $K = 2$. Furthermore, at $K = 2$, most individuals display $\geq$90% genomic identity associated with one of the two populations, whereas at $K > 2$ no individuals genome is comprised of more than a fraction of any identified population (Figure 6). In addition, these two SRUCTURE populations represent clusters that are geographically contiguous and allopatric (Figure 7), perfectly congruent with the two clades recovered in the NJ tree. Most localities contain individuals that fall strongly into one of the STRUCTURE identified populations ($\geq$90% of their genome). A small number of localities contained individuals exhibiting more admixture, meaning >10% of their genomic identity derived from the other

population. These populations were located mostly near the geographic boundary separating the two populations.

Results for Forest Group

Phylogenetic and population genetic analyses

The phylogeny revealed species polyphyly in the Forest Group, although to a much lesser extent than in the Meadow Group (Figure 2.2). Two major clades were recovered and were separated by 0.81% corrected net sequence divergence. The first of these clades (termed Forest1) consisted of 83% (71) *C. sexguttata*, 8% (7) *C. denikei*, and 9% (8) *C. patruela*. The second clade (termed Forest2) consisted of 100% (22) *C. patruela* individuals. These two clades overlap geographically (Figure 2.8) and the Forest 2 Clade encompasses a much smaller geographic area that is contained entirely within the distribution of the more widespread Forest 1 Clade. Moreover, eight of the Forest Group sampling localities contain individuals belonging to both clades. A single haplotype (#162) was shared between the species *C. sexguttata* and *C. patruela* (Figure 2.2). This haplotype occurred in eight northeastern *C. sexguttata* populations and three of five individuals within a geographically proximate population of *C. patruela* from MA. Additionally, one haplotype was shared between *C. sexguttata* and *C. denikei* (#186). This was present in the two westernmost *C. sexguttata*

populations from the Black Hills (SD, WY) and shared with all but one of the individuals in the two *C. denikei* populations.

AMOVA revealed that 38.86% of the genetic variation was explained by taxonomic species boundaries (Table 4B). Population structure accounted for an addition 38.86% and the remaining 22.28% was attributable to within population variation. As above, SAMOVA was used to identify spatially and genetically differentiated populations. At $K = 3$ the increase in Φ_{CT} plateaued and these regional groups explained 52.18% of the genetic variation, although these groups did not correspond to the mtDNA clades or any non-overlapping geographic areas. Taken together, these results suggest that taxonomy accounts for a moderate amount of the mtDNA genetic variation, and is equivalent to that explained by population structure. Regional groups explain a similar proportion of the genetic variation, but did not appear to be particularly geographically meaningful.

Coalescent simulations

LAMARC estimates of the parameter θ revealed that Ne for the Forest Group was on the order of 1.2×10^5. MESQUITE coalescent simulations demonstrated that, in the absence of gene flow complete lineage sorting between *C. sexguttata* and *C. patruela* would require approximately 180,000 years (Table 2.5). The estimate of Slatkin and Maddison's *s* from our observed tree is consistent with seven gene flow events. Lineage sorting between *C. sexguttata*

and *C. denikei* should require on the order of 120,000 years, and the observed tree is consistent with four gene flow events.

In addition, I ran simulations with Ne estimated for each Forest Group species separately because 1) repeated field observations and published accounts suggest highly asymmetrical population sizes for these species, based on differences in ecological niche breadth, patchiness of habitat, and observed abundances of the species in their appropriate habitat (Knisley *et al* 1990; Kaulbars and Freitag 1993a, b; Schultz 1998), and 2) the two Forest clades more or less reflect species boundaries with a relatively small percentage of apparently introgressed haplotypes in the Forest 1 Clade. Estimated Ne was 1.1×10^5 for *C. sexguttata*, 3×10^4 for *C. patruela*, and 2.8×10^3 for *C. denikei*. Coalescent simulations based on unequal Ne's demonstrated that complete lineage sorting between *C. sexguttata* and *C. patruela* would require approximately 120,000 years, whereas lineage sorting between *C. sexguttata* and *C. denikei* would require approximately 7,700 years under the assumption of no gene flow.

AFLP Analyses

A total of 26 individuals were included in the final AFLP analyses for the group, and the four primer pairs yielded 1252 anonymous loci (Table 2.2). The Neighbor Joining tree recovered two major clades, corresponding to 1) *C. sexguttata* + *C. denikei*, and 2) *C. patruela* (Figure 9). Bootstrap analysis

revealed strong support, with 100% of 1000 replicates recovering both clades. No further geographic or taxonomic structuring was identified within these clades. STRUCTURE analyses of K = 2 – 10 were run and K = 2 was most strongly supported (Figure 10). As in the Meadow Group, the Evanno *et al* (2005) method revealed a sharp decrease in the improvement in log *P [D/M]* after K = 2. With two populations, the clusters matched exactly with the results of the NJ tree above. All individuals displayed 96-100% genomic identity associated with one of the two populations, demonstrating a clear lack of admixture between the two clusters. Higher K values did not contribute to any clustering into additional geographic or taxonomic groupings, and at K > 2 no individual's genome is comprised of more than a fraction of any additionally identified population.

Discussion

Species-level polyphyly

Patterns of species-level polyphyly were observed in all five taxonomic species within the North American *Cicindela sylvatica* group. Multiple processes can result in non-monophyletic species and among these, unrecognized paralogy (*e.g.* amplification of nuclear pseudogenes) is probably the least likely problem at low phylogenetic levels. Nuclear pseudogene copies of mitochondrial genes are

typically identifiable by the presence of at least one of several characteristics, including the occurrence of insertions/deletions or unusual rates of nucleotide substitution that differ from that of other similar protein coding gene copies (Kimura 1983; Zhang and Hewitt 1996; Bensasson *et al* 2001; Williams and Knowlton 2001). Our sequences contained no indels, and were translated and found to be consistent with typical nucleotide substitution patterns, with accumulation of substitutions occurring primarily in third positions of codons. As such, no evidence of paralogy was found.

Another cause of polyphyly is inadequate phylogenetic information resulting in poorly supported clades or a star-like unresolved phylogeny. In the mtDNA tree, both Bayesian and maximum parsimony analyses recover very similar tree topologies with deep phylogenetic structuring (Fig. 2.2). Moreover, the major clades are strongly supported statistically ($\geq$70 bootstrap and/or $\geq$0.95 posterior probabilities) demonstrating that the observed polyphyly is not a result of weak phylogenetic signal. In the AFLP-based NJ trees, bootstrap support was below 50% for the two main clades within the Meadow Group, however the topology was remarkably consistent with the STRUCTURE analysis and with the mtDNA tree (Figs 2.5, 2.7). In the Forest Group, statistical support for the AFLP NJ tree was very strong, with 100% bootstrap support for both major clades (Fig. 2.9). Given these results, the existence of polyphyly in the *C. sylvatica* group cannot be explained away as the result of phylogenetic uncertainty.

Lineage sorting vs. interspecific hybridization

Incomplete lineage sorting (ILS) and interspecific hybridization are two general causes of species level polyphyly that can reflect the true allelic history of the taxa being examined. I evaluated the potential impact of these processes using multiple lines of evidence from phylogenetic and coalescent-based analyses. First, the tree topology and distribution of polyphyly can be informative as to whether incomplete lineage sorting or gene flow may be occurring. If ILS is responsible for polyphyly then incongruent haplotypes would be expected to occur in phylogenetically basal star-like areas of a gene tree (Holder *et al* 2001). Alternatively, when these are positioned in derived positions in a structured tree, then hybridization and introgression is more compelling as a cause. Our mtDNA results show that within the Meadow Group there exists deep phylogenetic structure, and multiple individuals of both nominal species and their putative hybrids are recovered within each of these major clades (Fig. 2.2), a pattern inconsistent with ILS, but consistent with hybridization. Furthermore, individual haplotypes are shared between the species (Fig. 2.2, haplotype #'s 42-46, 95, 136-139, 149) an indication of very recent or ongoing hybridization. Similar patterns are observed in the Forest Group, where phylogenetic structuring occurs and the topology of polyphyly is also consistent with hybridization, but not ILS (Fig. 2.2), although hybridization appears to be much less frequent. In the Forest1 Clade there is a single incidence of allele sharing between *C. sexguttata* and each of the other two species (Fig. 2.2, haplotype #'s 162, 186).

The degree of incongruence between a species tree and a gene tree can be used to infer levels of gene flow, or alternatively to model the time required for lineage sorting to result in reciprocal monophyly of species (Slatkin and Maddison 1989; Maddison and Maddison 2009). Our coalescent simulations show that complete lineage sorting should take on the order of 300,000 generations between the Meadow Group species, *C. longilabris* and *C. nebraskana*. Assuming a generation time of three years (Spanton 1988) these species would have had to remain as isolated groups for 900,000 years for complete lineage sorting to have occurred (Table 2.5). No fossil evidence exists for the *C. sylvatica* group, however MDIV estimates of divergence between the three Meadow Group clades range from 1.3 to 2.4 million years (Table 2.3). Although these clades do not correspond to species boundaries, they demonstrate that there was relatively ancient subdivision and isolation within the group, predating the estimated amount of time for lineage sorting to occur. In the case of the Forest Group species, *C. sexguttata* and *C. patruela,* the expected time needed for lineage sorting to occur was estimated at 60,000 - 90,000 generations, depending on values of Ne used. Assuming a typical two-year life cycle (Knisley and Schultz 1997), lineage sorting should require approximately 120,000-180,000 years (Table 2.5). The phylogeny recovers nearly native clades for each of those species (Forest1 Clade is 83% sexguttata and the Forest2 Clade is 100% patruela) and MDIV estimated the time of splitting at 360,000 years before present, indicating that there should have been sufficient time for purging of shared ancestral polymorphisms between these two species. Finally,

estimates of the time for lineage sorting between *C. sexguttata* and *C. denikei* ranged from 7,700 – 120,000 years, depending on assumptions of unequal effective population sizes (Table 2.5). If *C. denikei* evolved *in situ* then divergence must have been very recent, less than 10,500 yrs before present, given that the species limited contemporary range was covered by the Wisconsonian ice during the Last Glacial Maximum (CLIMAP 1981; Sibrava *et al* 1986*;* Mix *et al* 2001). Sufficient time could have passed for complete lineage sorting, but only if the minimum estimate is correct.

It is important to note that these estimates on the impact of lineage sorting are based on the assumption of selective neutrality. If ancestral mtDNA haplotype diversity was maintained through balancing selection then time for lineage sorting could be lengthened, even indefinitely. Conversely, positive selection should hasten monophyly through the fixation of alleles. Empirical evidence suggests that positive selection is more frequently the cause of mtDNA non-neutrality in animals, including insects (MacRae and Anderson 1988; Garcia-Martinez *et al* 1998; James and Ballard 2000, 2003). *Wolbachia* bacteria commonly infect insect species (Werren and Windsor 2000) and one consequence of infection is a form of selective sweep on mtDNA (*e.g.* Gompert *et al* 2008). As a result of this purifying selection, only one or a few mtDNA haplotypes may persist in most populations, however this phenomenon may also promote introgression of mtDNA haplotypes across species boundaries (Nice *et al* 2009).

Imperfect taxonomy

I have discussed the phylogenetic and biological causes of species polyphyly, and the remaining explanation for discordance between species trees and gene trees is a failure of the taxonomic species definitions to reflect patterns of gene flow. Classically, species have been described based on morphological characteristics, implicitly assumed to represent differentiated evolutionary lineages. In the case of tiger beetles, many species differ by one or a handful of characters, identifiable with a dichotomous key (Wallis 1968), but rarely have these species boundaries been assessed in an evolutionary context (Morgan *et al* 2000). It is possible that some of the taxonomic species may be "oversplit", that is, that character differences are assumed to be representative of barriers to gene flow, when in fact they may simply represent polymorphism within a single species. The taxonomy of the North American *C. sylvatica* group has undergone multiple revisions (Spanton 1988; Kaulbars and Freitag 1993b) and workers have ascribed differing levels of confidence to the species boundaries in the group. The taxonomic implications will be discussed below in each account.

Species boundaries in the Meadow Group

The results can effectively rule out ILS as an explanation for species-level polyphyly in the Meadow Group. The remarkable extent of polyphyly exhibited in the mtDNA tree would suggest frequent hybridization or alternatively, extensive

gene flow due to inaccurately assessed species boundaries by previous workers. The multilocus AFLP tree and STRUCTURE clustering both strongly corroborate the results of the mtDNA tree, indicating that mito-nuclear discordance is not occurring. This effectively rules out the possibility of mitochondrial introgression in the face of otherwise genomically differentiated species. Taxonomy for the group has been somewhat contentious and species limits have never been conclusively understood (Wallis 1961; Leffler and Pearson 1979). Although some morphological characters appear to generally segregate with the taxonomic species, no single morphological character can completely reliably separate the two (Spanton 1988). Likewise the AMOVA results show that taxonomy explains only 0.04% of the genetic variation. Hypothesis Ia, that *C. longilabris* and *C. nebraskana* are separate species, can be rejected. All evidence is consistent with Hypothesis Ib, that these previously defined taxonomic entities are instead the result of intraspecific polymorphism, and not separate species.

Interestingly, both the mtDNA and AFLP analyses show deep phylogeographic structuring in the group, corresponding to broad geographic sections of the continent. Taken together with MDIV estimates of divergence times, these patterns indicate that deep historical sundering must have taken place during the Quaternary Ice Ages. STRUCTURE analyses further suggest that the secondary contact between these two groups is recent, as admixture levels are very low except in populations close to the contact boundary (Fig. 2.7). The geographic location of this boundary is consistent with well-established "suture zones" for hundreds of species of animals and plants (Remington 1968;

Svenson and Howard 2004, 2005), suggesting that the effects of Quaternary climate change effected the ancestral Meadow Group populations in a manner similar to many other North American species.

Species boundaries in the Forest Group

Although the Forest Group species were also found to be polyphyletic, the mtDNA phylogeny and multilocus AFLP trees differed in topology, resulting in different evolutionary inferences compared to the Meadow Group. In the case of *C. sexguttata* and *C. patruela* I found a pattern consistent with occasional hybridization and mtDNA introgression based on the fact that the multilocus AFLP phylogeny and STRUCTURE analysis recovered these two species are distinct (Fig. 2.9), whereas the mtDNA phylogeny contained apparently introgressed haplotypes in some areas of geographic contact (Fig. 2.2). Moreover, the direction of introgression is consistent with expectations based on the species demographics. *C. sexguttata* is widespread and abundant with an apparently broad ecological niche and the species inhabits nearly any deciduous or mixed forest opening or edge within its range (Kaulbars and Freitag1993a, b; Knisley and Schultz 1997; Schultz 1998). In contrast, *C. patruela* is typically at much lower population densities and ecologically more specialized into “barrens” habitat with a patchy distribution even within its preferred range (Knisley *et al* 1990; Kaulbars and Freitag 1993b; Knisley and Schultz 1997). The asymmetrical observed abundances are corroborated by the different LAMARC values of Ne

for both species. Consequently, it is much more likely during rare hybridization events that *C. sexguttata* would introduce haplotypes into the *C. patruela* gene pool. The expectation is that the introgressed haplotypes for the rarer species should show up backcrossed into the clade of the more common species (Avise *et al* 1987; Avise 1989), a pattern observed in our phylogeny (Fig. 2.2). Furthermore, the distribution of polyphyletic haplotypes was not consistent with the expectations of ILS (Holder *et al* 2001) and ample time existed for lineage sorting to complete, based on MDIV results (Table 2.5). Hypothesis IIb, that *C. patruela* and *C. sexguttata* are the result of intraspecific polymorphism, can be rejected. The alternative Hypothesis IIa, that the two are separate species that remain distinct in sympatry, is supported. These conclusions are consistent with the taxonomic treatment of all recent tiger beetle workers (Boyd 1982; Knisley and Schultz 1997; Freitag 1999; Wiesner 1999; Pearson *et al* 2006).

Also as a result of our congeneric phylogeography approach, I was able to identify the existence of occasional gene flow between *C. sexguttata* and *C. patruela*, which was surprising, given that morphologically intermediate specimens are not observed in areas of sympatry. Because the two species shared a specific haplotype (Fig. 2.2, #162) ancient hybridization can be ruled out as the cause. Even if hybridization is rare (MESQUITE simulations suggest seven gene flow events to explain discordance between the gene tree and species tree), it does occur, and this fact has important implications for 'DNA barcoding' and conservation (discussed below).

The cause of polyphyly between *C. sexguttata* and *C. denikei* is less clear. *C. denikei* haplotypes were found in a basal star-like area of the mtDNA phylogeny (Fig. 2.2). Similarly, the multilocus AFLP analyses recovered *C. denikei* individuals within clades and clusters that contained mostly *C. sexguttata*. Coalescent simulations determined a minimum time for ILS to occur at 7,700 years. Given the geological history of present day Manitoba, Ontario, and Minnesota, the entire 'pavement alvar' habitat inhabited by *C. denikei* (Kaulbars 1993a) was covered by the Wisconsonian ice sheet until approximately 10,500 years ago. Seven individuals were sampled for the species from two populations (Table 2.1), and it is possible that I may not have reached the sampling level sufficient to allow the coalescent simulations to converge on a more accurate estimate of Ne and time required for complete lineage sorting. Consequently, it is not possible to say conclusively whether sufficient time had passed for *C. sexguttata* and *C. denikei* to become reciprocally monophyletic.

It is important to note that *C. denikei* is peripatric (satellite distribution) with respect to the putative parental species, *C. sexguttata*, and as a result there are certain expectations regarding genetic signatures. To the degree that a parental species exhibits geographic substructure, and a peripherally speciating population is small and local, this population is predicted to initially possess a phylogenetically restricted subset of parental alleles (Funk and Omland 2003). Although relatively little phylogenetic structure exists overall in the Forest1 Clade there were sharing of alleles between the *C. denikei* populations and the geographically closest populations of *C. sexguttata* from the Black Hills of South

Dakota and Wyoming. It is also interesting to note that the Black Hills *C. sexguttata* populations may also be more or less disjunct from the rest of the species distribution as few individuals have ever been collected from the middle of South Dakota and Nebraska (Backlund and Weins *pers. comm*. 2005; Pearson *et al* 2006). As a result of this study, I intend to follow up by conducting a more thorough evaluation of morphological characters in these populations to re-evaluate the species limits and determine if they may in fact be more closely allied with *C. denikei*.

Lastly, the taxonomic distinctions between these two species are minimal and based on subtle differences in morphometrics and ecological preferences, and only recently was the latter upgraded to full species status (Kaulbars 1993b). It is possible that these observed differences are not due to fixed characters as a consequence of reproductive isolation, but instead the differences could be the result of relatively recent population isolation, or even phenotypic plasticity in the presence of subtly different environments. It is not possible to reject Hypothesis IIIa, that *C. sexguttata* and *C. denikei* are recently diverged species resulting from a peripatric speciation event, but it can be concluded that if speciation has occurred, then it was at the end of the Pleistocene and too recent to see unequivocally clear genetic signatures of the divergence.

Implications for DNA barcoding and insect conservation

DNA barcoding implicitly assumes that species are reciprocally monophyletic in their mitochondrial DNA, therefore If species polyphyly is occurring, the method will fail to identify such groups. Attempts have been made to barcode other insect species and validate the technique for identification or description of new species (*e.g.* Hebert *et al* 2004; Hajibabaei *et al* 2006; Janzen *et al* 2009; but see Elias *et al* 2007; Schmidt and Sperling 2008; Linares *et al* 2009). These have largely focused on butterfly (Lepidoptera) species and may have given misleading results due to the fact that mitochondrial introgression should be less frequent because female Lepidoptera are the heterogametic sex (Sperling 1993; Forister *et al* 2008). Haldane's rule predicts that the heterogametic sex should exhibit reduced viability as hybrids (Haldane 1922; Turelli and Orr 2000; Coyne and Orr 2004). Because mtDNA is maternally inherited, it is expected that mitochondrial introgression would be extremely rare in these taxa. Therefore, much of the barcoding success is based on a group of organisms that represents a poor test of the methodology. It is unlikely that DNA barcoding would work as well in other insects groups that do not have females as the heterogametic sex - no other insect orders besides Lepidoptera exhibit this property. Barcoding is likely to be attractive to those attempting to identify insects, due to the sheer number of species and lack of specialists who are able to identify them (Hebert *et al* 2003b; Will and Rubinoff 2003). However, the above results demonstrate that a DNA barcoding approach would have failed to

correctly identify all individuals and populations of tiger beetles, and none of the species examined were reciprocally monophyletic with respect to mtDNA. Importantly, that approach would have failed to identify populations and individuals of *C. patruela*, the most endangered species in the North American *C. sylvatica* group.

Conclusions

Congeneric phylogeography allows for a more rigorous testing of species boundaries than would otherwise be possible. The identification of species-level polyphyly can lead to a greater understanding of the evolutionary history of the study organisms and can result in new research questions. Had I undertaken a study of any of the five taxonomic species without sampling congeners, I could not have identified the existence and the extent of the species-level polyphyly. In spite of the fact that tiger beetles are one of the most well-studied non-pest insect groups (Knisley and Schultz 1997) and multiple taxonomic revisions and dozens of natural history accounts have been published on the group, I was able to make powerful new inferences about the species limits and evolutionary history using this approach. Moreover, I could not have accomplished this without the intensive sampling scheme. Although this requires the collection of large numbers of populations and samples within each, in the case of the Forest Group

species, *C. sexguttata* and *C. patruela*, the evolutionary history could easily been mischaracterized if I had not sampled intensively from all areas of sympatry between the two. I was also able to resolve a longstanding taxonomic dispute over the distinction between *C. longilabris* and *C. nebraskana*, using the congeneric approach in conjunction with multilocus data.

Acknowledgments

I would like to thank the National Park Service for granting collection permits (Olympic National Park and Glacier National Park), and for the field assistance and advice of Jerry Freilich (Olympic N.P.). Dave Brzoska, Todd Lawton, Mike Kippenhan, Barry Knisley, Jason Schmidt, and Richard Freitag provided invaluable information on collection localities and knowledge of species natural history. While most of this work was carried out in Daniel Funk's lab, Karl Kjer and Catherine Duckett supplied lab space and use of equipment for part of the project. April Brown, Mark Mandel, Ryan Waples, Austin Ayres, and Nuri Kodaman provided lab assistance. I would also like to thank the entire Duran family and Rodger Gwiazdowski for help with weather updates, road conditions, and assorted valuable miscellaneous assistance. This work was funded by a Vanderbilt University Discovery Grant to Dan Funk, and by the Department of Biological Sciences at Vanderbilt University.

References

Arnold, M. L. and Bennett, B. D. 1993. Natural Hybridization in Louisiana irises: genetic variation and ecological determinants. In: Harrison, R. G. (ed.) Hybrid Zones and Evolutionary Process, pp. 115-139. Oxford University Press, New York.

Arrigo, N., J.W. Tuszynski, D. Ehrich, T. Gerdes, and N. Alvarez. 2009. Evaluating the impact of scoring parameters on the structure of intra-specific genetic variation using RawGeno, an R package for automating AFLP scoring. BMC Bioinformatics 10: 33

Avise, J.C., J.F. Shapiro, S.W. Daniel, C.F. Aquadro, and R.A. Lansman. 1983. Mitochondrial DNA differentiation during the speciation process in Peromyscus. Molecular Biology and Evolution 1: 38-56.

Avise, J.C., J. Arnold, R. M. Ball, E. Bermingham, T. Lamb, J.E. Neigel, C. A. Reeb, and N. C. Saunders. 1987. Intrspecific phylogeography: The mitochondrial DNA bridge between population genetics and systematics. Annu. Rev. Ecol. Syst. 18: 489-522.

Avise, J.C. 1989. Gene trees and organismal histories: a phylogenetic approach to population biology. Evolution 43:1192-1208.

Avise J, Ball R. 1990. Principles of genealogical concordance in species concepts and biological taxonomy. Oxford Surv Evol Biol. 1(7):45–67.

Avise, J.C. 2000. Phylogeography: The history and formation of species. Harvard University Press., Cambridge, Mass.

Ballard, J.W. and M.C. Whitlock, The incomplete natural history of mitochondria, *Mol. Ecol.* 13 (2004), pp. 729–744.

Barroclough, T.G. and A. P. Vogler. 2002. Recent diversification rates in North American tiger beetles estimated from a dated mtDNA phylogenetic tree. Molecular Biology and Evolution 19(10): 1706-1716.

Bensasson, D., D.X. Zhang, D.L. Hartl, G.M. Hewitt. 2001. Mitochondrial pseudogenes: evolution's misplaced witnesses. Trends Ecol. Evol. 16:314-321.

Boyd, H. P. and Associates. 1982. Checklist of Cicindelidae: The Tiger Beetles. Plexus Publishing Inc., New Jersey. 31 pp.

Brower, A.V. 1994. Rapid morphological radiation and convergence in geographical races of the butterfly, *Heliconius erato*, inferred from patterns of mitochondrial DNA evolution. Proc. Natl. Acad. Sci. USA 91:6491-6495.

Brown, W.J. 1934. New species of Coleoptera, V. The Canadian Entomologist 66(1): 22-24.

Buckley, T.R., M. Cordeiro, D.C. Marshall, and C. Simon. 2006. Differentiating between Hypotheses of Lineage Sorting and Introgression in New Zealand Alpine Cicadas (*Magicicada dugdale*). Syst. Biol 55(3):411-425.

Chan and Levin, 2005 K.M. Chan and S.A. Levin, Leaky prezygotic isolation and porous genomes: rapid introgression of maternally inherited DNA, Evolution 59: 720–729.

CLIMAP. 1981. Seasonal reconstructions of the Earth's surface at the last glacial maximum in Map Series, Technical Report MC-36. Boulder, Colorado: Geological Society of America.

Coyne, J.A., and H.A.Orr. 2004. Speciation. Sinauer Associates. Sunderland, Mass. 545 pp.

Dupanloup, I., Schneider, S., Excoffier, L. (2002) A simulated annealing approach to define the genetic structure of populations. Molecular Ecology 11(12):2571-81.

Edwards, S.V. 1997. Relevance of microevolutionary processes for higher level molecular systematics. In: Mindell DP (ed). Avian Molecular Systematics and Evolution. Academic Press: New York. pp 251–278.

Elias, M., Hill, R. I., Willmott, K. R., Dasmahapatra, K. K., Brower, A. V. Z., Mallet, J. and Jiggins, C. D. 2007. Limited performance of DNA barcoding in a diverse community of tropical butterflies, *Proc Royal Soc B* 274: 2881-2889.

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

Excoffier, L. P.E. Smouse, J.M. Quattro. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131: 479–491.

Falush, D., M. Stephens, J.K. Pritchard. 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 164: 1567–1587.

Falush, D., M. Stephens, J.K. Pritchard. 2007. Inference of population structure using multilocus genotype data: dominant markers and null alleles. Molecular Ecology Notes 7(4): 574-578.

Felsenstein, J. 2004. Inferring Phylogenies. Sinauer Associates, Sunderland, Mass. 540 pp.

Forister, M.L., C.C. Nice, J.A. Fordyce, Z. Gompert, and A.M. Shapiro. 2008. Considering evolutionary processes in the use of single-locus genetic data for conservation, with examples from the Lepidoptera. J. Insect Conserv. 12: 37-51.

Freitag, R. 1999. Catalogue of the Tiger Beetles of Canada and the United States. NRC Research Press, Ottawa, Ontario, Canada

Funk, D.J. 1999. Molecular systematics of cytochrome oxidase I and 16S from *Neochlamisus* leaf beetles and the importance of sampling. Mol Biol Evol 16: 67-82.

Funk, D. J. and K. E. Omland. 2003. Species-level paraphyly and polyphyly: Frequency, causes, and consequences, with insights from animal mitochondrial DNA. AREES 34:397-423.

Garcia-Martinez, J., J.A. Castro, M. Ramon, A. Latorre, and A. Moya. 1998. Mitochondrial DNA Haplotype Frequencies in Natural and Experimental Populations of Drosophila subobscura. Genetics 149: 1377-1382.

Gelman, A. and D.B. Rubin. 1992. Inference from iterative simulation using multiple sequences, Statistical Science 7: 457-511.

Gompert, Z., M.L. Forister, J.A. Fordyce and C.C. Nice. 2008. Widespread mito-nuclear discordance with evidence for introgressive hybridization and selective sweeps in *Lycaeides*. Molecular Ecology 17:5231-5244.

Hajibabaei, M., D.H. Janzen, J.M. Burns, W. Hallwachs, and P.D.N. Hebert. 2006. DNA barcodes distinguish species of tropical Lepidoptera PNAS 103(4): 968-971.

Haldane, J.B.S. 1922. Sex-ratio and unisexual sterility in hybrid animals. J Genet 12:101–109.

Harrison, R. G. 1991. Molecular changes at speciation. Annual Review of Ecology and Systematics 22:281-308.

Hebert, P. D. N., Ratnasingham, S. and deWaard, J. R. 2003a. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. Proc. R. Soc. Lond. B 270(Suppl.), S96–S99.

Hebert, P. D. N., Cywinska, A., Ball, S. L. and deWaard, J. R. 2003b. Biological identifications through DNA barcodes. Proc. R. Soc. Lond. B 270, 313–321.

Hebert, P.D.N., E.H. Penton, J.M. Burns, D.H. Janzen, and W. Hallwachs. 2004. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator.* PNAS 101(41): 14812-14817.

Holder, M.T., J.A. Anderson, A.K. Holloway. 2001. Difficulties in detecting hybridization. Systematic Biology 50:978–82

Hubisz, M.J., D. Falush , M. Stephens, J.K. Pritchard. 2009. Inferring weak population structure with the assistance of sample group information. Molecular Ecology Resources 9(5): 1322-1332.

James, A.C. and J.W.O. Ballard. 2000. The expression of cytoplasmic incompatibility and its impact on population frequencies and the distribution of *Wolbachia* strains in *Drosophila simulans*. Evolution 54: 1661-1672.

Janzen, D. H. *and 44 authors*. 2009. 'Integration of DNA barcoding into an ongoing inventory of complex tropical biodiversity', Molecular Ecology Resources 9(s1): 1-26.

Joly, S., J.R. Starr, W.H. Lewis, and A. Bruneau. 2006. Polyploid and hybrid evolution in roses east of the Rocky Mountains. American Journal of Botany 93(3): 412-425.

Kamoun, S, and S.A. Hogenhout. 1996. Flightlessness and rapid terrestrial locomotion in tiger beetles of the *Cicindela* L. subgenus *Rivacindela* van Nidek from saline habitats of Australia (Coleoptera: Cicindelidae). Coleopterists Bulletin 50: 221-230

Kaulbars, M.M., and R. Freitag 1993a. Foraging behaviour of the tiger beetle *Cicindela denikei* Brown (Coleoptera: Cicindelidae). Can. Field-Nat. 107:53-58.

Kaulbars, M.M., and R. Freitag 1993b. Geographic variation, classification, reconstructed phylogeny, and geographic history of the *Cicindela sexguttata* group (Coleoptera: Cicindelidae). Can. Entomol. 125:267-316.

Kimura, M. 1983. The Neutral Theory of Molecular Evolution. Cambridge University Press, Cambridge.

Kippenhan, M. G. 1994. The Tiger Beetles (Coleoptera: Cicindelidae) of Colorado. Trans. Amer. Entomological Society 120(1):1-86.

Knisley, C. B., T. D. Schultz and T. R. Hasewinkel. 1990. Seasonal activity and thermoregulatory behavior of *Cicindela patruela*. Ann. Entomol. Soc. Amer. 83: 911-915.

Knisley, C. B. and T. D. Schultz. 1997. The Biology of Tiger Beetles and a Guide to the Species of the South Atlantic States. Virginia Museum of Natural History, Martinsville, VA. 210 pp.

Kubatko, L.S. 2009. Identifying Hybridization Events in the Presence of Coalescence via Model Selection. Systematic Biology 58: 478-488.

Kuhner, M.K. 2006. LAMARC 2.0: Maximum likelihood and Bayesian estimation of population parameters. Bioinformatics 22(6): 768-770.

Leffler, S.L. and D.L. Pearson. 1976. The tiger beetles of Washington. Cicindela 8:21-60.

Leffler, S. R. 1979. Tiger beetles of the Pacific Northwest (Coleoptera: Cicindelidae). Ph.D. Dissertation, University of Washington, Seattle. 731 pp.

Linares, M. C., Soto-Calderón, I. D., Lees, D. C. and Anthony, N. M. 2009. High mitochondrial diversity in geographically widespread butterflies of Madagascar: A test of the DNA barcoding approach Mol Phylog Evol 50(3): 485-495.

MacRae, A.F., and W.W. Anderson, 1988. Evidence of non-neutrality of mitochondrial DNA haplotypes in *Drosophila pseudoobscura*. Genetics 120:485-494

Maddison, W. P. and D.R. Maddison. 2009. Mesquite: a modular system for evolutionary analysis. Version 2.72 http://mesquiteproject.org

Mayr, E. 1942. Systematics and the origin of species. Columbia Univ. Press, New York, NY.

McGuire JA, Linkem CW, Koo MS, Hutchison DW, Lappin AK, Orange DI, Lemos-Espinal JA, Riddle BR, Jaeger JR. 2007. Mitochondrial introgression and incomplete lineage sorting through space and time: phylogenetics of crotaphytid lizards. Evolution 61:2879–2897

Mihaescu, R., D. Levy, and L. Pachter. 2009. Why neighbor-joining works. Algorithmica 54(1): 1-24.

Mix AC, Bard E, Schneider R. 2001. Environmental processes of the ice age: land, oceans, glaciers (EPILOG). Quaternary Science Reviews 20: 627-657.

Morgan, M., C. B. Knisley and A. P. Vogler. 2000. New taxonomic status of the endangered tiger beetle *Cicindela limbata albissima* (Coleoptera: Cicindelidae): Evidence from mtDNA. Annals of the Entomological Society of America 93: 1108-1115.

NatureServe. 2009. NatureServe Explorer: An online encyclopedia of life [web application]. Version 7.1. NatureServe, Arlington, Virginia. Available http://www.natureserve.org/explorer. (Accessed: May 2010)

Neigel, J.E. and J.C. Avise. 1986. Phylogenetic relationships of mitochondrial DNA under various demographic models of speciation. Pp. 515-534 in: Evolutionary Processes and Theory, E. Nevo and S. Karlin (eds.). Academic Press, New York.

Nice, C.C., Z. Gompert, M.L. Forister, J.A. Fordyce. 2009. An unseen foe in arthropod conservation efforts: The case of Wolbachia infections in the Karner blue butterfly. Biological Conservation 142: 3137-3146.

Pamilio, P. and M. Nei, 1988. Relationships between gene trees and species trees. Molecular Biology and Evolution 5: 568–583

Pearson, D. L. 1988. Biology of tiger beetles. Annual Review of Entomology 33: 123-147.

Pearson, D. L., C. B. Knisley and C. J. Kazilek. 2006. A field guide to the tiger beetles of the United States and Canada: Identification, natural history, and distribution of the Cicindelidae. Oxford Univ. Press, NY, 288 pp.

Pearson, D. L., T. G. Barraclough, and A. P. Vogler. 1997. Distibutional maps for North American species of tiger beetles (Coleoptera: Cicindelidae). Cicindela 20:33-84.

Peters, J.L., W. Gretes, K.E. Omland. 2005. Late Pleistocene divergence between eastern and western populations of wood ducks (*Aix sponsa*)

inferred by the ‘isolation with migration’ coalescent method. Mol Ecol 14:3407–3418

Pons, J., T.G. Barraclough, J. Gomez-Zurita, A. Cardoso, D. Duran, S. Hazell, S. Kamoun, W.D. Sumlin, A.P. Vogler. 2006. Evolutionary species delineation for the DNA taxonomy of undescribed insects. Systematic Biology 55(4): 595-609.

Pritchard, J.K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. Genetics 155(2):945-959.

Remington, C.L. 1968. Suture-zones of hybrid interaction between recently joined biotas. pp. 321–428. In Evolutionary Biology (Eds.) T. Dobzhansky, M.K. Hecht, and W.C. Steere, Plenum Press, New York.

Rivalier, E., 1950. Démembrement du genre Cicindela Linné (Travail préliminaire limité à la faune paléaretique). Rev. fr. Entomol. 17: 217–244.

Rivalier, E., 1954. Démembrement du genre Cicindela Linné. II. Faune américaine. Rev. fr. Entomol. 21: 248–268.

Ronquist, F. and J. P. Huelsenbeck. 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19: 1572-1574.

Russell, A. L., R. A. Medellín and G. F. McCracken, 2005. Genetic variation and migration in the Mexican free-tailed bat (*Tadarida brasiliensis mexicana*). *Molecular Ecology* 14(7): 2207-2222.

Ryder, O.A. 1986. Species conservation and the dilemma of subspecies. Trends Ecol. Evol. 1:9-10.

Sang T., Zhong Y. Testing hybridization hypotheses based on incongruent gene trees. Syst. Biol. (2000) 49:422–434.

Saitou, N., and M. Nei. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Molecular Biology and Evolution 4(4): 406–425.

Schultz, T.D. 1998. The utilization of patchy thermal microhabitats by the ectothermic insect predator, Cicindela sexguttata. Ecological Entomology 23: 444-450.

Schmidt, B.C. and F.A. Sperling. 2008. Widespread decoupling of mtDNA variation and species integrity in *Grammia* tiger moths (Lepidoptera: Noctuidae). Systematic Entomology 33: 613-634.

Shaw, K. L. 1999. A nested analysis of song groups and species boundaries in the Hawaiian cricket genus Laupala. Molecular Phylogenetics and Evolution 11:332-341.

Shaw, K. L. 2002. Conflict between mitochondrial and nuclear DNA phylogenies of a recent species radiation: what mitochondrial DNA reveals and conceals about modes of speciation in Hawaiian crickets. PNAS 99: 16122-16127

Shelford, V.E. 1917. Color and color pattern mechanism of tiger beetles. Illinois Biological Monographs 3:395-532.

Sibrava, V., Bowen, D.Q, and Richmond, G.M., 1986, Quaternary Glaciations in the Northern Hemisphere, Quaternary Science Reviews. vol. 5, pp. 1-514.

Simon, C., F. Frati, A. Beckenbach, B. Crespi, H. Liu, and P. Flook. 1994. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. Ann. Ent. Soc. Amer. 87:651-701.

Slatkin M, Maddison WP. A cladistic measure of gene flow inferred from the phylogenies of alleles. Genetics. 1989;123: 603-13.

Smith, C. I. and B.D. Farrell. 2005. Phylogeography of the longhorn cactus beetle *Moneilema appressum* Leconte (Coleoptera: Cerambycidae): Was the differentiation of the Madrean sky-islands driven by Pleistocene climate changes? Molecular Ecology. 14: 3049-3065.

Spanton, T.G. 1988. The *Cicindela sylvatica* group: Geographical variation and classification of the Nearctic taxa and reconstructed phylogeny and geographical history of the species (Coleoptera: Cicindelidae). Quaestiones Entomologicae 24: 51-161.

Spellman, G.M., and J. Klicka. 2007. Phylogeography of the white-breasted nuthatch (Sitta carolinensis): diversification in North American pine and oak woodlands. Molecular Ecology 16(8): 1729-1740.

Sperling, F.A.H. 1993. Mitochondrial DNA variation and Haldane's rule in the *Papilio glaucus* and *Papilio troilus* species groups. Heredity 71: 227–233

Swenson, N.G. and D.J. Howard. 2004. Do suture zones exist? Evolution 58(11): 2391-2397.

Swenson, N.G. and D.J. Howard. 2005. Clustering of contact zones, hybrid zones, and phylogeographic breaks in North America. Am. Nat. 166: 581-591.

Svenson, N.G. 2006. GIS-based niche models reveal unifying climatic mechanisms that maintain the location of avian hybrid zones in a North American suture zone. Journal of Evolutionary Biology19:717-725.

Swofford, D. L. 1999. PAUP: phylogenetic analysis using parsimony, version 4:0b2a. Sinauer, Sunderland, MA.

Tajima, F., 1983 Evolutionary relationship of DNA sequences in finite populations. Genetics 105: 437–460.

Takahata, N. 1989. Gene genealogy in three related populations: consistency probability between gene and population trees. Genetics 122: 957-966.

Tautz, D., Arctander, P., Minelli, A., Thomas, R. H. and Vogler, A. P. 2002. DNA points the way ahead in taxonomy. Nature 418: 479.

Turelli, M., and H.A. Orr. 2000. Dominance, epistasis and the genetics of postzygotic isolation. Genetics 154:1663–1679

Vogler, A.P. and A.Welsh 1997. Phylogeny of North American *Cicindela* tiger beetles inferred from multiple mitochondrial DNA sequences. Mol. Phylogenet. Evol. 6:321-338.

Vogler, A. P., A. Cardoso, T. G. Barraclough. 2005. Exploring rate variation among and within a densely sampled tree: Species level phylogenetics of North American tiger beetles (Genus *Cicindela*). Syst. Biol. 54(1): 4-20.

Vos,, P., R. Hogers, M. Bleeker, et al 1995. AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res. 23:4407-4414.

Wallis, J.B.1961. The Cicindelidae of Canada. University of Toronto Press, Toronto, Canada. 74 pp.

Watterson, G.A. 1975. On the number of segregation sites. Theoretical Population Biology 7: 256–276

Wiesner, J. 1999. Verzeichnis der Sandlaufkäfer der Welt, Checklist of the Tiger Beetles of the World,. Verlag Erna Bauer, Keltern, pp.364.

Will, K. and D. Rubinoff. 2003. Myth of the molecule: DNA barcodes for species cannot replace morphology for identification and classification. Cladistics 20, pp.47-55.

Williams, S.T., and N. Knowlton. 2001. Mitochondrial pseudogenes are pervasive and often insidious in the snapping shrimp genus Alpheus. Mol. Biol. Evol. 18(8):1484-1493.

Willis, H.L. 1968. Artificial key to the species of Cicindela of North America north of Mexico (Coleoptera: Cicindelidae). J. Kansas Entomol. Soc. 41:303-317.

Werren, J.H. and D.M. Windsor. 2000. *Wolbachia* infection frequencies in insects: evidence of a global equilibrium? Proc Royal Soc B 267: 1277-1285.

Zhang, D.X., and G.M. Hewitt. 1996. Nuclear integrations: challenges for mitochondrial markers. Trends Ecol. Evol. 11:247–251.

Zinner, D., L.F. Groeneveld, C. Keller, and C. Roos. 2009. Mitochondrial phylogeography of baboons (*Papio* spp.) – Indication for introgressive hybridization? BMC Evol Biol. (9): 83.

Table 2.1. Locality sampling information. Population identifiers correspond to the taxon labels used in the phylogeny (Figure 2.2). The sample size indicates the number of individuals that were sequenced from each population.

Popn ID	Locality	Species	Clade	Coordinates	Sample size	Collected by
Meadow Group						
ABbi	Alberta: Bindloss	*nebraskana*	Continental	50°53'N, 110°16'W	2	R. Dzenkiw
ABca	Alberta: Calgary	*nebraskana*	Continental	50°58'N, 113°55'W	6	R. Dzenkiw
ABfm	Alberta: Ft. McMurray	*longilabris*	Continental	56°46'N, 111°21'W	3	R. Dzenkiw
ABop	Alberta: Opal	*longilabris*	Continental	53°59'N, 113°13'W	4	T. Lawton
AKfa	Alaska: Fairbanks	*longilabris*	Continental	64°52'N, 147°49'W	6	C. Nice
AKko	Alaska: Koyukuk	*longilabris*	Continental	65°21'N, 157°09'W	1	B. Knisley
AKva	Alaska: Valdez	*longilabris*	Continental	61°26'N, 143°43'W	1	D. Brzoska
AZco1	Arizona: Coconino Co.(1)	*longilabris*	SW Cont.	35°19'N, 111°43'W	4	D. Duran
AZco2	Arizona: Coconino Co.(2)	*longilabris*	SW Cont.	36°30'N, 112°08'W	9	D. Duran
BCca	Brit. Columbia: Castlegar	*longi*, *nebr*, hybrids	Northwest	49°12'N, 117°27'W	10	T. Lawton
BCma	Brit. Columbia: Manning	*nebraskana*	Northwest	49°06'N, 120°40'W	4	T. Lawton
CAfr	California: Fresno Co.	*longilabris*	Southwest	37°16'N, 118°42'W	5	D. Taron
CAla1	California: Lassen Co.(1)	*longilabris*	Northwest	40°21'N, 121°00'W	10	D. Duran
CAla2	California: Lassen Co.(2)	*nebraskana*	Southwest	40°28'N, 120°56'W	10	D. Duran
CAne	California: Nevada Co.	*longilabris*	Southwest	39°18'N, 120°20'W	4	D. Duran
CAmo1	California: Mono Co.(1)	*longilabris*	Southwest	37°56'N, 119°15'W	4	D. Duran
CAmo2	California: Mono Co.(2)	*nebraskana*	Southwest	37°31'N, 118°12'W	1	P. Opler
CAsi	California: Siskiyou Co.	*longilabris*	Northwest	41°21'N, 122°14'W	4	D. Duran
CAte	California: Tehama Co.	*longilabris*	Northwest	40°24'N, 121°32'W	9	D. Duran
COep	Colorado: El Paso Co.	*nebraskana*	Continental	39°01'N, 104°44'W	2	J. Schmidt
COhi	Colorado: Hinsdale Co.	*longilabris*	Continental	37°59'N, 107°16'W	3	D. Duran
COla	Colorado: Larimer Co.	*longilabris*	Continental	40°42'N, 106°26'W	6	J. Schmidt, D.D.
COsj	Colorado: San Juan Co.	*longilabris*	Continental	37°44'N, 107°39'W	8	M. Bergolc
IDad	Idaho: Adams Co.	*longilabris*	Northwest	44°56'N, 116°23'W	4	D. Duran
IDbl	Idaho: Blaine Co.	*longilabris*	Northwest	43°49'N, 114°16'W	4	D. Duran
IDcar	Idaho: Caribou Co.	*nebraskana*	Northwest	42°20'N, 111°37'W	5	B. Knisley
IDcas	Idaho: Cassia Co.	*nebraskana*	Northwest	42°19'N, 113°37'W	10	D. Duran
IDfr	Idaho: Fremont Co.	*nebraskana*	SW(4), NW(1)	44°17'N, 111°30'W	5	M. Kippenhan

Table 2.1. (Continued)

Popn ID	Locality	Species	Clade	Coordinates	Sample size	Collected by
IDid	Idaho: Idaho Co.	*longilabris*	Northwest	45°53'N, 116°04'W	5	D. Duran
IDla	Idaho: Latah Co.	*longilabris*	Northwest	46°54'N, 116°38'W	4	D. Duran
MB19	Manitoba: Trail 19	*longi*, *nebr*	Continental	49°38'N, 96°16'W	2	T. Lawton
MBas	Manitoba: Ashern	*longilabris*	Continental	51°01'N, 98°11'W	2	T. Lawton
MBel	Manitoba: Elma	*longi*, *nebr*	Continental	49°51'N, 96°02'W	5	T. Lawton
MBfa	Manitoba: Fairford	*longi*, *nebr*	Continental	51°40'N, 98°36'W	6	T. Lawton
MBgr	Manitoba: Grand Rapids	*longilabris*	Continental	53°37'N, 99°17'W	2	T. Lawton
MBmr	Manitoba: Minago R.	*longilabris*	Continental	54°12'N, 99°10'W	2	T. Lawton
MBnt	Manitoba: North Two R.	*longilabris*	Continental	52°27'N, 99°03'W	3	T. Lawton
MBor	Manitoba: Orr Creek	*longilabris*	Continental	56°27'N, 99°03'W	4	T. Lawton
MBsa	Manitoba: Sandilands	*longi*, *nebr*	Continental	49°20'N, 96°19'W	7	T. Lawton
MByq	Manitoba: Yellow Quill	hybrids	Continental	49°41'N, 99°33'W	1	A. Stjernberg
MTbe	Montana: Beaverhead Co.	*nebraskana*	Northwest	44°33'N, 112°08'W	7	M. Kippenhan
MTcar	Montana: Carbon Co.	*nebraskana*	Continental	45°12'N, 109°21'W	1	D. Duran
MTcas	Montana: Cascade Co.	*longilabris*	Continental	46°51'N, 109°21'W	3	D. Duran
MTga1	Montana: Gallatin Co.(1)	*longilabris*	Continental	45°18'N, 111°26'W	5	D. Duran
MTga2	Montana: Gallatin Co.(2)	*longilabris*	Continental	45°28'N, 110°57'W	4	M. Kippenhan
MTga3	Montana: Gallatin Co.(3)	*nebraskana*	Continental	45°46'N, 110°59'W	4	M. Kippenhan
MTpo	Montana: Powell Co.	*longilabris*	Continental	46°51'N, 112°44'W	2	D. Duran
MTra	Montana: Ravalli Co.	*longilabris*	Northwest	46°30'N, 114°14'W	10	D. Duran
MTri	Montana: Richland Co.	*nebraskana*	Continental	48°08'N, 104°05'W	1	D. Duran
NBno	New Bruns.: Northumb.	*longilabris*	Continental	47°19'N, 65°25'W	4	R. Webster
NTgt	Newfound,: Gaff Topsails	*longilabris*	Continental	49°09'N, 56°44'W	1	B. Rodrigues
NMbe	New Mexico: Bernalillo Co.	*longi*, *nebr,* hybrids	Continental	35°12'N, 106°27'W	4	D. Duran
NVcl	Nevada: Clark Co.	*longilabris*	Southwest	36°18'N, 115°41'W	9	D. Duran
NVel	Nevada: Elko Co.	*nebraskana*	Northwest	40°36'N, 115°23'W	2	R. Gwiazdowski
ON17	Ontario: Hwy 17	*longilabris*	Continental	49°32'N, 92°07'W	2	T. Lawton
ONca	Ontario: Calstock	*longilabris*	Continental	49°44'N, 84°19'W	4	D. Duran
ONco	Ontario: Cochrane	*longilabris*	Continental	48°56'N, 80°59'W	4	D. Duran
ONge	Ontario: Geraldton	*longilabris*	Continental	49°41'N, 86°58'W	4	D. Duran

Table 2.1. (Continued)

Popn ID	Locality	Species	Clade	Coordinates	Sample size	Collected by
ONke	Ontario: Kenora	*longilabris*	Continental	49°51'N, 94°29'W	2	T. Lawton
ONte	Ontario: Terrace Bay	*longilabris*	Continental	48°48'N, 87°06'W	4	D. Duran
ONsi	Ontario: Sibley Peninsula	*longilabris*	Continental	48°21'N, 88°49'W	2	R. Freitag
ONsh	Ontario: Sherwood Lake	*longilabris*	Continental	49°45'N, 95°05'W	1	T. Lawton
ONst	Ontario: Stanley	*longilabris*	Continental	48°23'N, 89°35'W	3	D.D, R. Freitag
ONve	Ontario: Vermillion Bay	*longilabris*	Continental	49°51'N, 93°18'W	1	T. Lawton
ORcl	Oregon: Clackamas Co.	*longilabris*	Northwest	45°08'N, 121°42'W	4	M. Kippenhan
ORha	Oregon: Harney Co.	*nebraskana*	Northwest	42°45'N, 118°39'W	4	D. Duran
ORkl	Oregon: Klamath Co.	*longilabris*	Northwest	42°23'N, 122°12'W	4	D. Duran
ORli	Oregon: Linn Co.	*longilabris*	Northwest	44°24'N, 121°52'W	6	D. Duran
ORwh	Oregon: Wheeler Co.	*longilabris*	Northwest	44°28'N, 120°13'W	4	M.K., D.D.
QCch	Quebec: Chapais	*longilabris*	Continental	49°49'N, 75°02'W	1	D. Duran
QCdu	Quebec: Duparquet	*longilabris*	Continental	48°30'N, 79°09'W	4	D. Duran
QClt	Quebec: La Tuque	*longilabris*	Continental	48°03'N, 72°17'W	4	D. Duran
SDcu1	South Dakota: Custer Co.(1)	*longi*, *nebr*, hybrids	Continental	43°51'N, 103°24'W	4	D. Duran
SDcu2	South Dakota: Custer Co.(2)	*longi*, hybrids	Continental	43°45'N, 103°47'W	4	D. Duran
UTbe	Utah: Beaver Co.	hybrids	Southwest	38°20'N, 112°22'W	5	D. Duran
UTca	Utah: Cache Co.	hybrids	Southwest	41°57'N, 111°29'W	6	D. Duran
UTir	Utah: Iron Co.	*longilabris*	Southwest	37°34'N, 112°51'W	6	D. Duran
UTsj	Utah: San Juan Co.	*longilabris*	SW Cont.	38°29'N, 109°16'W	8	D. Duran
UTut	Utah: Utah Co.	hybrids	Southwest	39°53'N, 111°16W	4	D. Duran
UTwa	Utah: Wasatch Co.	hybrids	Southwest	40°29'N, 111°02'W	6	B. Knisley
VTes	Vermont: Essex Co.	*longilabris*	Continental	44°46'N, 71°46'W	7	R. Gwiazdowski
WAch	Washington: Chelan Co.	*longilabris*	Northwest	47°45'N, 121°02'W	4	D. Duran
WAcl	Washington: Clallam Co.	*longilabris*	Northwest	47°59'N, 123°31'W	9	D.D., J. Freilich
WAki	Washington: King Co.	*nebraskana*	Northwest	47°24'N, 121°25'W	4	D. Duran
WAok	Washington: Okanagon Co.	*nebraskana*	Northwest	48°43'N, 120°02'W	3	W. Steffens
WAsp	Washington: Spokane Co.	*nebraskana*	Northwest	47°54'N, 117°06'W	4	D. Duran
WIvi	Wisconsin: Villas Co.	*longilabris*	Continental	46°08'N, 89°35'W	1	D. Brzoska
WYca	Wyoming: Carbon Co.	*nebr*, hybrids	Continental	41°14'N, 105°25'W	4	D. Duran

Table 2.1. (Continued)

Popn ID	Locality	Species	Clade	Coordinates	Sample size	Collected by
WYcr	Wyoming: Crook Co.	hybrids	Continental	44°24'N, 104°09'W	4	T. Lawton
WYfr	Wyoming: Fremont Co.	*nebraskana*	Continental	42°21'N, 107°34'W	3	D. Duran
WYsh	Wyoming: Sheridan Co.	*longilabris*	Continental	44°46'N, 107°46'W	4	D. Duran
WYsu1	Wyoming: Sublette Co.(1)	*nebraskana*	Continental	43°22'N, 109°57'W	4	D. Duran
WYsu2	Wyoming: Sublette Co.(2)	*nebraskana*	Continental	43°17'N, 110°35'W	2	B. Knisley
Forest Group						
GAde	Georgia: Decatur Co.	*sexguttata*	Forest1	34°38'N, 85°06'W	1	G. Beaton
GAwa	Georgia: Walker Co.	*sexguttata*	Forest1	30°41'N, 83°11'W	2	D. Duran
ILco	Illinois: Cook Co.	*sexguttata*	Forest1	42°03'N, 88°10'W	1	T. Bentley
INbr	Indiana: Brown Co.	*sexguttata*, *patruela*	Forest1, Forest2	39°09'N, 86°13'W	4,2	D. Duran
INst	Indiana: Steubing Co.	*sexguttata*	Forest1	41°43'N, 85°02'W	4	D. Duran
LAbr	Louisiana: Baton Rouge Co.	*sexguttata*	Forest1	30°25'N, 91°07'W	2	A. Cline
LAna	Louisiana: Natchitoches Co.	*sexguttata*	Forest1	31°29'N, 93°00'W	1	A. Cline
MApl	Mass.: Plymouth Co.	*patruela*	Forest1, Forest2	41°51'N, 70°41'W	4	T. Simmons
MBfl	Manitoba: Falcon Lake	*denikei*	Forest1	49°43'N, 95°16'W	4	T. Lawton
MDal	Maryland: Allegheny Co.	*sexguttata*, *patruela*	Forest1, Forest2	39°43'N, 78°40'W	1,3	J. McCann
MNmo	Minnesota: Morrison Co.	*patruela*	Forest1	46°43'N, 94°11'W	1	W. Steffens
MObo	Missouri: Boone Co.	*sexguttata*	Forest1	38°58'N, 92°22'W	3	K. Simpson
MOsc	Missouri: Scott Co.	*sexguttata*	Forest1	37°03'N, 89°33'W	3	K. Tindall
MSla	Mississippi: Lafayette Co.	*sexguttata*	Forest1	34°22'N, 89°32'W	3	J. King
NByo	New Brunswick: York	*sexguttata*	Forest1	45°50'N, 66°44'W	4	R. Webster
NCmo	North Carolina: Moore Co.	*sexguttata*	Forest1	35°12'N, 79°21'W	1	D. Duran
NCwa	North Carolina: Walker Co.	*sexguttata*, *patruela*	Forest1, Forest2	36°23'N, 81°09'W	2,1	D. Duran
NEst	Nebraskana: Stanton Co.	*sexguttata*	Forest1	42°00'N, 97°03'W	1	M. Brust
NJat	New Jersey: Atlantic Co.	*sexguttata*	Forest1	39°22'N, 74°39'W	4	D.F. Duran
NJoc	New Jersey: Ocean Co.	*sexguttata*, *patruela*	Forest1, Forest2	39°49'N, 74°23'W	4,4	D. Duran
NJun	New Jersey: Union Co.	*sexguttata*	Forest1	40°38'N, 74°28'W	3	D. Duran
NYro	New York: Rockland Co.	*sexguttata*	Forest1	41°08'N, 74°10'W	3	D.D., J. Stamatov
NYul	New York: Ulster Co.	*sexguttata*, *patruela*	Forest1, Forest2	41°40'N, 74°21'W	1,4	P. Novak

Table 2.1. (Continued)

Popn ID	Locality	Species	Clade	Coordinates	Sample size	Collected by
OHpi	Ohio: Pike Co.	*sexguttata*, *patruela*	Forest1, Forest2	39°06'N, 82°48'W	4,4	D. Duran
ONsh	Ontario: Sherwood Lake	*denikei*	Forest1	49°45'N, 95°05'W	3	T. Lawton
PAca	Pennsylvania: Carbon Co.	*sexguttata*, *patruela*	Forest1, Forest2	40°54'N, 75°45'W	1,1	D. Duran
PAsc	Pennsylvania: Schuykill Co.	*patruela*	Forest2	40°44'N, 76°26'W	3	R. Waldrep
SDme	South Dakota: Meade Co.	*sexguttata*	Forest1	44°23'N, 103°29'W	4	D. Duran
TNda	Tennessee: Davidson Co.	*sexguttata*	Forest1	36°16'N, 86°54'W	4	D. Duran
TNse	Tennessee: Sevier Co.	*sexguttata*	Forest1	35°42'N, 83°31'W	1	D.D., M. Duran
TXba	Texas: Bastrop Co.	*sexguttata*	Forest1	30°13'N, 97°15'W	1	C. Nice
TXgo	Texas: Gonzales Co.	*sexguttata*	Forest1	29°35'N, 97°35'W	1	C. Nice
VAbo	Virginia: Botetourt Co.	*patruela*	Forest2	37°38'N, 79°59'W	1	B. Knisley
VTca	Vermont: Caledonia Co.	*sexguttata*	Forest1	44°21'N, 71°54'W	1	S. Egan, D. Funk
WIja	Wisconsin: Jackson Co.	*patruela*	Forest1	43o58'N, 90o03'W	1	D. Duran
WImo	Wisconsin: Monroe Co.	*patruela*	Forest1	44o00'N, 90o12'W	2	D. Duran
WYcr	Wyoming: Crook Co.	*sexguttata*	Forest1	44o29'N, 104o07'W	3	T. Lawton
Outgroups						
	United Kingdom: Dorset	*Cicindela sylvatica*		50°45'N, 2°29'W		T. Allen
	Japan: Mt. Hakusan	*Cicindela sachalinensis*		36°09'N, 136°46'E		L. Bocak
	Japan: Osaka	*Cicindela japonica*		34°41'N, 135°32'E		L. Bocak
	Canada: Ont.: Terrace Bay	*Cicindela tranquebarica*		48°48'N, 87°06'W		D. Duran
	USA: Wisc.: Monroe Co.	*Cicindelidia punctulata*		44°00'N, 90°12'W		D. Duran

Table 2.2. Primers used for AFLP analysis, and number of loci for each primer combination (in parenthesis) that were used for analyses.

Primer	(Sequence 5' – 3')
Preselective	
Eco + C	GACTGCGTACCAATTCC
Mse + C	GATGAGTCCTGAGTAAC
Selective	
Eco + CTC	GACTGCGTACCAATTCCTC
Eco + CAG	GACTGCGTACCAATTCCAG
Mse + CTG	GATGAGTCCTGAGTAACTG
Mse + CGA	GATGAGTCCTGAGTAACGA
Mse + CAA	GATGAGTCCTGAGTAACAA
Mse + CCT	GATGAGTCCTGAGTAACCT
Primer Combinations	
F (296)	Eco + CTC / Mse + CGA
G (331)	Eco + CTC / Mse + CAA
I (293)	Eco + CAG / Mse + CGA
J (332)	Eco + CAG / Mse + CAA

Table 2.3. Pairwise sequence divergences between major clades in the mtDNA tree. Pairwise comparisons corrected by accounting for the genetic variation present in the ancestral population (Edwards 1997). Divergence times estimated with MDIV.

Clade comparison	corrected pairwise sequence divergence	Time of divergence (95% credibility interval) in millions of years
Meadow Group / Forest Group	3.26	2.51 (1.96 – 3.05)
Meadow Group		
Northwest Clade / Continental Clade	1.97	1.73 (1.34 – 2.26)
Northwest Clade / Southwest Clade	1.95	1.78 (1.34 – 2.39)
Continental Clade / Southwest Clade	1.46	1.66 (1.32 – 2.38)
Forest Group		
Forest1 Clade (mostly *C. sexguttata*) / Forest2 (*C. patruela* only) Clade	0.81	0.51 (0.36 – 0.76)

Table 2.4A. AMOVA for the mitochondrial gene region COI/COII for the Meadow Group. (i) AMOVA for populations grouped according to species identifications based on morphological and ecological differences. (ii) AMOVA for populations grouped according to regions identified by SAMOVA to maximize F_{CT}

A. Source of variation	d.f.	Sum of squares	Variance component	% of total	*P* value
Among species	1	39.505	0.003	0.04	0.397
Among populations / within species	99	2815.084	7.580	90.27	<0.001*
Within populations	270	219.772	0.814	9.69	<0.001*

B. Source of variation	d.f.	Sum of squares	Variance component	% of total	*P* value
Among regional groups	2	2106.391	8.571	72.25	<0.001*
Among populations / within groups	93	1013.130	2.443	20.60	<0.001*
Within populations	304	257.734	0.848	7.15	<0.001*

Table 2.4B. AMOVA for the mitochondrial gene region COI/COII for the Forest Group. (i) AMOVA for populations grouped according to species identifications based on morphological and ecological differences. (ii) AMOVA for populations grouped according to regions identified by SAMOVA to maximize F_{CT}

A. Source of variation	d.f.	Sum of squares	Variance component	% of total	*P* value
Among species	2	92.340	1.500	38.86	<0.001*
Among populations / within species	38	180.921	1.500	38.86	<0.001*
Within populations	68	58.500	0.860	22.28	<0.001*

B. Source of variation	d.f.	Sum of squares	Variance component	% of total	*P* value
Among regional groups	2	132.854	2.080	52.18	<0.001*
Among populations / within groups	32	82.140	0.328	8.23	<0.001*
Within populations	74	116.767	1.578	39.59	<0.001*

Table 2.5. Results of MESQUITE coalescent simulations to evaluate the impact of lineage sorting and gene flow on observed polyphyly. The expected time required for complete lineage sorting (assuming no gene flow) is based on simulations of gene trees constrained within species trees using Slatkin and Maddison's (1989) *s*. The number of inferred gene flow events is based on the *s* value for the observed phylogeny.

Species comparison	Expected time for CLS (in years)	Number of gene flow events
C. longilabris / C. nebraskana	900,000	46
C. sexguttata / *C. patruela* (equal Ne's)	180,000	7
C. sexguttata / C. patruela (unequal Ne's)	120,000	7
C. sexguttata / C. denikei (equal Ne's)	120,000	4
C. sexguttata / C. denikei (unequal Ne's)	7,700	4

Figure 2.1. Species distributions and sampling localities. (**A)** *C. longilabris* known range distribution, with the three currently recognized subspecies (Freitag 1999) each shaded differently. Red dots indicate sampled populations; violet dots represent populations in which putative hybrids exist between *C. longilabris* and *C. nebraskana*. (**B)** *C.nebraskana* known range distribution with black dots indicating sampled populations; violet dots indicate putative hybrid populations as in Fig. 2.1A. (**C)** *C. sexguttata* and *C. denikei* known range distributions. Dark blue and light blue dots represent sampled populations for the two species respectively. (**D)** *C. patruela* known range distribution with the two currently recognized subspecies shaded differently. Green dots indicate sampled populations. Maps adapted from Pearson *et al.* (2006).

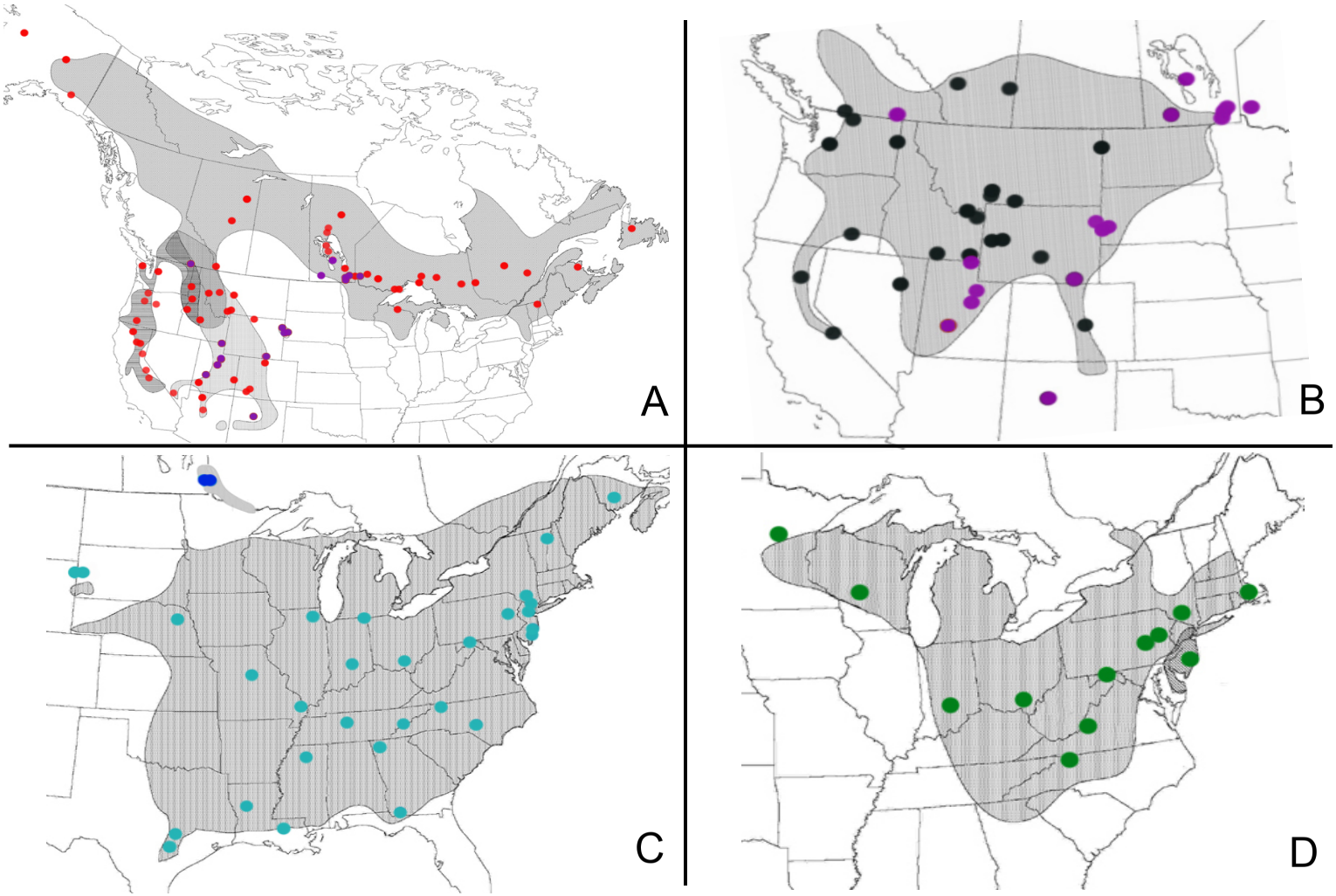

A
B
C
D

Figure 2.2 (part 1). Phylogenetic tree from Bayesian and Maximum Parsimony (MP) analyses. Support values indicate the Bayesian posterior probability of a node out of a possible 1.0 (*left value*), and the bootstrap percentage (out of 1000 replicates) of the MP analysis (*right value*). Numerical values are only indicated when there was strong statistical support for a node in either analysis ($\geq$ 0.95 / $\geq$ 70) or when a node was recovered in both analyses regardless of support values. Taxon IDs correspond to populations in Table 2.1.
***** The Continental Clade contains a subclade, the SW Continental Clade, a group of populations Identified as distinct in the SAMOVA analysis.

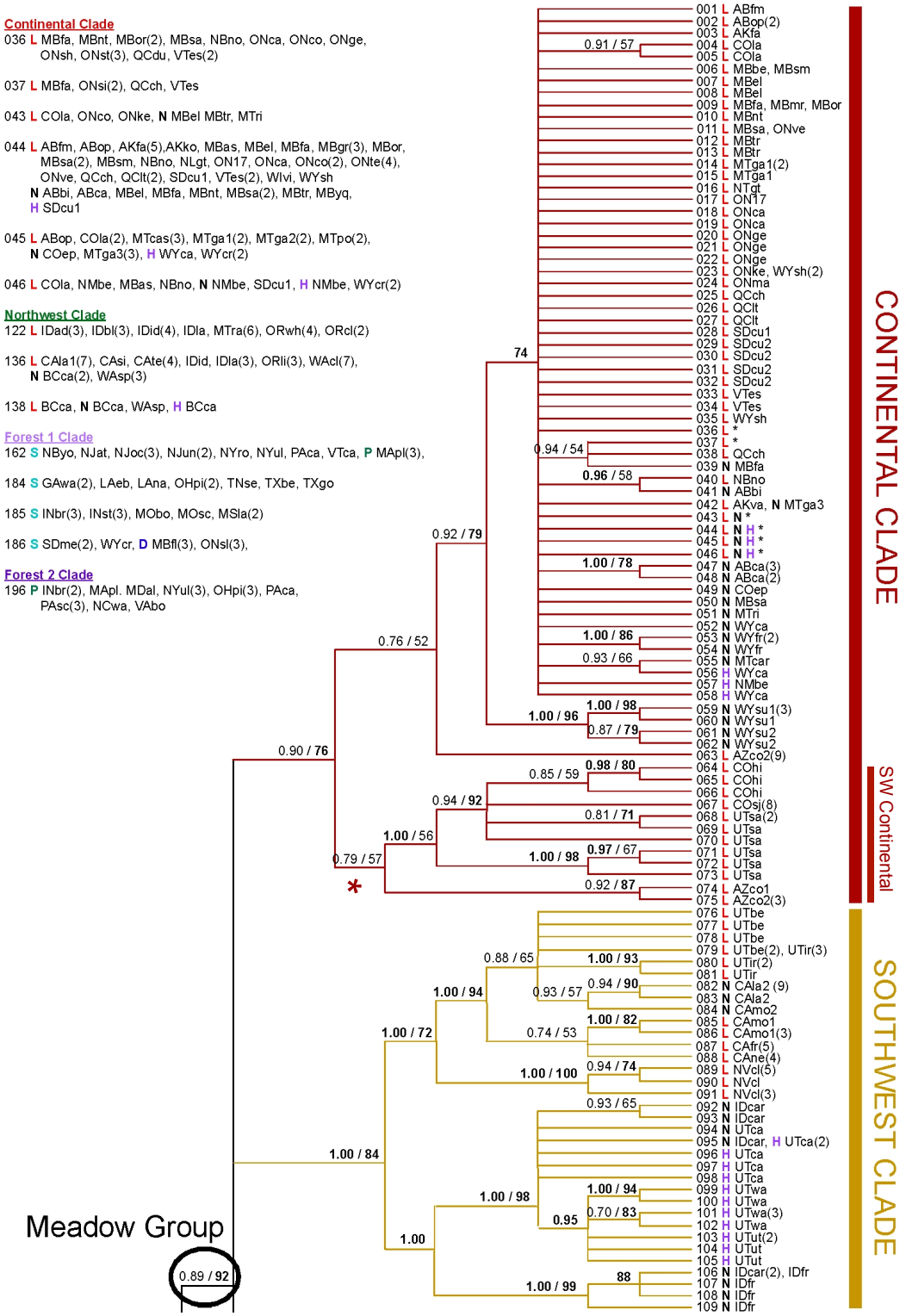

Continental Clade
036 L MBfa, MBnt, MBor(2), MBsa, NBno, ONca, ONco, ONge, ONsh, ONst(3), QCdu, VTes(2)
037 L MBfa, ONsi(2), QCch, VTes
043 L COla, ONco, ONke, N MBel MBtr, MTri
044 L ABfm, ABop, AKfa(5),AKko, MBas, MBel, MBfa, MBgr(3), MBor, MBsa(2), MBsm, NBno, NLgt, ON17, ONca, ONco(2), ONte(4), ONve, QCch, QClt(2), SDcu1, VTes(2), Wlvi, WYsh
N ABbi, ABca, MBel, MBfa, MBnt, MBsa(2), MBtr, MByq,
H SDcu1
045 L ABop, COla(2), MTcas(3), MTga1(2), MTga2(2), MTpo(2),
N COep, MTga3(3), H WYca, WYcr(2)
046 L COla, NMbe, MBas, NBno, N NMbe, SDcu1, H NMbe, WYcr(2)
Northwest Clade
122 L IDad(3), IDbl(3), IDid(4), IDla, MTra(6), ORwh(4), ORcl(2)
136 L CAla1(7), CAsi, CAte(4), IDid, IDla(3), ORli(3), WAcl(7),
N BCca(2), WAsp(3)
138 L BCca, N BCca, WAsp, H BCca
Forest 1 Clade
162 S NByo, NJat, NJoc(3), NJun(2), NYro, NYul, PAca, VTca, P MApl(3),
184 S GAwa(2), LAeb, LAna, OHpi(2), TNse, TXbe, TXgo
185 S INbr(3), INst(3), MObo, MOsc, MSla(2)
186 S SDme(2), WYcr, D MBfl(3), ONsl(3),
Forest 2 Clade
196 P INbr(2), MApl. MDal, NYul(3), OHpi(3), PAca, PAsc(3), NCwa, VAbo
001 L ABfm
002 L ABop(2)
003 L AKfa
004 L COla
005 L COla
006 L MBbe, MBsm
007 L MBel
008 L MBel
009 L MBfa, MBmr, MBor
010 L MBnt
011 L MBsa, ONve
012 L MBtr
013 L MBtr
014 L MTga1(2)
015 L MTga1
016 L NTgt
017 L ON17
018 L ONca
019 L ONca
020 L ONge
021 L ONge
022 L ONge
023 L ONke, WYsh(2)
024 L ONma
025 L QCch
026 L QClt
027 L QClt
028 L SDcu1
029 L SDcu2
030 L SDcu2
031 L SDcu2
032 L SDcu2
033 L VTes
034 L VTes
035 L WYsh
036 L *
037 L *
038 L QCch
039 N MBfa
040 L NBno
041 N ABbi
042 L AKva, N MTga3
043 L N *
044 L N H *
045 L N H *
046 L N H *
047 N ABca(3)
048 N ABca(2)
049 N COep
050 N MBsa
051 N MTri
052 N WYca
053 N WYfr(2)
054 N WYfr
055 N MTcar
056 H WYca
057 H NMbe
058 H WYca
059 N WYsu1(3)
060 N WYsu1
061 N WYsu2
062 N WYsu2
063 L AZco2(9)
064 L COhi
065 L COhi
066 L COhi
067 L COsj(8)
068 L UTsa(2)
069 L UTsa
070 L UTsa
071 L UTsa
072 L UTsa
073 L UTsa
074 L AZco1
075 L AZco2(3)
076 L UTbe
077 L UTbe
078 L UTbe
079 L UTbe(2), UTir(3)
080 L UTir(2)
081 L UTir
082 N CAla2 (9)
083 N CAla2
084 N CAmo2
085 L CAmo1
086 L CAmo1(3)
087 L CAfr(5)
088 L CAne(4)
089 L NVcl(5)
090 L NVcl
091 L NVcl(3)
092 N IDcar
093 N IDcar
094 N UTca
095 N IDcar, H UTca(2)
096 H UTca
097 H UTca
098 H UTca
099 H UTwa
100 H UTwa
101 H UTwa(3)
102 H UTwa
103 H UTut(2)
104 H UTut
105 H UTut
106 N IDcar(2), IDfr
107 N IDfr
108 N IDfr
109 N IDfr
0.91 / 57
74
0.94 / 54
0.96 / 58
0.92 / 79
1.00 / 78
0.76 / 52
1.00 / 86
0.93 / 66
1.00 / 98
1.00 / 96
0.87 / 79
0.90 / 76
0.98 / 80
0.85 / 59
0.94 / 92
0.81 / 71
1.00 / 56
0.97 / 67
0.79 / 57
1.00 / 98
0.92 / 87
*
0.88 / 65
1.00 / 93
1.00 / 94
0.93 / 57
0.94 / 90
1.00 / 82
1.00 / 72
0.74 / 53
1.00 / 100
0.94 / 74
0.93 / 65
1.00 / 84
1.00 / 98
1.00 / 94
0.95
0.70 / 83
1.00
88
1.00 / 99
Meadow Group
0.89 / 92
CONTINENTAL CLADE
SW Continental
SOUTHWEST CLADE

Figure 2.2 (part 2). Phylogenetic tree from Bayesian and Maximum Parsimony (MP) analyses. Support values indicate the Bayesian posterior probability of a node out of a possible 1.0 (*left value*), and the bootstrap percentage (out of 1000 replicates) of the MP analysis (*right value*). Numerical values are only indicated when there was strong statistical support for a node in either analysis ($\geq$ 0.95 / $\geq$ 70) or when a node was recovered in both analyses regardless of support values. Taxon IDs correspond to populations in Table 2.1.
***** The Continental Clade contains a subclade, the SW Continental Clade, a group of populations Identified as distinct in the SAMOVA analysis.

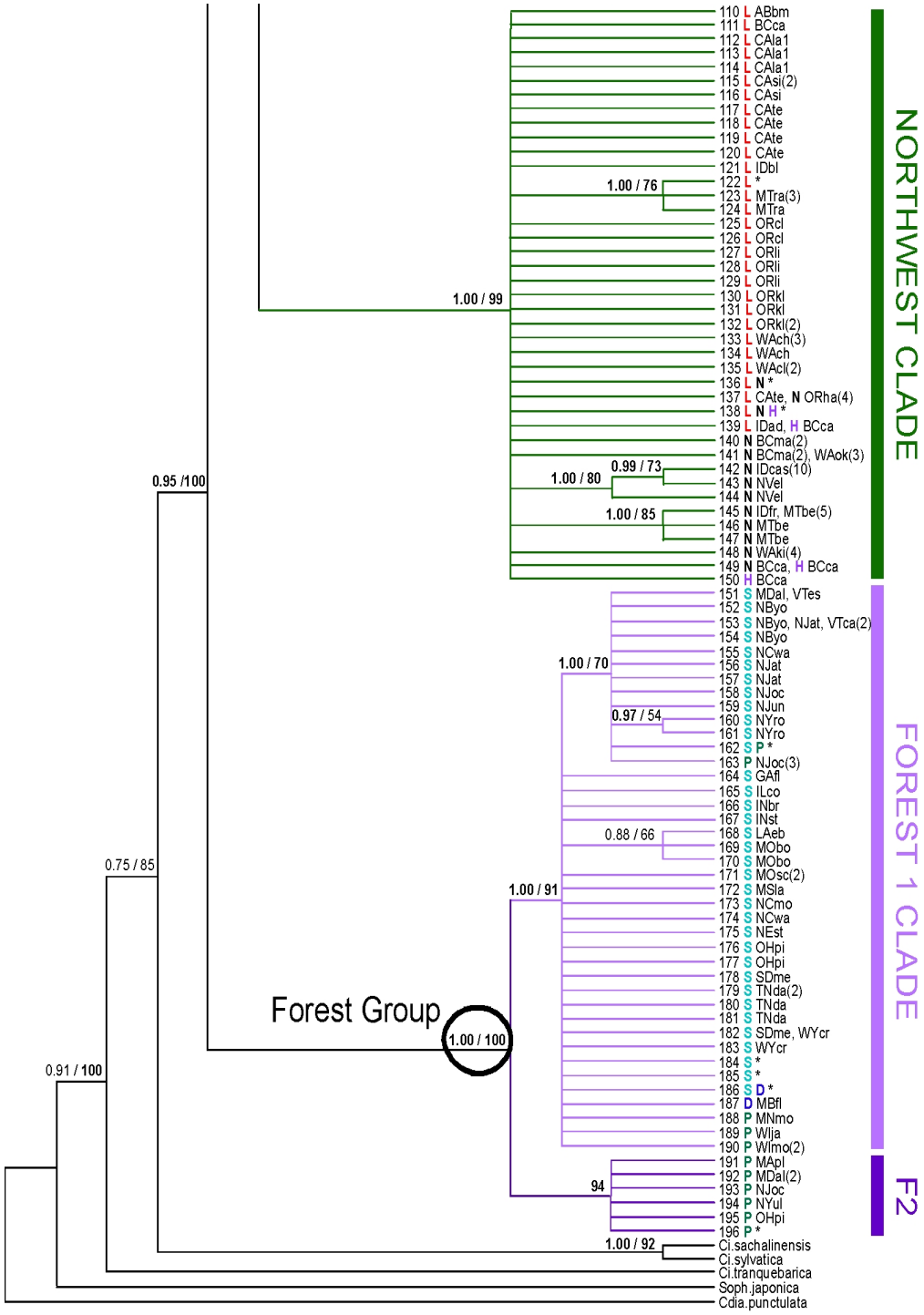

110 L ABbm
111 L BCca
112 L CAla1
113 L CAla1
114 L CAla1
115 L CAsi(2)
116 L CAsi
117 L CAte
118 L CAte
119 L CAte
120 L CAte
121 L IDbl
122 L *
123 L MTra(3)
124 L MTra
125 L ORcl
126 L ORcl
127 L ORli
128 L ORli
129 L ORli
130 L ORkl
131 L ORkl
132 L ORkl(2)
133 L WAch(3)
134 L WAch
135 L WAcl(2)
136 L N *
137 L CAte, N ORha(4)
138 L N H *
139 L IDad, H BCca
140 N BCma(2)
141 N BCma(2), WAok(3)
142 N IDcas(10)
143 N NVel
144 N NVel
145 N IDfr, MTbe(5)
146 N MTbe
147 N MTbe
148 N WAki(4)
149 N BCca, H BCca
150 H BCca
151 S MDal, VTes
152 S NByo
153 S NByo, NJat, VTca(2)
154 S NByo
155 S NCwa
156 S NJat
157 S NJat
158 S NJoc
159 S NJun
160 S NYro
161 S NYro
162 S P *
163 P NJoc(3)
164 S GAfl
165 S ILco
166 S INbr
167 S INst
168 S LAeb
169 S MObo
170 S MObo
171 S MOsc(2)
172 S MSla
173 S NCmo
174 S NCwa
175 S NEst
176 S OHpi
177 S OHpi
178 S SDme
179 S TNda(2)
180 S TNda
181 S TNda
182 S SDme, WYcr
183 S WYcr
184 S *
185 S *
186 S D *
187 D MBfl
188 P MNmo
189 P Wlja
190 P Wlmo(2)
191 P MApl
192 P MDal(2)
193 P NJoc
194 P NYul
195 P OHpi
196 P *
Ci.sachalinensis
Ci.sylvatica
Ci.tranquebarica
Soph.japonica
Cdia.punctulata
1.00 / 76
1.00 / 99
0.95 /100
0.99 / 73
1.00 / 80
1.00 / 85
1.00 / 70
0.97 / 54
0.88 / 66
0.75 / 85
1.00 / 91
Forest Group
1.00 / 100
0.91 / 100
94
1.00 / 92
NORTHWEST CLADE
FOREST 1 CLADE
F2

Figure 2.3. Geographic distribution of major mtDNA clades within the Meadow Group: geographic areas shaded in red, green, and gold correspond to the Continental, Northwest, and Southwest clades from Figure 2.2. Only a single sampling locality contained individuals belonging to more than one clade (denoted by asterisk). This population contained five individuals from the Southwest Clade and one from the Northwest Clade. Sampling localities are colored by species ID: red = *C. longilabris*, black = *C. nebraskana*, violet = both species occur, including putative 'hybrids' based on morphology.

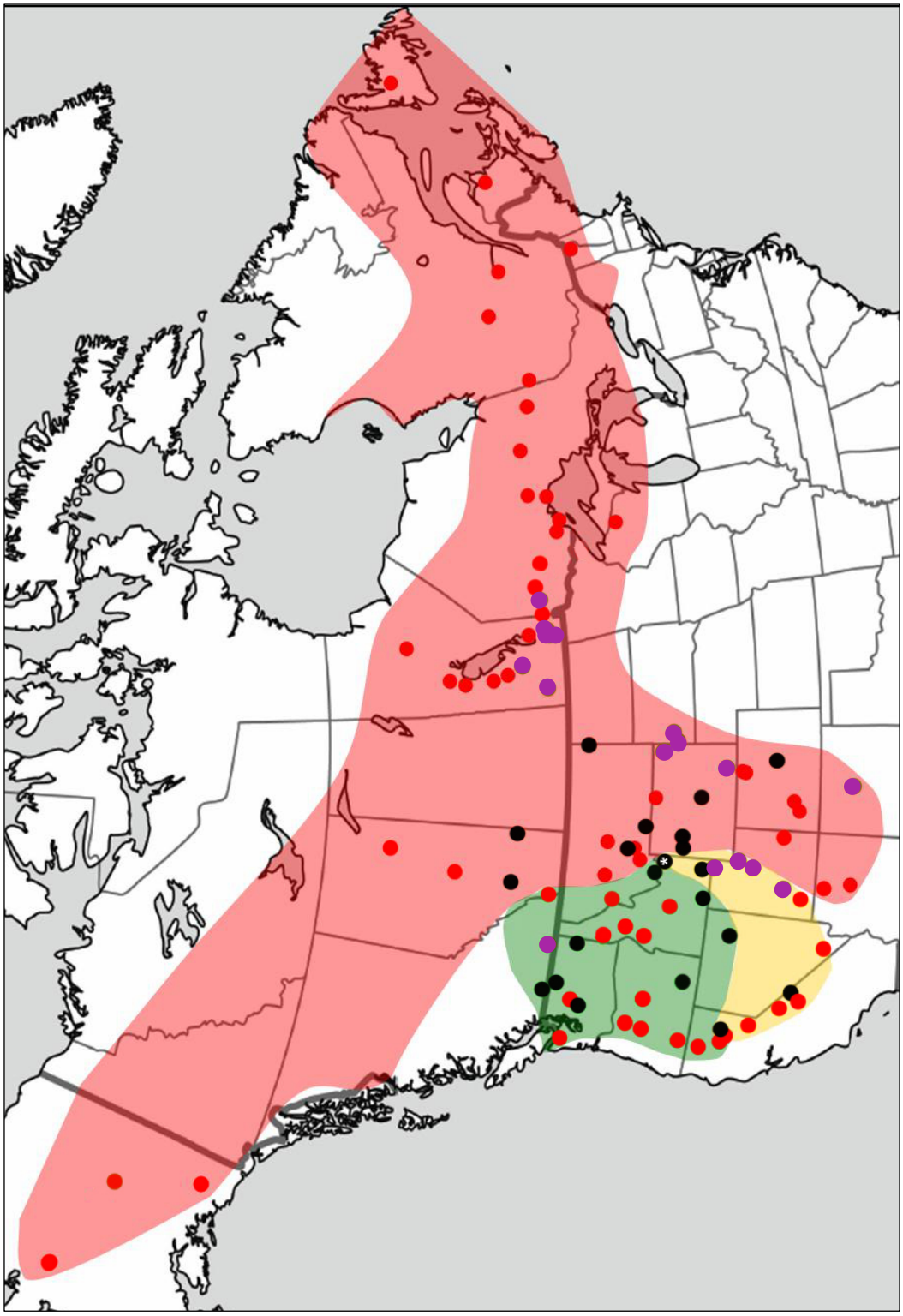
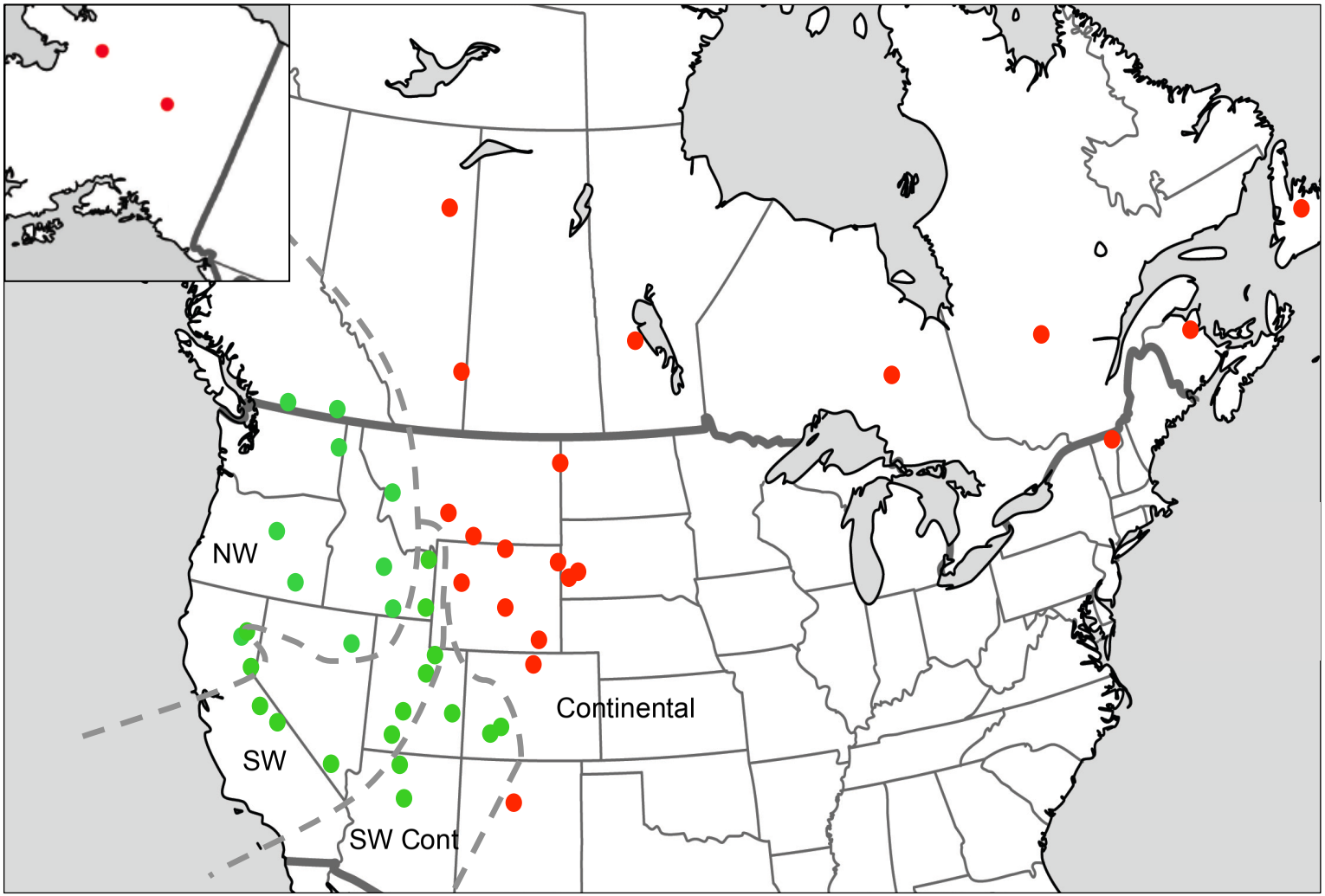


Figure 2.4. **"**Meadow Group" Neighbor-Joining tree based on 1252 AFLP loci. Taxon names are identical to those used in Figure 2.1. Colors of branches correspond to those in Figure 2.5.

Figure 2.5. Meadow group sampling localities colored by clade membership in the Neighbor-Joining tree (Fig. 2.4) based on 1252 AFLPs. Dotted lines indicate the geographical distribution of mtDNA clades (NW, SW, SW Cont, and Continental). Note that the AFLP clades correspond exactly with the geographic break between the mtDNA Continental Clade and all other mtDNA clades. This cytonuclear concordance indicates that the mtDNA history was not entirely idiosyncratic. Taken as a whole, these data strongly support the existence of a deep historical division between populations in the two geographic areas. However, these results do **not** support the existence of the nominal taxa *C. longilabris* and *C. nebraskana*.


NW
Continental
SW
SW Cont

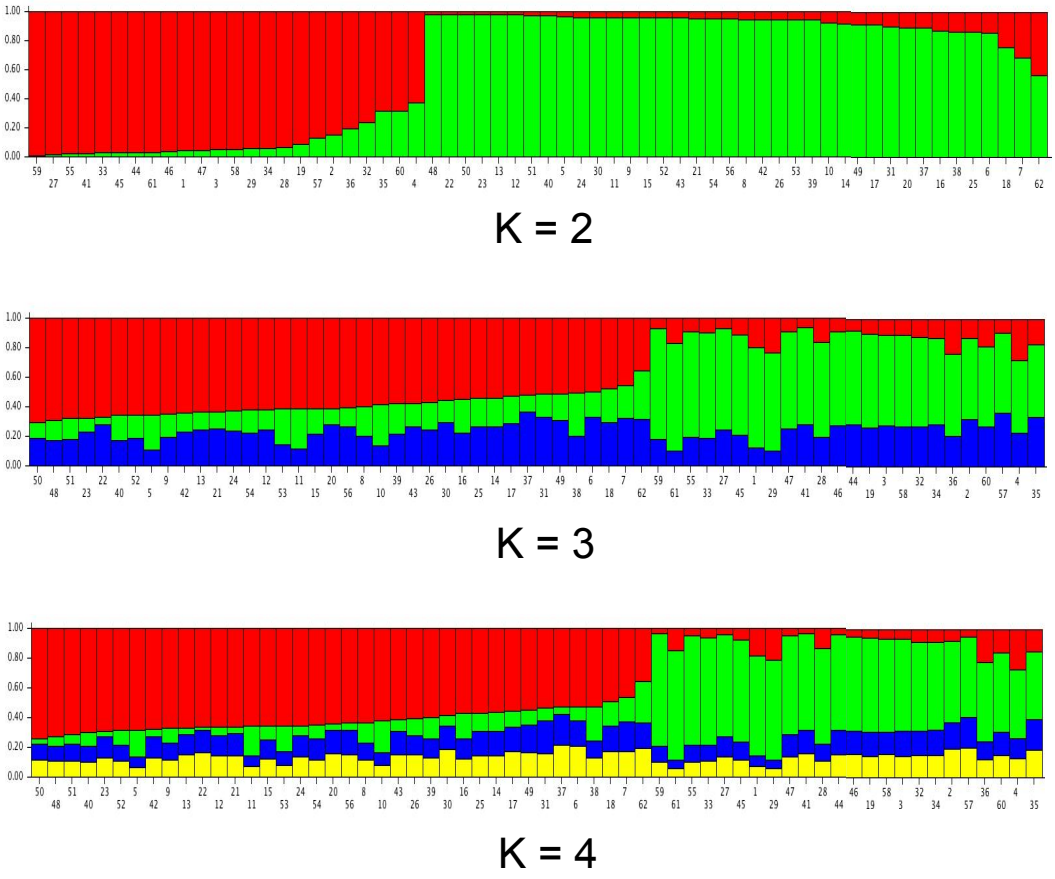


Figure 2.6. Results of STRUCTURE analysis for "Meadow Group" with *K* = 2-4 populations, based on 1252 AFLP loci. Each bar plot is sorted by genomic membership and individuals are not in the same order for each bar. At *K* = 2, most individuals display ≥90% genomic identity associated with one of the populations. Those individuals exhibiting a greater amount of admixture are found near the geographic boundary of the two identified populations (see Fig. 2.7). With increasing *K* populations, groupings do not correspond to biologically, taxonomically, or geographically meaningful groups. Moreover, at *K* > 2 no individual's genome is comprised of more than a fraction of any additionally identified population. Based the biological interpretation and statistical estimates using the Evanno *et al.* (2005) method, *K* = 2 is selected as the best explanation for population structure in the group.

Figure 2.7. Results of STRUCTURE analysis for "Meadow Group" when *K* = 2 populations, based on 1252 AFLP loci. Localities are colored by population membership, with dot size proportional to sampling. Most localities contain individuals that fall strongly into one of the STRUCTURE identified populations (≥90% of their genome); those which exhibit more than 10% genomic identity from the other population are colored by the proportion of their genome belonging to each. Letters indicate taxonomic identity of specimens at each site: L = *C. longilabris*, N = *C. nebraskana*, H = "hybrid" (i.e. low assignment confidence to either taxonomic species based on intermediate morphology). Grey dotted lines indicate the distribution of mtDNA clades. Also note that increasing *K* in STRUCTURE analyses did not recover populations corresponding to other mtDNA clades.

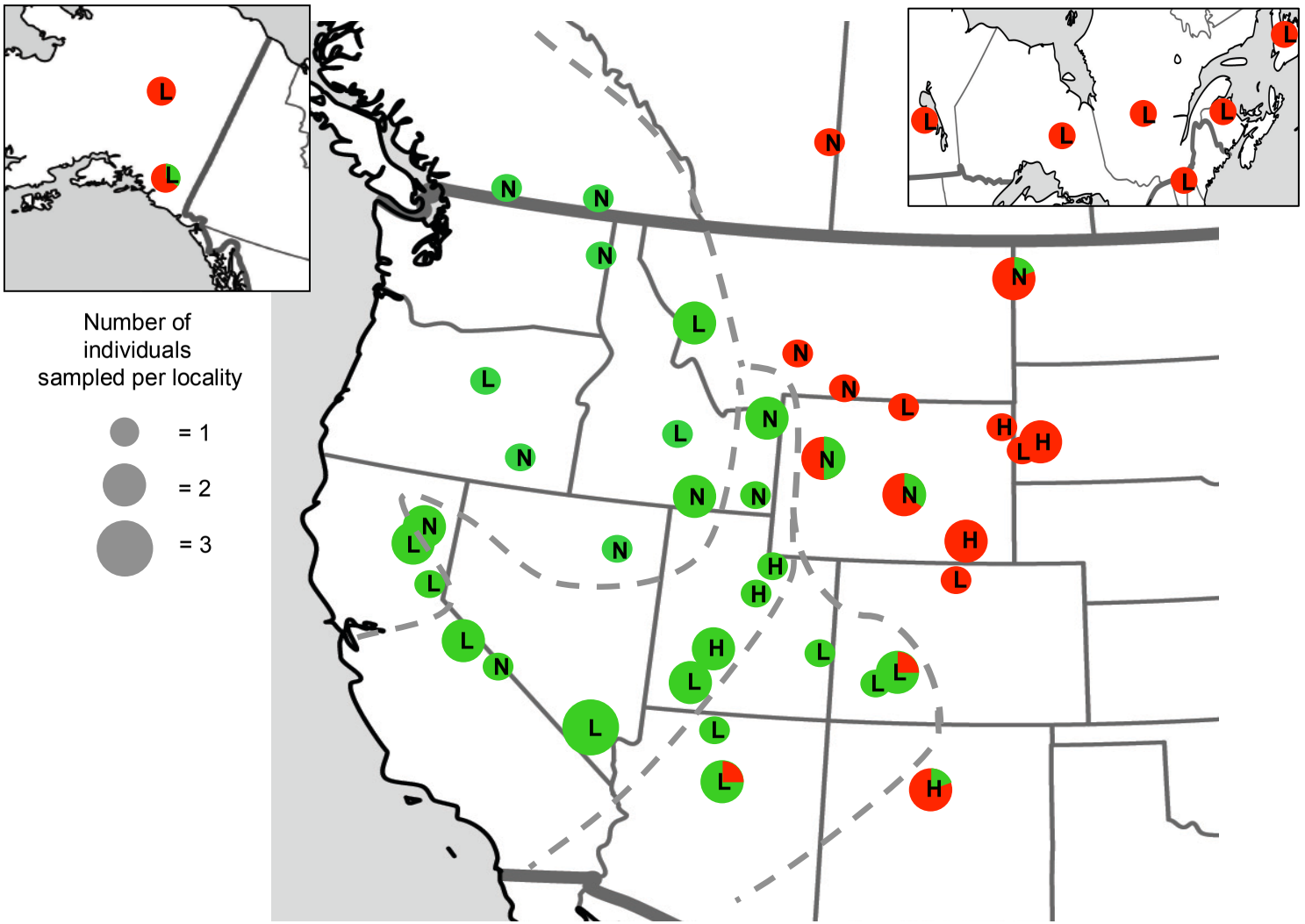

Number of individuals sampled per locality
= 1
= 2
= 3

Figure 2.8. Geographic distribution of major mtDNA clades within the Forest Group: the geographical area shaded in lavender corresponds to the Forest1 clade (see Fig. 2.2), consisting of *C. sexguttata*, *denikei*, and *patruela*; the geographical area shaded in darker purple corresponds to the Forest2 clade, comprised entirely of *C. patruela* individuals. In contrast to the Meadow Group, the clades are sympatric and some sampling localities contain individuals that belong to both clades. Sampling localities are colored by species ID: light blue = *sexguttata*, dark blue = *denikei*, green = *patruela*, sympatric sites are indicated by circles with two colors.

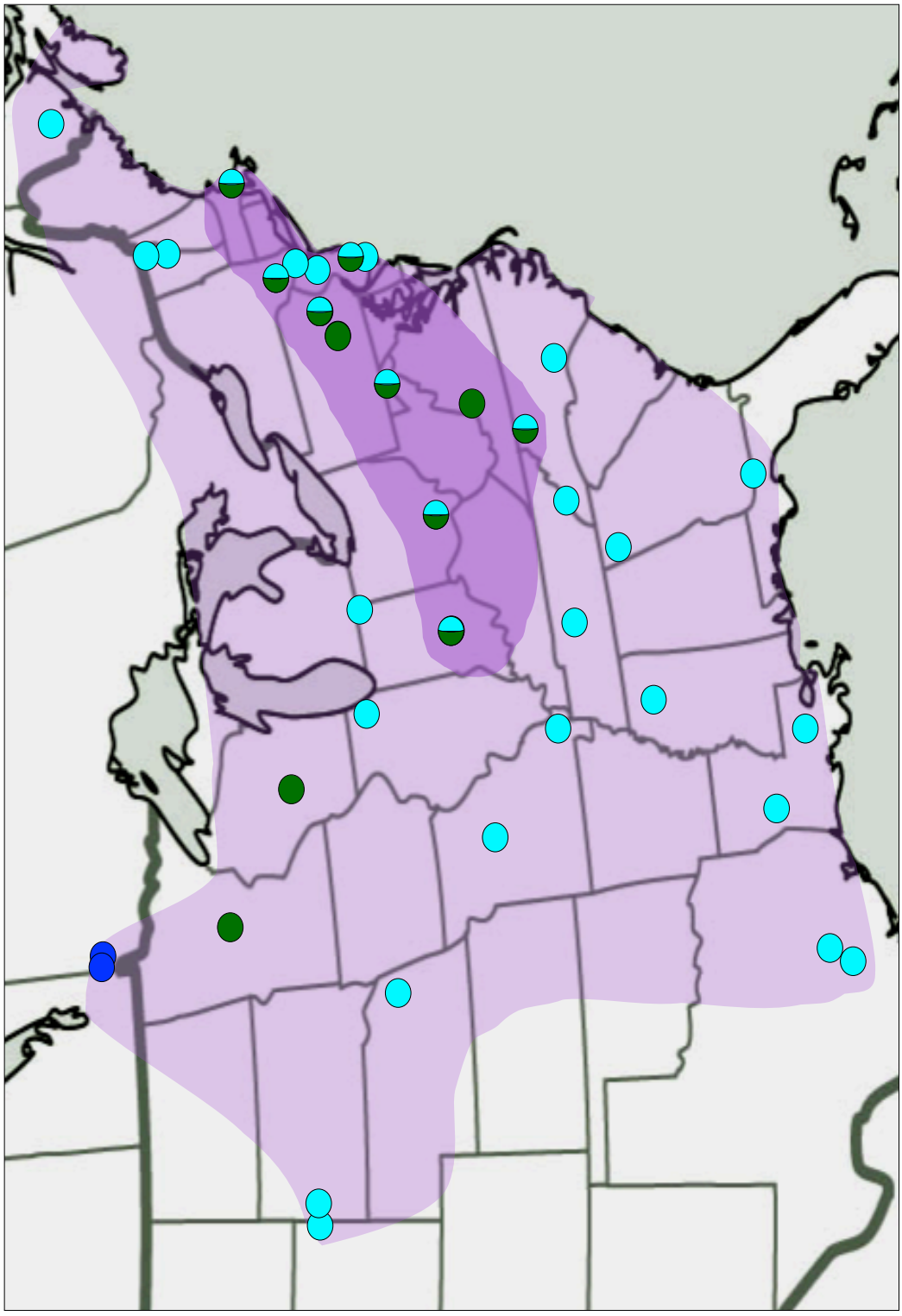
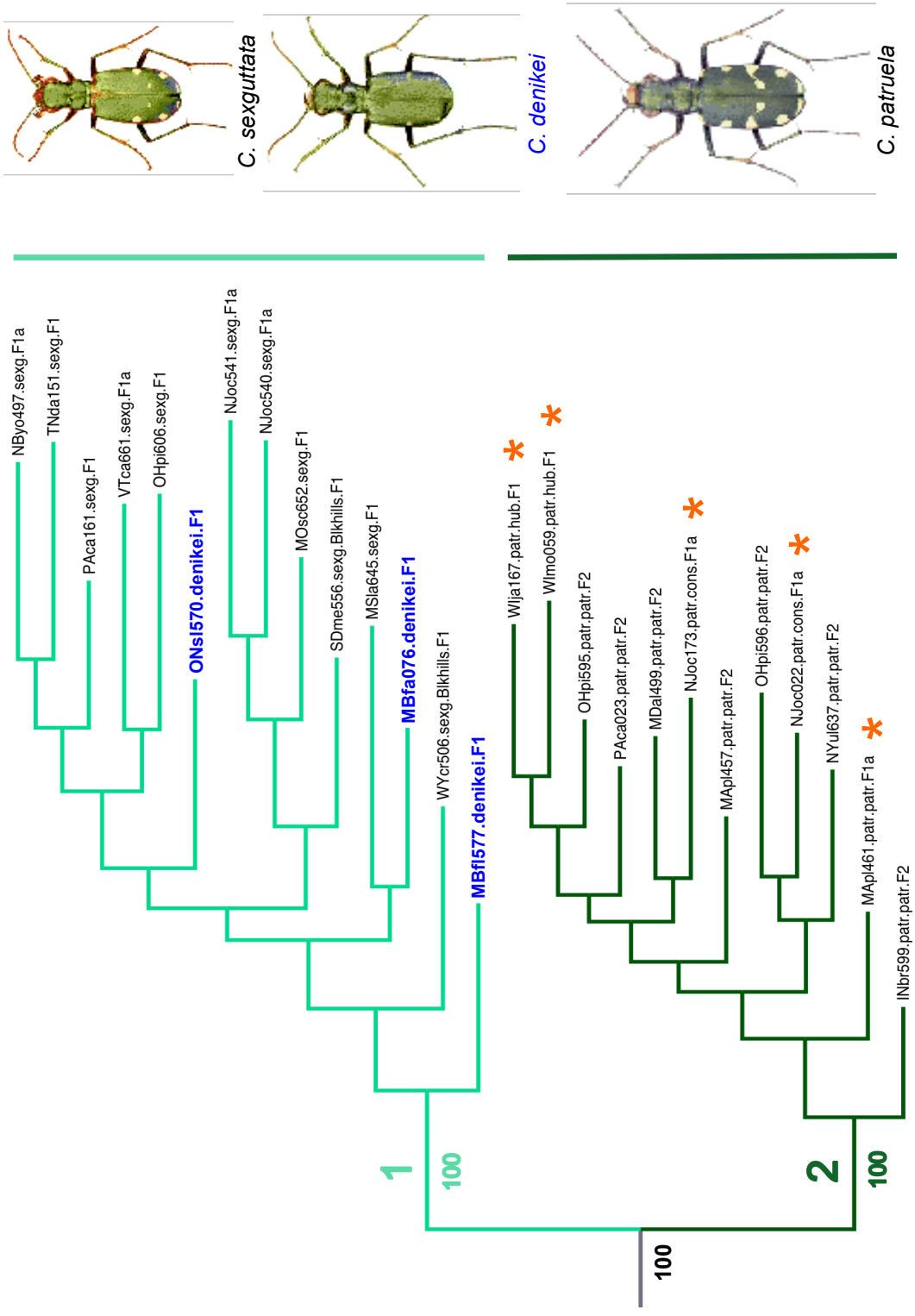


Figure 2.9. **"**Forest Group" Neighbor-Joining tree based on 1252 AFLP loci. Clades correspond to 1) *C. sexguttata* + *denikei*, 2) *C. patruela*. *C. denikei* is not recovered as a separate monophyletic clade in either AFLP or mtDNA trees. Asterisks denote *C. patruela* individuals that contained mtDNA haplotypes more closely related to *C. sexguttata*. Taken together, these patterns support the existence of *C. patruela* and *C. sexguttata* as separate species with occasional hybridization and mtDNA introgression. There is no support for the existence of *C. denikei* as a separate species, however it may be so recently separated as to be undetectable based on overall genomic divergence. Numbers underneath nodes correspond to percent bootstrap support, based on 1000 replicates.


C. sexguttata
C. denikei
C. patruela
NByo497.sexg.F1a
TNda151.sexg.F1
PAca161.sexg.F1
VTca661.sexg.F1a
OHpi606.sexg.F1
ONsl570.denikei.F1
NJoc541.sexg.F1a
NJoc540.sexg.F1a
MOsc652.sexg.F1
SDme556.sexg.Blkhills.F1
MSla645.sexg.F1
MBfa076.denikei.F1
WYcr506.sexg.Blkhills.F1
MBfl577.denikei.F1
Wlja167.patr.hub.F1
Wlmo059.patr.hub.F1
OHpi595.patr.patr.F2
PAca023.patr.patr.F2
MDal499.patr.patr.F2
NJoc173.patr.cons.F1a
MApl457.patr.patr.F2
OHpi596.patr.patr.F2
NJoc022.patr.cons.F1a
NYul637.patr.patr.F2
MApl461.patr.patr.F1a
INbr599.patr.patr.F2
1
100
2
100
100

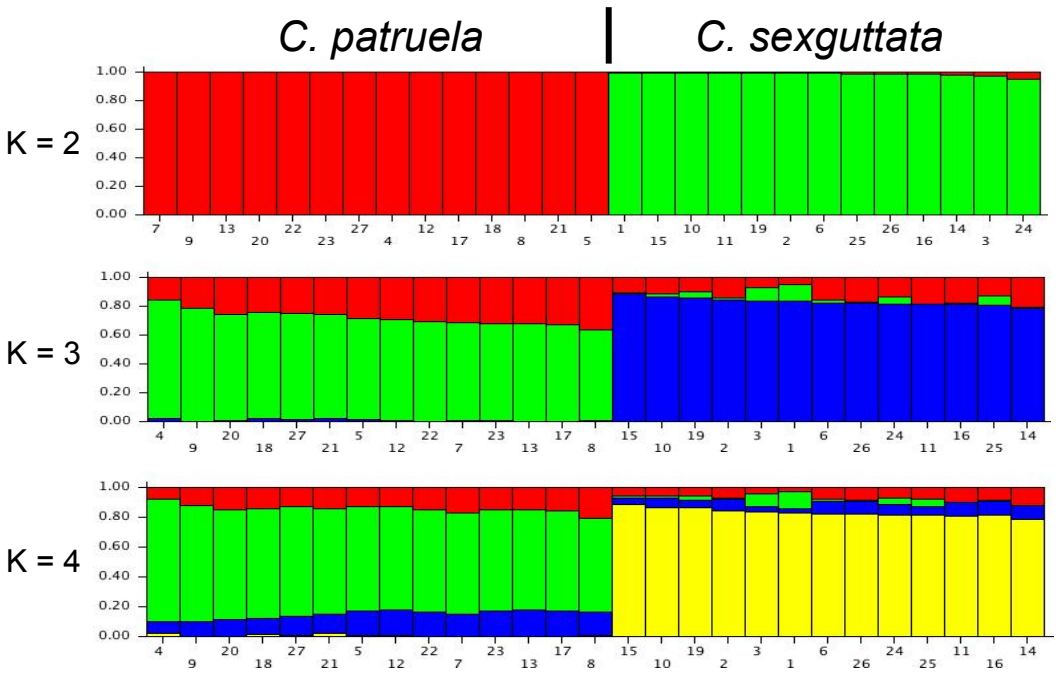


Figure 2.10. Results of STRUCTURE analysis for "Forest Group" with *K* = 2 - 4 populations, based on 1252 AFLP loci. At *K* = 2, all individuals display 96 - 100% genomic identity associated with one of the populations. These populations are identical to clades recovered in the NJ tree (See Fig. 2.1). With increasing *K* populations, groupings do not correspond to biologically, taxonomically, or geographically meaningful groups. Moreover, at *K* > 2 no individual's genome is comprised of more than a fraction of any additionally identified population. Based the biological interpretation and statistical estimates using the Evanno *et al* (2005) method, *K* =2 is selected as the best explanation for population structure in the group.

CHAPTER III

THE EFFECTS OF QUATERNARY ICE AGES ON THE HISTORICAL DEMOGRAPHY OF THE NORTH AMERICAN *CICINDELA SYLVATICA* SPECIES GROUP

Introduction

The distribution of extant species and their current population structure can best be understood within the context of historic events. In particular, Pleistocene glacial cycles had profound effects on many species' present day distributions (*e.g.* Huntley and Webb 1989; Hewitt 1996; Klicka and Zink 1997; Knowles 2001; Rowe *et al* 2004; Burg *et al* 2005; Nice *et al* 2005; Mardulyn *et al* 2009). With the advent of coalescent theory (Kingman 1982) and improved mathematical models and methodologies we can now to perform reliable tests to reject or support specific hypotheses about the historical and contemporary causes of population structure (*e.g.* Knowles 2001; Knowles and Maddison 2002; Templeton 2004). Furthermore, by comparing phylogeographic and historical demographic patterns across a wide range of taxa prior hypotheses about the locations of glacial refugia have been more rigorously supported (Avise 1996;

Swenson and Howard 2005). The results of these studies also contribute to our general knowledge of the context, causes, and timing of speciation.

Phylogeographic patterns are primarily determined through genealogical information, but inclusion of other population genetic data can be particularly revealing as to the demographic history of populations (Avise 2000). One approach involves the examination of two different genetic diversity measures (Nei 1987): haplotype diversity (*h*), a measure of the numbers and frequencies of haplotypes among individuals, and nucleotide diversity (π), a measure of the average weighted sequence divergence between haplotypes. By comparing the patterns of each, we can infer the effects of recent bottlenecks, long-term isolation, and rapid population growth (Grant and Bowen 1998). A limitation of these tests however is that demographic expansion can leave genetic diversity signatures that are similar to those of selective sweeps (Avise 2000). To overcome this problem careful comparison of neutrality tests can potentially establish whether population genetic patterns are caused by population growth or the effects of selection (Fu 1997; Ramos-Onsins and Rozas 2002). Additional demographic information can be gleaned from the frequency of pairwise sequence divergences between individuals in a population. This 'mismatch distribution' (Rogers and Harpending 1992) can be plotted and compared to expectations of distributions under models of rapid demographic growth. Moreover, it is possible to use these population demographic patterns in conjunction with biogegraphic data to infer the existence and location of glacial refugia (Spellman and Klicka 2006). These types of analyses already have been

used successfully to test patterns of historical demography for various North American animal species (*e.g.* Avise et al 1987; Grant and Bowen 1998; Rowe *et al* 2004; Hull and Girman 2005; Russell *et al* 2005). Taken together, these studies allow for a type of "concordance phylogeography" (Avise 1996) whereby we can eventually elucidate the impacts of general historical events on shaping entire biotas. Additional detailed studies will further reveal the degree to which these patterns can be generalized.

The tiger beetles in the North American *Cicindela sylvatica* group are well suited for examining the effects of Quaternary climate change on historical demography. Collectively these species are distributed throughout the majority of the continent, including areas impacted by glacial cycles (Sibrava *et al* 1986) and other areas hypothesized as glacial refugia (Remington 1968; Swenson and Howard 2005). My previous studies on the phylogeography of the North American *C. sylvatica* group (see Chapter II) showed that they were deeply divided genetically into two major groups that corresponded to ecological differences. One branch, the "Meadow Group" (*C. longilabris* and *C. nebraskana*) was most often associated with alpine meadows and grasslands (Leffler and Pearson 1976; Spanton 1988). The second branch, the "Forest Group" (*C. sexguttata, C. patruela,* and *C. denikei*), was typically found in forested areas and adjacent ecotones (Kaulbars and Freitag 1993a, b). However, the patterns of phylogenetic structuring within each group were quite different. The Meadow Group showed deeply separated (1.5 – 2% corrected pairwise divergence) allopatric clades that could not be explained by taxonomic

species boundaries. These clades corresponded to broad geographic areas and were named the Continental Clade (including a distinct subclade, the Southwest Cont Clade), Southwest Clade, and Northwest Clade (see Chapter II). In contrast, the Forest Group mtDNA clades were shallow (0.8% corrected pairwise divergence), overlapped in distribution, and were more congruent with taxonomic boundaries. As such, it was of interest to subsequently focus on the more enigmatic underlying historical demographic processes that might explain the Meadow Group mtDNA patterns and geographic distributions. The goal of this chapter will be to test hypothesis that can explain the patterns observed in the Meadow Group, which are not explicable by taxonomy. Specific questions to be addressed are as follows: 1) Are Quaternary glaciations responsible for historic population fragmentation in the Meadow Group? 2) If so, are there genetic signatures that demonstrate the existence of long-term stable populations in glacial refugia?

I examined the signatures of historical demographic events using multiple population genetic and coalescent-based approaches, implemented in Arlequin 3.1 (Excoffier *et al* 1992; Excoffier *et al* 2005), and DNASP 4.1 (Rozas *et al* 2003). To conduct these tests I will use estimates of divergence times and geographic distribution of unique haplotypes (*i.e.* "private alleles") (Slatkin 1985) to infer population isolation. These combined results will then determine whether genetic patterns are consistent with historical fragmentation during Quaternary glacial cycles. They will also help identify specific signatures of long-term stable glacial refugia and other areas of rapid demographic expansion.

Materials and Methods

Sampling

I sampled specimens from populations representing most of the range of the Meadow Group (*i.e.* combined localities for nominal species *C. longilabris* and *C. nebraskana*). The sampling localities (Table 2.1, Meadow Group: 93 localities, 253 specimens) were based on published records (Leffler 1979; Spanton 1988; Kippenhan 1994), localities provided by other North American tiger beetle researchers and collectors, and additional populations that I located through exploratory collecting of appropriate habitat. Colleagues contributed specimens from additional populations that I was unable to sample personally. We captured specimens using aerial nets or insecticidal sprays and preserved them in $\geq$ 95% ethyl alcohol.

DNA extraction and sequence analysis

Genomic DNA was extracted from most specimens by separating the head + prothorax from the rest of the body then removing flight muscles with sterilized forceps. DNA was isolated from the muscle tissue using the protocol of the DNeasy DNA isolation kit (Qiagen Corp.). Upon removal of tissue for extraction, the voucher specimens were stored in 70% ethyl alcohol for eventual

pinning and vouchering. A small fraction of samples was obtained from dried specimens, and their DNA was extracted by perforating the abdomen using small "minuten" pins; then the entire specimen was placed in 1.5ml microcentrifuge tubes containing lysis buffer and soaked overnight in a 55^{o}C water bath. No specimens were destroyed in the DNA extraction process, and the reassembled vouchers exist for each. Using combinations of several standard insect mtDNA primers (Simon *et al* 1994), initial sequence data was obtained. The following degenerate primers were designed with the program Primer3 (http://www.bioinformatics.nl/primer3plus) for PCR amplification and sequencing: CicF1 5'-AAA GGA AAC ATT TGG TTC ATT (A/G)GG-3', and CicR2 5'-AGT CGA AGA GAT GGA AG(C/T) GC-3'. These primers were used to amplify a 1.1 kilobase fragment of the mitochondrial genes *cytochrome oxidase c* subunit I + tRNA + coII corresponding to positions 2212-3342 of the *Drosophila yakuba* sequence. PCR conditions consisted of an initial denaturation at 96^{o}C for 2 mins, then 10 cycles at 96^{o}C for 30 s, 46^{o}C for 30 s, 72^{o}C for 1 min, and an additional 30 cycles at 96^{o}C for 30 s, 48^{o}C for 30 s, 72^{o}C for 1 min. Negative controls were used in all PCR reactions, and no amplifications were observed in them. The target fragments were sequenced using the Dyenamic Terminator Sequencing Kit (Amersham Biosciences), run out on a polyacrylamide gel using an MJ Basestation, and analyzed with the program Cartographer (MJ Research, Waltham, MA). Depending on the number of specimens available from each site and population, the DNA of between one and ten individuals was extracted and sequenced for use in the analyses. Sequences were edited using Sequencher

4.2 (GeneCorp) and unambiguously aligned, as no indels were present in the fragment. The alignment was verified by eye and trimmed to 972bp for all 523 taxa. I used DNASP 4.0 to identify unique haplotypes for use in the private allele analysis.

Demographic analyses and coalescent-based hypothesis testing

The historical demography of the group was inferred from haplotype diversity (*h*) and nucleotide diversity (π) and calculated for each clade using DNASP 4.0. I determined specific demographic patterns by comparison of the two diversity measures. The most likely cause of a population with low *h* and low π is a recent or severe bottleneck, or a founder event by a small number of individuals (Grant and Bowen 1998). Populations with high values for both diversity measures may have been caused by sustained large populations sizes (Ne), or may result from an admixed sample of individuals from historically isolated populations. When a high value of *h* and a low value of π are observed, this may indicate a recent population expansion. To distinguish some of the alternative explanations, I also calculated the expansion coefficient (S/d) or the ratio of the number of variable sites (S) to the average number of pairwise nucleotide differences (Peck and Congdon 2004); population growth is indicated by larger expansion coefficients.

The above population genetic diversity measures assume selective neutrality. However, non-neutral processes (*e.g.* selective sweeps) can cause

low haplotype and nucleotide diversity patterns, and if unrecognized can confound the interpretation of demographic results (Fu 1997). Furthermore, not all neutrality tests are equally capable of distinguishing between expansion and selection (Ramos-Onsins and Rozas 2002). Fu (1997) demonstrated that comparisons of some neutrality tests could potentially distinguish between expansion and selection. If Fu and Li's (1993) F^* and D^* are significant but Fu's (1997) F_S is not, then background selection is indicated. If the reverse is true, then population expansion is supported. Ramos-Onsins and Rozas (2002) R_2 is a more recently derived neutrality test that can be useful in detecting population expansion, however it may also be influenced by selective sweeps. I used DNASP 4.0 to implement the above set of tests, including 1000 coalescent simulations to test for the statistical significance of each.

I also used mismatch distributions (Rogers and Harpending 1992) to uncover clade demographic histories. DNASP 4.0 was used to compare observed frequencies of pairwise differences with those expected under a model of exponential population growth. Under this model, a smooth unimodal "wavelike" distribution is expected. Alternatively a "ragged" multimodal distribution is expected for a population that has experienced long term stable history. The raggedness statistic (r) was used test the null hypothesis that populations have experienced sustained stable demographic history, using 1000 coalescent simulations in DNASP 4.0.

Lastly, I used two additional methods to characterize population structure and isolation in the Meadow Group. Analysis of molecular variance (AMOVA;

Excoffier *et al* 1992) was implemented with Arlequin 3.1 to examine the degree to which genetic variance is explained by different partitions of the data. In this case, I compared the amount of genetic variation between and within populations for each clade. For population localities where four or more individuals were sampled (Table 3.1), I used DNASP 4.0 to identify private alleles (unique haplotypes found in no other populations). For each of these populations I also calculated genetic diversity measures, *h* and π, to characterize demographic patterns at a more local level.

Results

The mismatch distribution analysis of the Continental Clade revealed a multimodal distribution (Figure 3.1A) that was as ragged as expected under a model of population stationarity ($P = 0.07$), *i.e.* it did not show a pattern significantly associated with overall rapid population expansion. Results from phylogenetic and SAMOVA analyses demonstrated that five southwestern populations were a geographically and genetically distinct subclade within the Continental Clade, termed the "SW Cont" clade (see Chapter II). As such, the mismatch distribution analysis was re-run after removing these populations and the resulting mismatch distribution was more unimodal (Figure 3.1B), and found to be significantly less ragged than null expectations ($P = 0.04$), interpreted to

represent a history of rapid demographic growth for the remaining populations of the Continental Clade. The Northwest Clade displayed a similarly unimodal distribution (Figure 3.1C) that was also significantly less ragged than expected (P = 0.02), a result consistent with recent expansion. In contrast, the Southwest Clade exhibited a multimodal mismatch distribution (Figure 3.1D) that did not deviate significantly from expectations of population stationarity (P = 0.21), indicating a history of long-term stability.

I compared the results from the genetic diversity and neutrality tests (Table 3.2) to expected patterns given a model of exponential population growth (Grant and Bowen 1998; Peck and Congdon 2004). The Continental Clade exhibited high (π > 0.5%) nucleotide diversity (π = 0.626%) and very high (h > 0.9) haplotype diversity (h = 0.927), consistent with long-term stable population size. However, when the five SW Cont populations are removed from the analysis, the remaining group exhibits a pattern of low nucleotide diversity (π = 0.323) and very high haplotype diversity (h =0.904), consistent with recent population expansion. The Northwest Clade displayed a similar signature of expansion, also exhibiting a low nucleotide diversity and high haplotype diversity (π = 0.383, h = 0.903). The Southwest Clade had very high (π > 1) nucleotide diversity (π = 1.279), along with very high haplotype diversity (h = 0.961), a combination that indicates overall long-term population stability. The magnitude of the demographic expansions was characterized by the expansion coefficient (S/d). The Continental Clade minus the SW Cont populations displayed the highest expansion coefficient (S/d = 20.389), followed by the total Continental

Clade (S/d = 12.495), the Northwest Clade (S/d = 11.823), and the Southwest Clade (S/d = 4.902). With the exception of the Southwest Clade, these observed expansion coefficients are high compared to other taxa (see Discussion). All neutrality tests were found to be significant ($P < 0.05$) for the Continental Clade, Continental Clade minus SW Cont, and the Northwest Clade. The R_2 statistic was also found to be highly significant ($P < 0.01$) for both the Continental Clade minus SW Cont, and the Northwest Clade. No neutrality tests were found to be significant for the Southwest Clade.

AMOVA results demonstrated that most genetic variation can be explained by population structure, as opposed to within population variation (Table 3.3.). The Continental and Northwest Clades displayed similar levels of population structure, with ~ 70% of genetic variation explaining by this partition, and the remaining 30% within populations. The Southwest Clade exhibited an even higher degree of population structuring, >85%. This general pattern was further illustrated by the distribution of private alleles (Figure 3.2). All Southwest Clade populations but one were found to exhibit very low gene flow (0.667-1.000 private alleles) with other populations. Similarly, the populations making up the SW Cont Clade (subclade within the Continental Clade) contained very high levels of private alleles. Most of the Continental Clade populations exhibited lower levels of population structuring. Some populations had no private alleles, indicating that all of the haplotypes present in these localities were more geographically widespread (Table 3.1, *e.g.* MT: Gallatin Co (3)). The Northwest Clade showed a large proportion of private alleles in the western part of its

distribution, although they were confined to the Cascade Range, including many higher elevation sites. Insular populations were found to be much less isolated from each other.

Discussion

Historical inferences for mtDNA clades

The demographic analyses demonstrated considerably highest genetic diversity measures (*h* and π values) for the Southwest Clade than for either the Continental or Northwest Clades. This difference indicates overall long-term isolation and stable population size (Grant and Bowen 1998) for the Southwest Clade. All neutrality tests were non-significant, consistent with the interpretation that minimal demographic expansion had occurred overall. Mismatch distribution further corroborated this pattern, as the observed distribution was both multimodal and did not deviate from the null expectation of raggedness (Harpending 1993). AMOVA and private allele distributions illustrated that this pattern of long-term stability and demographic stasis can be extended to the population level as well. Nearly all populations exhibited evidence of reduced gene flow, inferred from the high percentages of private alleles at each sampling site (Slatkin 1985). Most Southwest Clade populations are located in several

mountain ranges within the Great Basin area, and given the above results, these appear to be “sky islands” (*e.g.* Knowles 2001). It is not surprising that these high elevation populations are isolated by the surrounding deserts, given evidence that the Meadow Group is affected by temperature as a primary limiting factor (see Chapter IV Results and Discussion). However, the combined demographic signatures of long-term population structure and isolation for the Southwest clade indicate that the constituent populations may have been isolated more anciently, and their lack of current gene flow may not simply reflect the effects of recent and contemporary climatic conditions. MDIV coalescent simulations estimate the time of splitting between the Southwest and the other clades at 1.66 - 1.78 million years ago (1.32 – 2.39 mya, 95% credibility interval) (see Chapter II). This estimate is consistent with divergence during the Quarternary (mid – early Pleistocene). Since that point, North America has undergone multiple glacial and interglacial periods (Sibrava *et al* 1986) causing many habitats to repeatedly shrink and expand, and these processes may have resulted in additional population fragmentation and isolation for the Southwest Clade.

Population genetic patterns in the Continental and Northwest Clades appear generally more consistent with recent population expansions. Mismatch distribution was particularly informative with respect to the Continental Clade and the SW Cont subclades, a group of genetically differentiated populations identified by prior SAMOVA analyses (see Chapter II). When this population was removed, the remaining Continental Clade exhibited a more unimodal and

smooth mismatch distribution, and a statistically significant (P = 0.04) pattern of rapid population growth based on the resulting raggedness index. The Northwest Clade also displayed mismatch distribution characteristics indicative of expansion, and it was also significantly different from the null expectation of population stability (P = 0.02). However, the results of neutrality tests were less clear-cut with respect to both Continental and Northwest Clades. All neutrality tests were significant, and therefore it is not possible to conclusively say whether they were due to population expansion or to the effects of selective sweeps. Fu (1997) demonstrated that when Fu and Li's (1993) F^* and D^* are significant but Fu's (1997) F_S is not, then background selection is indicated. He also showed that if the reverse is true, then population expansion is supported. In the observed cases, all were significant, a scenario that has less defined expectations. Even so, the selective sweep hypothesis is less likely when one considers the geographic distribution of the population genetic structure. Given that most of the private alleles are generally found at lower latitudes and higher elevations (Figure 3.2), the post-glacial expansion hypothesis is still supported. Under the scenario of population growth following glacial retreat, the diversity measures yielded estimates of the expansion coefficient (S/d) that were high, even when compared to other organisms that displayed Quaternary fragmentation and subsequent expansion. The values observed for the Continental (S/d = 12.495), Continental minus SW Cont (S/d = 20.389) and Northwest Clades (S/d = 11.823) were higher than those found in migratory Mexican free-tailed bats (S/d = 3.955 – 8.185) (Russell *et al* 2005), migratory

elephant seals (S/d = 3.50 - 11.77) (de Bruyn *et al* 2009) and of similar magnitude to that of Black-backed gulls (S/d = 5.5 – 17.8) (Liebers and Helbig 2002) and sharp-shinned hawks (S/d = 9.87 – 28.85) (Hull and Girman 2005). The comparison to vagile species of birds and mammals demonstrates that these tiger beetles were quickly able to follow the retreating glacial ice and expand their populations into newly available habitat.

Inferences about glacial refugia

Results from the Continental Clade mismatch distribution analyses strongly support the existence of a glacial refugium corresponding to the SW Cont populations (northern AZ to southwestern CO) (Figure 3.2). When included, the SW Cont populations caused the overall Continental Clade mismatch distribution to display expected raggedness and multimodality. Yet when these populations were excluded, the signature of Continental Clade population expansion was instead observed. Not surprisingly, this set of populations includes the southernmost extent of the known distribution of the Continental Clade, and typically refugial locations are believed to be in or near the southern limits of species distributions (Hewitt 1996). For the Southwestern Clade, the mismatch distribution and population-level characterization of private alleles indicate that essentially the entire clade could be considered a “refugium”. Patterns suggest that the clade has a demographic history of long-term stabile population size and more ancient fragmentation and persistence. The population

in the most southern extreme (NV: Clark Co.) is also the most geographically isolated from any other population in the clade (Spring Mts. "sky island"), and the most genetically divergent based on average pairwise sequence differences. Interestingly, this population (N = 9) displayed the most pronounced degree of maculation in the phenotype analyses (see Chapter V). Although the Spring Mt. population exhibits complete genetic isolation from other populations, it also displays a pattern of high h and low π, suggesting either a previous population bottleneck or founder effect from an adjacent population. Additional populations may occur in other parts of the Spring Mts., but are unknown as of the present. For the Northwest Clade, refugia locations were less clear. No populations displayed both high h and π, but most populations in the Cascade Range displayed low levels of genetic connectivity as indicated by the distribution of private alleles. Although the entire mountain range could have served as a refugium, it is also possible that the persisting populations may have been concentrated near the southern extent of that more general region. Populations in the CA: Lassen Co. and Tehama Co. area are located in a region that has been repeatedly identified as a glacial refuge based on meta-analyses of hundreds of faunal and floral biogeographic and phylogeographic studies (Remington 1968; Swenson and Howard 2005).

Acknowledgments

I would like to thank the National Park Service for granting collection permits (Olympic National Park and Glacier National Park), and for the field assistance and advice of Jerry Freilich (Olympic N.P.). Dave Brzoska, Todd Lawton, Mike Kippenhan, Barry Knisley, Jason Schmidt, and Richard Freitag provided invaluable information on collection localities and knowledge of species natural history. April Brown, Mark Mandel, Ryan Waples, Austin Ayres, and Nuri Kodaman provided lab assistance. I would also like to thank the entire Duran family and Rodger Gwiazdowski for help with weather updates, road conditions, and assorted valuable miscellaneous assistance. This work was funded by a Vanderbilt University Discovery Grant to Dan Funk.

References

Avise, J.C., J. Arnold, R. M. Ball, E. Bermingham, T. Lamb, J.E. Neigel, C. A. Reeb, and N. C. Saunders. 1987. Intraspecific phylogeography: The mitochondrial DNA bridge between population genetics and systematics. Annu. Rev. Ecol. Syst. 18: 489-522.

Avise, J.C. 1996. Three fundamental contributions of molecular genetics to avian ecology and evolution. Ibis 138:16-25.

Avise, J.C. 2000. Phylogeography: The history and formation of species. Harvard University Press., Cambridge, Mass.

Barroclough, T.G. and A. P. Vogler. 2002. Recent diversification rates in North American tiger beetles estimated from a dated mtDNA phylogenetic tree. Molecular Biology and Evolution 19(10): 1706-1716.

Brower, A.V. 1994. Rapid morphological radiation and convergence in geographical races of the butterfly, *Heliconius erato*, inferred from patterns of mitochondrial DNA evolution. Proc. Natl. Acad. Sci. USA 91:6491-6495.

Brunsfeld, S.J., J. Sullivan, D.E. Soltis, and P.S. Soltis. 2001. Comparative phylogeography of northwestern North America: a synthesis. *In* Integrating ecological and evolutionary processes in a spatial context. (ed. J. S. Antonovics). pp. 319–339. Oxford: Blackwell Science.

Burg, T.M., A.J. Gaston, K. Winker, and V.L. Friesen. 2005. Rapid divergence and postglacial colonization in western North American Steller's jays (Cyanocitta stelleri). Mol. Ecol. 14(12):3745-3755.

CLIMAP. 1981. Seasonal reconstructions of the Earth's surface at the last glacial maximum in Map Series, Technical Report MC-36. Boulder, Colorado: Geological Society of America.

de Bruyn M, Hall BL, Chauke LF, Baroni C, Koch PL. 2009. Rapid Response of a Marine Mammal Species to Holocene Climate and Habitat Change. PLoS Genet 5(7): e1000554. doi:10.1371/journalpgen.1000554

Excoffier, L., Smouse, P., Quattro, J. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131(2): 479–91

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

Fu, Y.X. and W.H. Li. 1993. Statistical tests of neutrality of mutations. Genetics 133: 693-709.

Fu, Y.X. 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics 147: 915-925.

Grant, W.S. and B.W. Bowen. 1998. Shallow population histories in deep evolutionary lineages of marine fishes: insights from the sardines and anchovies and lessons for conservation. J. of Heredity 89: 415-426.

Harpending, R.C. 1994. Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. Human Biology 66:591–600.

Hewitt, G. M. 1996. Post-glacial re-colonization of European biota. Biological Journal of the Linnean Society 68: 87-112.

Hull, J., and D.J. Girman. 2005. Population structure and historical demography in migrating Sharp-shinned hawks (*Accipiter velox*): Effects of Holocene climatic variability. Molecular Ecology , 14:159-170.

Huntley, B. and T. Webb. 1989. Migration: species' responses to climatic variations caused by changes in the earth's orbit. Journal of Biogeography 16: 5-19.

Kaulbars, M.M., and R. Freitag 1993a. Foraging behaviour of the tiger beetle *Cicindela denikei* Brown (Coleoptera: Cicindelidae). Can. Field-Nat. 107:53-58.

Kaulbars, M.M., and R. Freitag 1993b. Geographic variation, classification, reconstructed phylogeny, and geographic history of the *Cicindela sexguttata* group (Coleoptera: Cicindelidae). Can. Entomol. 125:267-316.

Kingman, J.F.C. 1982. The coalescent. Stoch. Process. Appl. 13:235-248.

Kippenhan, M. G. 1994. The Tiger Beetles (Coleoptera: Cicindelidae) of Colorado. Trans. Amer. Entomological Society 120(1):1-86.

Klicka, J. and R.M. Zink. 1997. The importance of recent Ice Ages in speciation: A failed paradigm. Science 277: 1666-1669.

Knowles, L.L. 2001. Did the Pleistocene glaciations promote divergence? Tests of explicit refugia models in montane grasshoppers. Mol. Ecol. 10: 691-701.

Knowles, L.L., and W.P. Maddison. 2002. Statistical phylogeography. Molecular Ecology 11(12): 2623-2635.

Leffler, S. R. 1979. Tiger beetles of the Pacific Northwest (Coleoptera: Cicindelidae). Ph.D. Dissertation, University of Washington, Seattle. 731 pp.

Leffler, S.L. and D.L. Pearson. 1976. The tiger beetles of Washington. Cicindela 8:21-60.

Liebers, D., and A.J. Helbig. Phylogeography and colonization history of Lesser Black-backed Gulls (Larus fuscus) as revealed by mtDNA sequences.

Mardulyn, P., Y.E. Mikhailov, J.M. Pasteels. 2009. Testing phylogeograghic hypotheses in a Euro-Siberian cold-adapted leaf beetle with coalescent simulations. Evolution 63(10): 2717-2729.

Nei, M. 1987. Molecular Evolutionary Genetics. Columbia University Press, New York, NY.

Nice, C.C., N. Anthony, G. Gelembiuk, D. Raterman, and R. Ffrench-Constant. 2005. The history and geography of diversification within the butterfly genus Lycaeides in North America. Mol. Ecol. 14:1741-1754.

Peck, D.R., and B.C. Congdon. 2004. Reconciling historical processes and population structure in the sooty tern *Sterna fuscata*. Journal of Avian Biology 35: 327-335.

Ramos-Onsins, S. E. and Rozas, J. 2002. Statistical properties of new neutrality tests against population growth. *Mol. Biol. Evol.* **19:** 2092-2100

Remington, C.L. 1968. Suture-zones of hybrid interaction between recently joined biotas. pp. 321–428. In Evolutionary Biology (Eds.) T. Dobzhansky, M.K. Hecht, and W.C. Steere, Plenum Press, New York.

Rogers, A.R. 1995. Genetic evidence for a Pleistocene population explosion. Evolution 49:608–615.

Rogers, A.R., and R.C. Harpending 1992. Population growth makes waves in the distribution of pairwise genetic difference. Molecular Biology and Evolution 9:552–569

Rowe, K.C., E.J. Heske, P.W. Brown, and K.N. Paige. 2004. Surviving the ice: Northern refugia and postglacial colonization. Proc. Nat. Acad. Sciences. 101(28):10355-10359.

Rozas, J., Sánchez-DelBarrio, J. C., Messeguer, X. and Rozas, R. 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics 19: 2496-2497.

Sibrava, V., Bowen, D.Q, and Richmond, G.M., 1986, Quaternary Glaciations in the Northern Hemisphere, Quaternary Science Reviews. vol. 5, pp. 1-514.

Simon, C., F. Frati, A. Beckenbach, B. Crespi, H. Liu, and P. Flook. 1994. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. Ann. Ent. Soc. Amer. 87:651-701.

Slatkin, M. 1985. Rare alleles as indicators of gene flow. Evolution 39: 53-65.

Spanton, T.G. 1988. The *Cicindela sylvatica* group: Geographical variation and classification of the Nearctic taxa and reconstructed phylogeny and geographical history of the species (Coleoptera: Cicindelidae). Quaestiones Entomologicae 24: 51-161.

Spellmann, G.M., and J. Klicka. 2006. Testing hypotheses of Pleistocene population history using coalescent simulations: phylogeography of the pygmy nuthatch (*Sitta pygmaea*). Proc. R. Soc. B 273: 3057-3063

Swenson, N.G. and D.J. Howard. 2005. Clustering of contact zones, hybrid zones, and phylogeographic breaks in North America. Am. Nat. 166: 581-591.

Templeton, A. R. 2004. Statistical phylogeography: methods of evalutating and minimizing inference errors. Mol. Ecol. 13: 789-809.

Table 3.1. Population-level genetic diversity statistics for the Meadow Group. Only populations with four or more individuals sampled were included. Shown are the population IDs (same as Table 2.1), mtDNA clade, locality, number of individuals sampled (N), number of alleles observed in the population, number of alleles found exclusively in the population (private alleles), haplotype diversity (h), nucleotide diversity (π), and the proportion of private alleles in the population. DNASP 4.0 (Rozas *et al* 2003) was used to estimate h, π, and identify private alleles.

Popn ID	clade	locality	*N*	# of alleles	private alleles	haplotype div. (*h*)	nucleotide div. (π)	proportion of private alleles
ABca	Continental	Alberta: Calgary	6	3	2	0.733	0.130	0.667
ABop	Continental	Alberta: Opal	4	3	1	0.833	0.120	0.333
AZco1	Continental	Arizona: Coconino Co. (1)	4	2	2	0.500	0.051	1.000
AZco2	Continental	Arizona: Coconino Co. (2)	9	1	1	0.000	0.000	1.000
COla	Continental	Colorado: Larimer Co.	6	5	2	0.933	0.247	0.400
COsj	Continental	Colorado: San Juan Co.	8	1	1	0.000	0.000	1.000
MBel	Continental	Manitoba: Elma	5	4	2	0.900	0.267	0.500
MBfa	Continental	Manitoba: Fairford	6	5	1	0.933	0.329	0.200
MBor	Continental	Manitoba: Orr Creek	4	3	0	0.833	0.171	0.000
MBsa	Continental	Manitoba: Sandilands	7	4	1	0.714	0.118	0.250
MBtr	Continental	Manitoba: N. Two Rivers	4	4	2	1.000	0.223	0.500
MTga1	Continental	Montana: Gallatin Co. (1)	5	3	2	0.800	0.103	0.667
MTga3	Continental	Montana: Gallatin Co. (3)	4	2	0	0.500	0.103	0.000
NBno	Continental	New Bruns: Northumberland	4	4	1	1.000	0.326	0.250
NMbe	Continental	New Mexico: Bernalillo Co.	4	2	1	0.500	0.051	0.500
ONca	Continental	Ontario: Calstock	4	4	2	1.000	0.309	0.500
ONco	Continental	Ontario: Cochrane	4	3	0	0.833	0.103	0.000
ONge	Continental	Ontario: Geraldton	4	4	3	1.000	0.206	0.750
ONte	Continental	Ontario: Terrace Bay	4	1	0	0.000	0.000	0.000
QCch	Continental	Quebec: Chapais	4	4	2	1.000	0.429	0.500
QClt	Continental	Quebec: La Tuque	4	3	2	0.833	0.223	0.667
SDcu1	Continental	South Dakota: Custer Co. (1)	4	3	1	0.833	0.257	0.333
SDcu2	Continental	South Dakota: Custer Co. (2)	4	4	4	1.000	0.412	1.000
UTsa	Continental	Utah: San Juan Co.	8	6	6	0.929	0.716	1.000
VTes	Continental	Vermont: Essex Co.	7	5	2	0.905	0.225	0.400
WYca	Continental	Wyoming: Carbon Co.	4	4	3	1.000	0.446	0.750

Table 3.1. (continued) *IDfr population also contains a single Northwest Clade haplotype.

Popn ID	clade	locality	*N*	# of alleles	private alleles	haplotype div. (*h*)	nucleotide div. (π)	proportion of private alleles
WYcr	Continental	Wyoming: Crook Co.	4	2	0	0.667	0.137	0.000
WYsh	Continental	Wyoming: Sheridan Co.	4	3	1	0.833	0.257	0.333
WYsu	Continental	Wyoming: Sublette Co.	6	4	4	0.800	0.302	1.000
BCca	Northwest	Br.Columbia: Castlegar	10	6	3	0.889	0.217	0.500
BCma	Northwest	Br.Columbia: Manning	4	2	1	0.667	0.274	0.500
CAla1	Northwest	California: Lassen Co. (1)	10	4	3	0.533	0.062	0.750
CAsi	Northwest	California: Siskiyou Co.	4	3	2	0.833	0.120	0.667
CAte	Northwest	California: Tehama Co.	9	6	4	0.833	0.114	0.667
IDad	Northwest	Idaho: Adams Co.	4	2	0	0.500	0.103	0.000
IDbl	Northwest	Idaho: Blaine Co.	4	2	1	0.500	0.154	0.500
IDcas	Northwest	Idaho: Cassia Co.	10	1	1	0.000	0.000	1.000
IDid	Northwest	Idaho: Idaho Co.	5	2	0	0.400	0.123	0.000
IDla	Northwest	Idaho: Latah Co.	4	2	0	0.500	0.154	0.000
MTbe	Northwest	Montana: Beaverhead Co.	7	3	2	0.524	0.059	0.667
MTga1	Northwest	Montana: Gallatin Co. (1)	5	3	2	0.800	0.103	0.667
MTra	Northwest	Montana: Ravalli Co.	10	3	2	0.600	0.069	0.667
ORcl	Northwest	Oregon: Clackamas Co.	4	3	2	0.833	0.309	0.667
ORha	Northwest	Oregon: Harney Co.	4	1	0	0.000	0.000	0.000
ORkl	Northwest	Oregon: Klamath Co.	4	3	3	0.833	0.120	1.000
ORli	Northwest	Oregon: Linn Co.	6	4	3	0.800	0.206	0.750
ORwh	Northwest	Oregon: Wheeler Co.	4	1	0	0.000	0.000	0.000
WAch	Northwest	Washington: Chelan Co.	4	2	2	0.500	0.154	1.000
WAcl	Northwest	Washington: Clallam Co.	9	2	1	0.389	0.080	0.500
WAki	Northwest	Washington: King Co.	4	1	1	0.000	0.000	1.000
WAsp	Northwest	Washington: Spokane Co.	4	2	0	0.500	0.103	0.000
CAla2	Southwest	California: Lassen Co. (2)	10	2	2	0.200	0.021	1.000
CAmo	Southwest	California: Mono Co.	4	2	2	0.500	0.051	1.000
CAne	Southwest	California: Nevada Co.	4	1	1	0.000	0.000	1.000
IDcar	Southwest	Idaho: Caribou Co.	5	4	2	0.900	0.720	0.500
IDfr	Southwest*	Idaho: Fremont Co.	4	4	3	1.000	0.360	0.750

Table 3.1. (continued)

Popn ID	clade	locality	*N*	# of alleles	private alleles	haplotype div. (*h*)	nucleotide div. (π)	proportion of private alleles
NVcl	Southwest	Nevada: Clark Co.	9	3	3	0.639	0.074	1.000
UTbe	Southwest	Utah: Beaver Co.	5	4	3	0.900	0.165	0.750
UTca	Southwest	Utah: Cache Co.	6	5	4	0.933	0.206	0.800
UTir	Southwest	Utah: Iron Co.	6	3	2	0.733	0.219	0.667
UTut	Southwest	Utah: Utah Co.	4	3	3	0.833	0.377	1.000
UTwa	Southwest	UT: Wasatch Co.	6	4	4	0.800	0.295	1.000

Table 3.2. Results of multiple demographic analyses, including diversity statistics, neutrality tests, and mismatch distributions. Analyses were conducted for each of the major mtDNA clades in the Meadow Group (see Fig 2.2 – 2.3). Expected data patterns are given for a model of exponential population growth (Grant and Bowen 1998; Peck and Congdon 2004).

Demographic analysis	Continental	Cont.(-SW Cont)	Northwest	Southwest	Expectation
Diversity statistics					
Nucleotide diversity (π)%	0.626	0.323	0.383	1.279	Low
Haplotype diversity (h)	0.927	0.904	0.903	0.961	High
Expansion coefficient (S/d)	12.495	20.389	11.823	4.902	High
Neutrality tests					
Fu & Li's (1993) F*	-2.796**	-4.271**	-3.051**	-0.744	Not significant
Fu & Li's (1993) D*	-2.920*	4.623**	-3.174**	-0.900	Not significant
Fu's (1997) F_S	-61.727**	-74.783**	-28.419**	-4.798	Significant
Ramon-Onsins and Rozas (2002) R_2	0.040*	0.003**	0.041**	0.098	Significant
Mismatch distribution					
Raggedness (r)	0.005	0.003*	0.002*	0.017	Significant
Shape	Multimodal	Unimodal	Unimodal	Multimodal	Unimodal

Significant results are indicated by asterisks: *P < 0.05, **P < 0.01.

Table 3.3. Analysis of Molecular Variance (AMOVA) (Excoffier *et al.* 1992) for Meadow Group mtDNA clades. Comparison of the genetic variance explained by population structure versus within population variation. *P* values demonstrate that each partition of the data explains a significantly non-zero percentage of the genetic variance. The Southwest Clade exhibits the highest degree of population structure, although in each clade the population structure explains most of the variation.

Clade	Category description	% variance	*P* value
Continental	Among populations	69.26	<0.001*
	Within populations	30.74	<0.001*
Northwest	Among populations	73.20	<0.001*
	Within populations	26.80	<0.001*
Southwest	Among populations	85.65	<0.001*
	Within populations	14.35	<0.001*

Figure 3.1. Mismatch distributions of mtDNA sequences for major phylogeographic clades within the Meadow Group (see Figs. 2.2, 2.3). **A)** Continental Clade including SW Continental Clade. **B)** Continental Clade with SW Continental Clade removed, **C)** Northwest Clade, **D)** Southwest Clade. Solid green lines indicate the expectations under a model of population growth. Red dotted lines indicate the observed mismatch distribution of haplotypes. Recently expanded populations are expected to display a smooth unimodal distribution (Rogers and Harpending 1992). Deviations from this expectation were tested with the raggedness index (Harpending 1994), a statistical test of the null hypothesis that populations experienced long-term stability and constant size. Significant *p*-values (shown in each panel) demonstrate rapid demographic expansion.

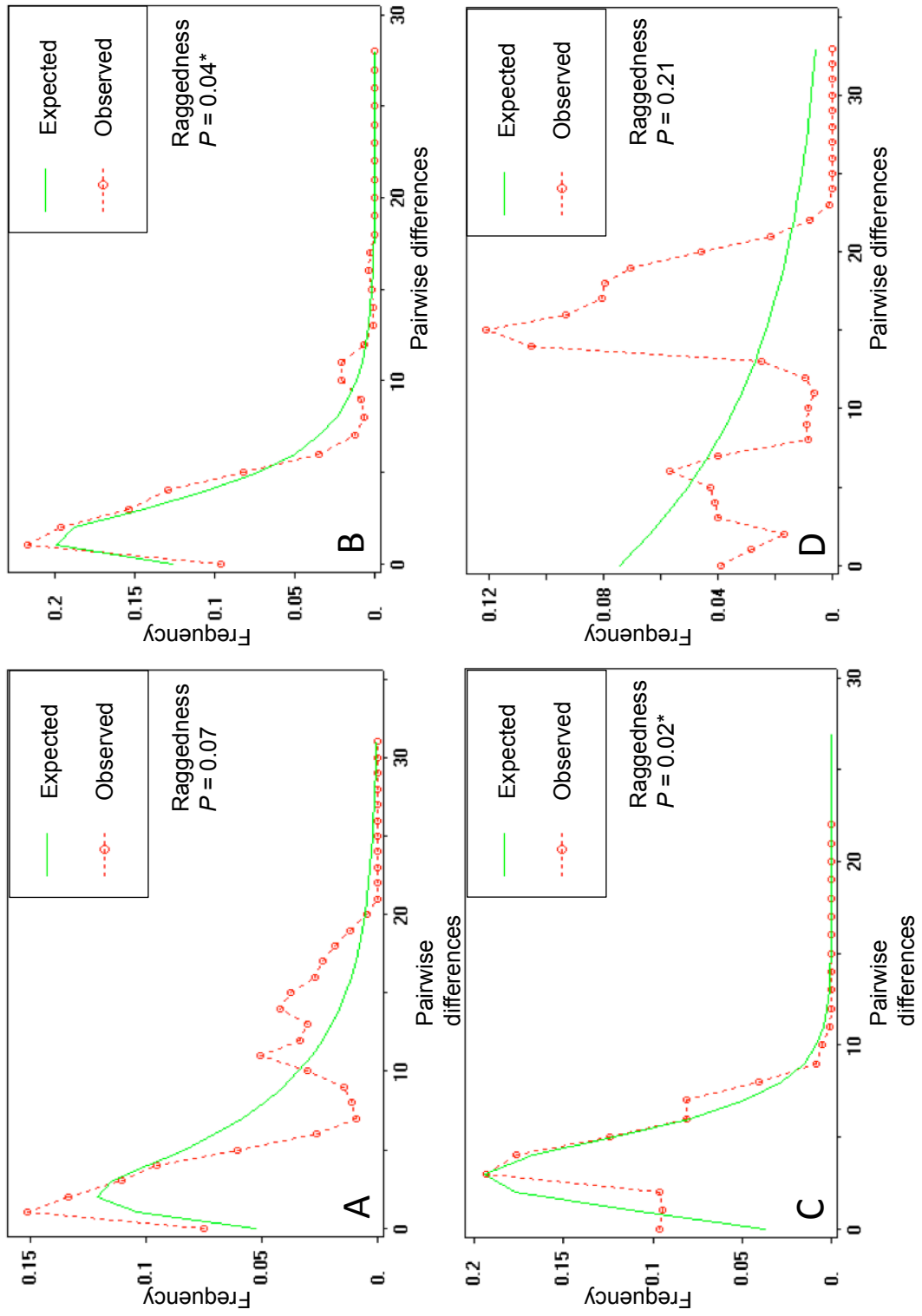
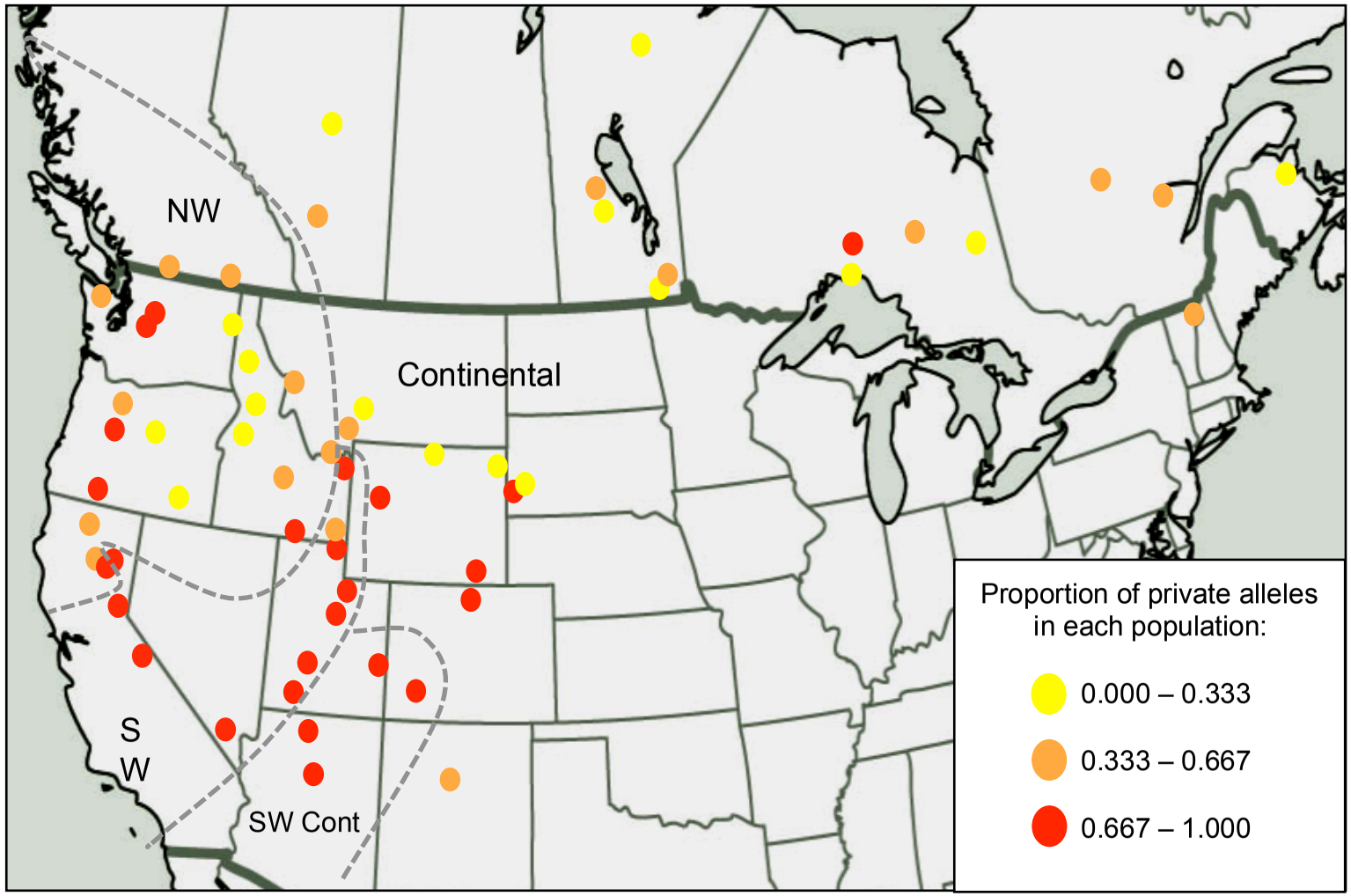

A
Expected
Observed
Raggedness
P = 0.07
Frequency
Pairwise differences
B
Expected
Observed
Raggedness
P = 0.04*
Frequency
Pairwise differences
C
Expected
Observed
Raggedness
P = 0.02*
Frequency
Pairwise differences
D
Expected
Observed
Raggedness
P = 0.21
Frequency
Pairwise differences

Figure 3.2. Distribution of private alleles (mtDNA) in Meadow Group populations. Private alleles are haplotypes that are found in only a single population. A high percentage of private alleles in a population is indicative of restricted gene flow (Slatkin 1985). DNASP (Librado and Rozas 2009) was used to identify private alleles for all populations in which four or more individuals had been sampled. Colored dots show the percentage of private alleles at each sampling site as: low 0-0.333 (yellow), medium 0.333-0.667 (orange) and high 0.667-1 (red) proportions. Dotted lines show the distribution of major phylogeographic clades, the Northwest Clade, Southwest Clade , Continental Clade, and SW Continental Clade (SAMOVA identified subclade within Continental Clade).


NW
Continental
S
W
SW Cont
Proportion of private alleles
in each population:
0.000 – 0.333
0.333 – 0.667
0.667 – 1.000

CHAPTER IV

REVEALING THE FACTORS THAT LIMIT SPECIES DISTRIBUTIONS: INTEGRATING NICHE MODELS WITH PHYLOGEOGRAPHIC HYPOTHESES USING TIGER BEETLES IN THE NORTH AMERICAN *CICINDELA SYLVATICA* GROUP

Introduction

A fundamental question in evolutionary ecology is what factors limit species geographic distributions. The range of a species can be influenced by a complex interplay of evolutionary, ecological, and physiological processes (Cicero 2004). Separating these effects has been a challenge, but theoretical and computational advances have resulted in the development of Ecological Niche Models (ENMs) (Nix 1986; Stockwell and Nobel 1992; Phillips *et al* 2006) that can be used to identify the principal factors predicting and thus potentially determining species ranges. These GIS-based methods relate known species occurrences to multiple environmental data layers across a geographic landscape in order to infer the environmental requirements that determine where a species lives. The ENM can be used to predict the potential species distribution across a geographic area of interest by identifying all areas that

contain similar conditions (Soberon and Peterson 2005). ENMs assume that a species is in equilibrium with its environment and that it will occur in any suitable habitat. Therefore, concordance between the observed and predicted distribution can demonstrate that variables in the model are able to help explain the species occurrence, presumably as a result of intrinsic physiological tolerances to a set of environmental conditions (Swenson 2006).

Absence of a species from an ENM predicted area can be informative, and comparisons between observed and predicted species distributions can be used to evaluate hypotheses about the evolutionary ecology of the study taxa. Discordance between ENMs and known ranges can be the result of extrinsic factors, either biotic or abiotic. For example, a species could be predicted to occur in a geographic area containing suitable habitat, but dispersal limitations may have prevented colonization, either through insufficient time for dispersal (*i.e.* since glacial retreat), insurmountable distance, or due to a discreet barrier (Svenning *et al* 2006). Ecological interactions with other species could also play a role. Species with strong host associations may be limited by the boundaries of the host range (Levine *et al* 2007). Competitive exclusion has been invoked when a species is predicted to occur in an area but its observed distribution ends abruptly where it comes in contact with another ecologically similar species (Anderson *et al* 2002; Sanchez-Cordero *et al* 2008). In some instances ENM models have predicted species occurrences into previously unsampled areas, and these have led to the identification of new populations for rare or poorly

known groups (Menon *et al* 2010), or resulted in the discovery of undescribed species that were closely related to the study taxa (Raxworthy *et al* 2003).

Another interesting application of ecological niche models is the corroboration of previously hypothesized Pleistocene glacial refugia (Waltari *et al* 2007; Waltari and Guralnick 2009). This can be accomplished by creating an ENM for a species and then "backcasting", *i.e.,* comparing to paleoclimate reconstructions during Last Glacial Maximum (~18,000 years ago). Under the assumption of niche conservatism (reviewed in Weins and Graham 2005), this method can be used to predict a species' past distribution by identifying geographic regions that contained similar environmental conditions to those where the species occurs today. Resulting inferences can be especially informative when compared to phylogeographic patterns, as a form of cross-validation of hypotheses. Similarly, ENMs have been used to reveal other historical phenomena such as the creation of hybrid zones following post-glacial range shifts (Swenson 2006), and identification of the environmental factors maintaining them into the present. Integration of ENMs with other biological subfields has the potential to yield many new insights into the causes of evolutionary patterns (Kozak *et al* 2008).

The tiger beetles (Coleoptera: Carabidae: Cicindelinae) are a group of generalist predatory insects that are well studied ecologically (Knisley and Schultz 1997). Both adults and larvae are predaceous, and most species are diurnally active in open habitats, such as sand dunes, open fields, alkali flats, and patches of bare soil or rock (Pearson 1988). Larvae are ambush predators that

live in burrows in soil or rarely in other substrates such as rock crevices (Kaulbars and Freitag 1993b). The larvae lie in wait at the top of their burrows with their mandibles open and their heads and pronota flush with the ground surface (Knisley and Schultz 1997). When a small invertebrate comes within reach they fling their heads in the direction of the prey, grasp it with their mandibles, and drag the prey down into the burrow to feed. Adult tiger beetles run on the ground after prey, capturing and killing them with their mandibles. . Larvae and adults typically occur in the same habitats, with only a few known exceptions (Knisley and Schultz 1997).

Tiger beetles are well-suited to niche modeling because they are some of the best-studied non-pest insects (Knisley and Schultz 1997) and dependable range maps and detailed occurrence data exist for all North American species of *Cicindela* (Pearson *et al* 1997). Niche models are most robust when reliable presence data exist (Pearson *et al* 2006; Phillips *et al* 2006). Moreover, there is a wealth of natural history literature on tiger beetles, and hypotheses have been put forth about factors believed to be most critical in limiting species ranges. Dispersal limitations have been proposed as an explanation for some species range boundaries (Kaulbars and Freitag 1993b), and others have invoked tiger beetles' strong affinities with specific habitats and soil types (Leffler 1979; Spanton 1988; Schincariol and Freitag 1991). Tiger beetles have been used in studies of environmental tolerances including thermoregulation (Dreisig 1980; Schultz and Hadley 1987; Schultz 1998), metabolism and water balance (Hadley and Schultz 1987; Schultz *et al* 1992). Other, previously uninvestigated, abiotic

variables might also be important in this regard. Yet the ecological process of competitive exclusion could also prevent species from co-occurring at fine spatial scales or over larger geographic areas (Connell 1980; Vitt *et al* 1999). Some tiger beetle studies have demonstrated that competitive interactions could be important in shaping community structure (Pearson and Mury 1979; Pearson 1980; Pearson and Juliano 1991), and one study identified food availability as a potentially limiting resource in a mesic grassland habitat (Pearson and Knisley 1985). Despite this body of work, to date there has never been a study quantifying the ecological niches of tiger beetles using a GIS-based niche modeling approach.

In the case of the tiger beetle species in the North American *C. sylvatica* group, interesting phylogeographic patterns have been recently uncovered (see Chapter II) and ecological niche models could be informative when compared to the current hypotheses about their evolutionary history. The group includes five nominal taxa, which correspond to two ecologically and phylogenetically distinct clades: The "Meadow Group" (*C. longilabris* and *C. nebraskana*) often associated with alpine meadows and grasslands (Leffler and Pearson 1976; Spanton 1988), and the "Forest Group" (*C. sexguttata, C. patruela,* and *C. denikei*), typically found in forested areas and adjacent ecotones (Kaulbars and Freitag 1993a, b). The phylogeographic work demonstrated that the Meadow Group nominal taxa were not separate species, and as such they appear to represent phenotypic polymorphism within a single species. Nonetheless, deep genetic breaks were discovered within the Meadow Group, and population

genetic analyses supported the possibility of fragmentation during the Quaternary Ice Ages with evidence for recent demographic expansion following glacial retreat (see Chapter III). Given these results, ecological niche modeling could be revealing as to whether suitable habitat existed in hypothesized glacial refugia for each of the mtDNA clades. In addition, even if *C. longilabris* and *C. nebraskana* are not distinct species, they could represent populations containing specific genes under strong selection against an otherwise undifferentiated genomic background. It is possible that particular areas of the genome may represent "genomic islands" associated with ecological and morphological traits (Nosil *et al* 2009; Michel *et al* 2010) under divergent selection. Therefore, it could be informative to quantify the niches of the nominal species and identify whether they may be ecologically differentiated, even if phylogenetically interchangeable.

Also of interest was the distinction between *C. sexguttata* and *C. denikei*. Phylogenetic and coalescent-based analyses were inconclusive as to the validity of the latter taxon, which exists as a small set of populations that are disjunct from the presumed parental species, *C. sexguttata* (see Chapter II). If *C. denikei* evolved *in situ* in its current habitat, then it could only have diverged less than 11,000 years ago (CLIMAP 1981; Sibrava *et al* 1986; Mix *et al* 2001), potentially insufficient time to exhibit reciprocal monophyly with respect to *C. sexguttata*. However, it is possible that *C. denikei* could have diverged from *C. sexguttata* prior to the LGM in a separate geographic area and later dispersed to its current location following glacial retreat. If this second scenario were true, then the case

for incomplete lineage sorting is less compelling, as it would indicate that the divergence was more ancient.

To address the above questions, I generated ENMs for the North American *Cicindela sylvatica* group and conducted statistic tests to determine which factors are most important in limiting the species ranges. I also created models based on the results of my prior phylogeographic work, treating geographically discrete mtDNA clades as taxa (Swenson 2006) and evaluating alternative explanations for their current distributions. Statistical tests of niche identity and niche overlap were conducted that allowed for more rigorous testing of alternative hypotheses regarding species boundaries and evolutionary inferences.

Materials and Methods

Species occurrence data

For each species, I gathered locality data from published sources (Leffler 1979; Spanton 1988; Kippenhan 1994), museum records (Table 4.1), localities contributed by numerous North American tiger beetle workers (see Acknowledgements), and localities sampled during the molecular portion of my dissertation (see Table 2.1). At each of my collection sites I recorded the latitude

and longitude using a Garmin GPSMap 60-CSX. The majority of locality data from other sources were in the form of textual descriptions, and these were converted to decimal latitude and longitude coordinates through the use of Google Earth (http://earth.google.com) and Garmin MapSource (Garmin Ltd., Olathe, KS). Unfortunately, many historic occurrence records were insufficiently detailed to be used in the analysis (*e.g.* state and county only) and to avoid basing the ENMs on imprecise data, only records that could be georeferenced to within a 5 km^2 area were included. A total of 782 presences were included in the dataset (Table 4.2), out of ~2000 records examined.

In addition to the five nominal species, I generated ENMs for the three *C. longilabris* subspecies. Spanton (1988) described these as generally allopatric in distribution except in the northern Rockies where they come into contact and apparently "hybridize" forming phenotypic intergrades. I also created ENMs for the Meadow Group, as phylogenetic and STRUCTURE analyses revealed that *C. longilabris* and *C. nebraskana* were not separate species (see Chapter II Discussion). Finally, I created ENMs for the three major mtDNA clades within the Meadow Group. These were found to be tightly allopatric and non-overlapping and consequently I sought to uncover the underlying basis for their distributions.

Environmental data

To create ENMs, I used climate layers based on the 19 bioclimatic variables in the WorldClim dataset (Hijmans *et al* 2005a). These represent

means and extremes of environmental conditions that are likely to limit species distributions (Phillips *et al* 2006). I also included layers corresponding to World Wildlife Fund ecoregions (Olson *et al* 2001) and major groups of soil orders (http://soildatamart.nrcs.usda.gov). These last two were chosen because habitat type and soil conditions have been suggested as key factors limiting the distribution of *Cicindela* species (Leffler 1979; Schincariol and Freitag 1991), including the *C. sylvatica* group (Spanton 1988; Kaulbars and Freitag 1993b). The 19 bioclimatic variables exhibited a spatial resolution of 30 arc-seconds (~1 km^2 grid cells at the equator), the highest resolution available to date (WorldClim Version 1.4, last accessed May 2010). The ecoregion and soil layers were originally at 2.5 arc-minute spatial resolution (~5 km^2 grid cells at the equator) and were re-gridded to match the cell size of the other layers using a nearest-neighbor algorithm in DIVA-GIS (Hijmans *et al* 2005b). All layers were formatted for use with niche modeling software using DIVA-GIS and cropped to a geographic area corresponding to the Nearctic biogeographic realm, including northern Mexico. Ranges of the cropped areas include the following: longitude range: -176.667, -50.825, and latitude range: 75, 21.658.

Generation of Ecological Niche Models

Species ecological niches were modeled using MAXENT version 3.3 (Phillips *et al* 2006), a machine learning type algorithm that has been demonstrated effective at generating ENMs with presence-only data (Elith *et al*

2006). I used the default settings for the program, and maximum number of iterations (500), as these have been shown to be robust across taxa (Phillips *et al* 2006). The program was also used to calculate the area under the receiver operating characteristic (AUC), a widely used measure of model predictive performance (Fielding and Bell 2002; Phillips *et al* 2006; but see Lobo *et al* 2008). The AUC can range from 0-1, with higher values indicative of a greater fit of the model to the data. An AUC of 0.5 would be consistent with a random prediction, and 1.0 indicates a perfect fit between the model and the data. AUC values over 0.9 have been considered high support for the fit of the model to the data (Fielding and Bell 2002). Presence data were divided randomly into 75% training and 25% model testing partitions, following Pearson (2007). The training points are used to generate the fundamental niche model and following this, the model is checked against the test points. The degree to which the training model fits the test data (the 'test AUC') is an important indicator of model predictivity (Phillips *et al* 2006). Statistical tests of significance were implemented within MAXENT to evaluate the null hypothesis that test points are predicted no better than random.

For each taxon, jackknife analyses were conducted in MAXENT to assess which environmental variables were most predictive of the species ENM. This approach was used to conduct three types of analyses: 1) models were created using all variables except one, with each such possible model evaluated, 2) models were created using each variable in isolation, and 3) models were created using the total set of variables. Comparison of the resulting jackknife

AUC values can be very informative as to the relative importance of each environmental layer in predicting species distributions. The first test demonstrates the negative impact on the predictive power of the model when a particular variable is omitted. The second test shows how predictive each variable is when it is the *only* layer used to create the ENM. The third is the entire model AUC for comparison to the first two tests.

MAXENT outputs are continuous probability values ranging from 0-100 across each cell grid of the study area. These are converted to binary predictions of species occurrence by selecting a threshold. The choice of threshold should be based on the details of the study organisms, including dispersal capabilities and the reliability and accuracy of presence data (Liu *et al* 2005; Jimenez-Valverde and Lobo 2007; Pearson 2007; Pearson *et al* 2007). Some tiger beetle species are known to be vagile, and in the extreme, *C. trifasciata* has been observed flying to oil rigs 160km from the nearest land (Graves 1981). However, most tiger beetles in the *C. sylvatica* group rarely appear to stray from suitable habitat and collectors visit many of the published localities year after year (Knisley, Bzroska, Lawton, Schmidt, *pers. comm.* 2005), validating these presences. For this reason, I used the Lowest Predicted value Threshold (LPT) of environmental suitability (Phillips *et al* 2006; Pearson 2007) as my less stringent threshold. The LPT predicts the minimum area at which a species occurs while ensuring that no observed species presences are omitted by the model. This method will necessarily be affected by the presence of outlier localities (*e.g.* dispersers into non-suitable habitat) if they exist, for they will widen

the predicted niche to accommodate them. Consequently, I also chose a second more stringent threshold for comparison, the M10 (*sensu* Waltari *et al* 2007). This is a fixed sensitivity threshold that treats 10% of occurrence points (those at the edges of the fundamental niche) to be discarded. Necessarily it will predict a more restricted and conservative fundamental niche and subsequent geographic distribution for the species. A large discrepancy between the amount of habitat predicted under the two thresholds is indicative of the strong effect of outliers (Waltari *pers. comm*. 2010), or alternatively the result of niche divergence between populations.

Paleoclimate reconstructions

In order to compare to phylogeographic hypotheses, I used LGM climatic reconstructions based on the two existing models available from the Paleomodeling Intercomparison Project II (Braconnot *et al* 2007): the Community Climate System Model (CCSM) (Collins *et al* 2006) and the Model for Interdisciplinary Research on Climate (MIROC) (Hasumi and Emori 2004). The original data were downloaded from the PMIP2 website (http://www.pmip2.cnrs-gif.fr), with a spatial resolution of 2.5 arc-minutes. These layers included the same 19 bioclimatic variables used to create present-day ENMs, although at lower resolution. I used DIVA-GIS to re-grid and crop the LGM layers to match the parameters of the previous layers. Soil data and WWF ecoregions could not be included in LGM reconstructions because those data do not exist for

paleoclimate models. To investigate the Meadow Group ENM I combined the georeferenced occurrences for both *C. longilabris* and *C. nebraskana* into a single occurrence file.

For the "backcasting" several issues had to be addressed. First, ENMs were based on the contemporary environmental conditions that exist in the Nearctic, although the range of conditions during the LGM was different and included environmental extremes not encountered today. MAXENT's default setting is to treat variables outside the training range as if they were at the edge of their training range, referred to as "clamping" (Phillips *et al* 2006). The most conservative method of dealing with this issue is to remove these areas altogether, as they can be considered suspect if extensive. However, the latest version of MAXENT has incorporated an intermediate solution to the problem, implemented in the "fadebyclamping" option. This method calculates the degree of clamping and subtracts it from the total cumulative probability of occurrence (Phillips 2010) and my LGM reconstructions were generated using this option. Given that there were two separate LGM models used to create ENMs, I reconciled the outputs using a conservative approach modified after Waltari *et al* (2007) in order to avoid overprediction. First, models were created for both CCSM and MIROC, implementing the LPT and M10 thresholds as described above. Consensus models were created in DIVA-GIS that show the predicted areas as shades of green, with darker shades indicating increasingly stronger prediction by both models. The darkest green areas are those regions that were predicted as suitable habitat under the more stringent threshold by both models.

Testing for niche overlap and differentiation

In order to quantify niche similarity and test for significant differences between species ENMs, I used ENMTools (Warren *et al* 2010), a Perl script that allows for statistical comparisons between MAXENT outputs, employing randomization and resampling methods. First, I used the script to generate estimates of niche overlap incorporating all variables, conceptually based on two measures of similarity, Schoener's (1968) *D*, and a derivative of Hellinger's distance called *I* (Warren *et al* 2008). These distance-based similarity measures are obtained by comparing estimates of habitat suitability calculated for each grid cell of the study area, and are threshold independent. Following this, I used ENMTools to conduct tests of significance for niche differentiation. The "Identity test" addresses the null hypothesis that ENMs from two populations are identical. his test pools the occurrence points for a pair of populations, randomizing the population identities of the points, and extracting new population samples with the same sizes as the two original samples. Through a series of pseudoreplicated datasets the program creates a null distribution of niche similarity indices to which it compares the observed measures of niche overlap to test whether they are significantly outside of the null distribution. Lastly, I used ENMTools to implement "background tests" for allopatric taxa of interest. The previous Identity tests are most meaningful when species have the same suite of environmental conditions available to them, and this is unlikely to hold for allopatric species (Warren *et al* 2008). The background test can be used to

determine if two allopatrically distributed groups are more different than would be expected by chance, given the underlying environmental differences between the regions in which they occur. To accomplish this, ENMTools again generates a null distribution for the ENM differences expected from randomizations of the occurrence points. If the observed value of niche similarity is significantly *higher* than expected based on the null distribution, then the null hypothesis that similarity between the species is no more than expected based on the availability of habitat can be rejected. I conducted background tests between the generally allopatric subspecies of *C. longilabris*, and also between *C. sexguttata* and *C. denikei*.

Results

Model validation and variable contribution

In all cases, training and test AUC values indicated strong support for the fit and predictive power of the models (Table 4.2). Training and test AUC values for each taxon ranged from 0.897 – 0.999, and in every case the null hypothesis that test points are predicted no better than random could be rejected (all $p < 0.05$). Jackknife analyses of variable contribution revealed which environmental layers were most critical in limiting species distributions (Figures 4.1 - 4.5). In

most cases, the removal of any single variable had little negative effect on the ENM for any species, a result indicating that no single variable contains a substantial amount of useful variation that is not already contained in the other variables (Phillips *et al* 2006). This suggests that there is correlation between the environmental variables. However, In the case of *C. nebraskana*, some variables (*i.e.* precipitation of the coldest quarter) actually had a slight negative effect on the predictive power of the total model (Figure 4.2) when these were included.

A stronger indicator of variable importance is how each performs in isolation. For *C. longilabris*, annual mean temperature and ecoregion were most predictive of the species distribution, and both displayed a test AUC >0.85 when they were the only variable used (Figure 4.1), *e.g.* each was >85% predictive of the species distribution. Soil orders and maximum temperature of the warmest month were the next most predictive factors, both with values just under 0.85. For *C. nebraskana*, results were similar, with annual mean temperature and ecoregion again the two most predictive factors, both with AUC values also >0.85 (Figure 4.2). Soil orders and maximum temperature were also predictive at >0.80, but in contrast to *C. longilabris*, so were a number of other environmental variables. The results for *C. sexguttata and C. patruela* indicated that ecoregion and soil orders were the two most critical environmental factors in delineating the species range (Figures 4.3, 4.4). Maximum temperature during the warmest month was the third most predictive factor for *C. patruela* (0.93) compared to annual mean precipitation for *C. sexguttata* (0.89). For *C. denikei*, all but one variable displayed greater than 80% predictivity, and most were over 90%. This

is likely attributable to the species very limited geographic distribution and the inherent effect that has on the ENM process (see Discussion).

ENM predicted distributions

Maps of the species predicted distributions were in general concordance with known ranges (Figures 4.6 – 4.10). Differences in the amount of predicted area between LPT and M10 thresholds were most pronounced in *C. longilabris* and *C. sexguttata*. When *C. longilabris* populations were categorized according to Spanton's (1988) subspecies (Figure 4.11), each ENM predicted extensive suitable habitat beyond the observed distributions (Figure 4.12). For both *C. l. pervirids* and *C. l. laurentii*, over 50% of each predicted range was not occupied. Moreover, all three taxa overlapped in their predicted distributions in areas of the northern Rockies and adjacent highlands. This area of overlap included the distribution of phenotypic "intergrade" populations, as well as additional geographic areas where intergrades are not observed. ENMs were generated for the Meadow Group mtDNA phylogeographic clades (see Figure 2.2) and in each case there were extensive areas of predicted habitat where they do not occur (Figure 4.13). The predicted distributions of mtDNA clades overlapped throughout most of the higher elevation areas of the western United States.

Last Glacial Maximum reconstructions

Backcasting the Meadow Group ENM onto LGM conditions yielded predictions of suitable habitat throughout the southern part of the western U.S., with additional bands of potential habitat extending into the southeast (Figure 4.14). There were areas predicted under the more stringent threshold by both paleoclimate models, and these were distributed on the east slope of the Sierra Nevada and Cascade ranges, fragmented areas along the southern Great Basin, and additional areas mostly east and south of the Rockies. These stringently predicted areas include the hypothesized refugia for the each of the three major mtDNA clades (Figure 4.14B and C) based on prior phylogeographic and historical demographic results (Chapters II and III).

LGM models for *C. denikei* predicted very little suitable habitat under the CCSM model, corresponding to a small area of present-day western OK and TX panhandle (Figure 4.15). Most of this area was predicted only under the LPT threshold, and within this, an extremely small area (~10km wide) was predicted as suitable using the M10 threshold. No areas were predicted as habitable under either threshold using the MIROC model.

Niche overlap and differentiation

Niche overlap estimates varied considerably between taxon comparisons and between the two distance measures (Table 4.3). However, Identity tests

determined that each of the taxa (species, subspecies, and mtDNA clades) displayed unique niches; the null hypothesis of identical niche could be rejected in all comparisons (p <0.05). The lowest degree of species niche overlap was between *C. sexguttata* and *C. denikei* (D = 0.047, I = 0.342), whereas *C. sexguttata* and *C. patruela* exhibited the most similarity (D = 0.583, I = 0.716). The Background test revealed that *C. sexguttata* and *C. denikei* were more ecologically divergent than predicted by chance (p <0.05). For the *C. longilabris* subspecies, niche overlap was much higher between *C. l. laurentii* and *C. l. perviridis* (D = 0.580, I = 0.709) than between *C. l. longilabris* and either of these (D = 0.197, I = 0.470, and D = 0.165, I = 0.449 respectively). Furthermore, background tests revealed that *C. l. laurentii* and *C. l. perviridis* were not more significantly different in niche than expected (e.g. they were *more* ecologically similar than the null expectation), but in contrast *C. l. longilabris* was significantly more different in niche to these other subspecies than expected (both p <0.05). Meadow Group mtDNA clade comparisons showed results similar to the *C. longilabris* subspecies. Niche overlap was much higher between the Southwest and Northwest Clades (D = 0.559, I = 0.707) than between the Continental Clades and either of these (D = 0.208, I = 0.474, and D = 0.153, I = 0.430 respectively). The Background test further supported the similarity between the Southwest and Northwest Clades, as they were not significantly more different than would be expected by chance. Conversely, the Continental Clade was significantly different from the other two (both p <0.05).

Discussion

Factors limiting species ranges

Results of the jackknife tests were informative as to the environmental factors that may be delimiting the species ranges. Despite a lack of monophyly (see Chapter II) the Meadow Group nominal taxa, *C. longilabris* and *C. nebraskana*, were investigated as separate entities on the basis that they may represent "ecomorphs" or populations with genes under divergent selection for ecological characteristics (Michel *et al* 2010). Results indicated similar, but not identical, niches and limiting factors for *C. longilabris* and *C. nebraskana*. In the case of *C. longilabris*, annual mean and maximum temperature were two of the four environmental factors that best predicted occurrence (Figure 4.1). Given the observed distribution restricted to high latitudes (up to 67.4^{o} N, above the Arctic Circle) or high elevation (up to 3800m / 12,500ft), it is not surprising that *C. longilabris* is greatly affected by temperature. I have collected this species along the edge of snowmelt and glaciers and found it active at ambient air temperatures as low as 12.5^{o} C (55^{o} F), but very rarely observed the species active at temperatures above 30^{o} C (86^{o} F). Schultz *et al* (1992) performed physiological and behavioral studies in the field and lab, and found that *C. longilabris* had a lower tolerance for high body temperatures than any other *Cicindela* species examined. Furthermore, the authors repeated their

observations at four field sites (AZ, CO, WI, ME) and found no significant differences in preferred body temperature or thermoregulatory behavior between different geographic areas. In addition, the variables of ecoregion and soil type were also highly predictive for *C. longilabris* (AUC = 0.89, 0.84), consistent with Spanton's (1988) conclusions. He characterized the species as being associated with specific boreal and montane habitat types and edaphic conditions that were limited to particular soil orders. Spanton also hypothesized that differences in soil preference were responsible for divergence between *C. longilabris* and *C. nebraskana*. Jackknife analyses did show soil orders and ecoregion as highly predictive of the range of *C. nebraskana* as well (AUC = 0.81, 0.87) (Figure 4.2), although it is not possible from the ENM results to say how different the habitat and soil preferences are between the species, simply that these factors were key in explaining both distributions. For *C. nebraskana*, annual mean temperature was the most predictive single variable (AUC = 0.90), although in contrast to *C. longilabris*, little was previously known about the species physiology.

For the Forest Group, predictions by Kaulbars and Freitag (1993b) were supported by the jackknife analyses. They had concluded that "For all species of the group, habitats occupied and limits of distribution to eastern Canada and the USA appeared to be governed by soil and forest types". Consistent with this, jackknife tests demonstrated that ecoregion was the most predictive factor in all three species (Figures 4.4 - 4.6), and soil order was the second most predictive variable for both *C. sexguttata* and *C. patruela*, although not in *C. denikei*. It should be noted however, that in the latter species nearly all factors were highly

predictive (AUC >0.90 for all but two variables), so rank order of variable predictivity is less informative (see additional discussion below). For *C. sexguttata*, annual total precipitation was the third most predictive factor, and this is not surprising given that it inhabits “moist, and loamy soils” according to the previous authors (Kaulbars and Freitag 1993b). The third most predictive variable for *C. patruela* was maximum temperature of the warmest month. Although this species is largely sympatric with *C. sexguttata*, it has a more restricted range (Figures 4.8, 4.9). South of Ohio, *C. patruela* is only found at higher elevation areas of the Appalachian Mountains and Cumberland Plateau, presumably as a result of the species inability to tolerate higher temperatures at lower latitudes. By contrast, the variable of elevaton is actually the poorest single predictor of *C. patruela* distribution, however this is not surprising given that the species occurs at sea level or low elevation in the northern part of its range and therefore this variable would not be particularly predictive in the absence of other factors. For *C. denikei*, nearly all variables predicted the species distribution well, and nine of these exhibited AUC values over 0.95 (Figure 4.5). The restricted distribution of the species presents an inherent challenge to the inference capabilities of ENM methods (Phillips *et al* 2006). Because all occurrences occur within close proximity to each other, there is a high degree of specificity in the prediction, as the model has a very narrow set of conditions to train from. This specificity may represent the true fundamental niche (*e.g.* the species may actually have a very narrow of a range of tolerances), or it could be

that the model is underpredicting the true geographic distribution due to tight correlation of conditions at the few known occurrences.

Species distribution models

In general, the ENM predicted species distributions fit the known distributions well (Figures 4.6 – 4.10) suggesting that abiotic environmental factors are primarily responsible for limiting species ranges, not biotic interactions (*i.e.* competitive or mutualistic). Although in some cases there were notable differences in the extent of predicted habitat between the two thresholds. The discrepancy was most pronounced in *C. longilabris*, and the LPT predicted greater than three times the amount of suitable habitat as the M10 threshold (fractional predicted area 37% vs. 11%). One possibility is that outlier occurrences are responsible (Phillips *et al* 2006). If some of the *C. longilabris* localities represent dispersers into suboptimal habitat, then those points may exhibit an atypical set of environmental conditions. The more stringent M10 threshold will identify these as outliers and predict a reduced subset of habitat compared to the LPT. Alternatively, the ENM may have mischaracterized part of the niche as the result of ecological divergence within the species. Statistical tests demonstrated that northern populations (the putative subspecies *C. l. longilabris*) were more ecologically different from the rest of the species (*C. l. perv*iridis and *C. l. laurentii*) than expected by chance ($p < 0.05$). As such, entire sets of populations may be “outliers” because the niche is not identical across the

species range. It is also worth noting that there are parts of Alaska, Yukon, and northern British Columbia that were not predicted as suitable habitat, although *C. longilabris* is generally believed to be common and widespread in much of that region (Spanton 1988; Knisley *pers. comm*. 2005; Sikes *pers. comm*. 2006). This result is likely due to collector sampling bias overrepresenting the more accessible areas in the United States and southern Canada, and underrepresenting the extreme northwestern corner of the continent. These unsampled areas may contain environmental conditions that fall outside of the range encountered by MAXENT's training model, and consequently the ENM will fail to accurately assess their suitability.

In the case of *C. sexguttata*, the thresholds differed in habitat prediction primarily west of the Mississippi River (Figure 4.8). It is possible that this discrepancy is due to the patchier nature of moist deciduous forest in the region, which the more stringent threshold fails to accomodate. Due to the geographic scale of the MAXENT analyses, habitats smaller than 1km^2, or suitable microhabitats within otherwise suboptimal habitat cannot be identified. In addition, there is again the possibility of some degree of local adaptation at the edge of the species range into the Midwest. Lastly, there may also be a collector sampling bias towards the more heavily populated East Coast metropolitan areas.

ENMs have been used to identify new populations of species (*e.g.* Menon *et al* 2010), but could also be used to validate older or dubious records. Raxworthy *et al* (2003) used ENMs for chameleon species in Madagascar and

discovered additional unsampled habitat in geographically disjunct areas. Subsequent sampling yielded undescribed species that were closely related and ecologically similar to the modeled taxa. The potential exists for similar discoveries with tiger beetles. For *C. longilabris*, two mountain ranges in southeastern AZ were predicted as highly suitable, Mt. Graham, and the Chiracahua Mts. (highest points), although the species has not been recorded from either. If *C. longilabris* were found at these sites it would extend the known range south, and the results of the study warrant investigations into these areas. In addition, the ENM predicted several kilometers of habitat in the vicinity of Spruce Knob, WV, potentially validating a dubious record in the University of West Virginia collection (Knisley *pers. comm*. 2010). If *C. longilabris* were present there, it would extend the known range ~450 miles from the nearest established population in the Adirondacks.

Last Glacial Maximum reconstructions

Phylogeography and historical demography analyses uncovered deeply separated allopatric mtDNA clades within the Meadow Group (Chapters II, III), consistent with divergence during the early-mid Pleistocene Ice Ages. If this hypothesis about the history of the North American *C. sylvatica* group is correct, then populations representing each of these mtDNA clades must have persisted in refugia when glacial retreat began at the LGM, 18,000 years ago. Results from the consensus paleoclimate model indicated that suitable habitat did exist at

this point, and occurred in the geographic areas where glacial refugia were hypothesized to exist (Figure 4.14). During the LGM, habitat appears to have been most extensive along the eastern slope of the Sierra Nevada and Cascade ranges, and east of the front range of the Rockies. Suitable habitats within the Great Basin were more extensive than in the present, but also fragmented. Post-glacial dispersal routes exist, as these LGM habitats are broadly connected to the present-day Meadow Group range.

Paleoclimate models were able to address the hypotheses pertaining to divergence between *C. denikei* and *C. sexguttata.* If *C. denikei* evolved *in situ* in its current habitat, then it could only have diverged less than 11,000 years ago (CLIMAP 1981; Sibrava *et al* 1986; Mix *et al* 2001), potentially insufficient time to exhibit reciprocal monophyly with respect to *C. sexguttata* based on estimates from coalescent analyses. Alternatively, it is possible that *C. denikei* could have diverged from *C. sexguttata* prior to the LGM in a separate geographic area and later dispersed to its current location following glacial retreat. If this second scenario were true, then the case for incomplete lineage sorting is less compelling, as it would indicate that the divergence was more ancient. However, the consensus paleoclimate model did not support the existence of suitable habitat during the LGM (Figure 4.15), because no area was predicted by both models. This finding is consistent with the hypothesis that if *C. denikei* is a distinct species, it evolved *in situ* after glacial ice retreated and created the species current alvar habitat (Kaulbars and Freitag 1993a,b).

Ecological niche overlap and differentiation

Statistical tests of niche overlap indicated that no two taxa were identical in niche (Table 4.3). This result bears on the systematic relationship of the Meadow Group nominal taxa. Although prior work demonstrated that *C. longilabris* and *C. nebraskana* were extensively polyphyletic with respect to mtDNA and did not form discreet clusters based on genome-wide AFLP analyses, these ENM results suggest that there may still be a genetic basis underlying the ecological differences in these taxa. Following up on these results, I have begun a genome-scan study to potentially identify loci under selection that could be associated with ecological or morphological divergence (*e.g.* Bonin *et al* 2006, Egan *et al* 2008). Within the Forest Group, results were also informative as to the hypothesized relationships between taxa. *C. sexguttata* and *C. denikei* displayed the smallest degree of niche overlap of any comparison in the study (D = 0.047, I = 0.342), although this is expected when species are allopatric (Warren et al 2008). The Background test is a more relevant test in such a case, and accounts for the inherent underlying differences between non-overlapping geographic areas. Still, it was found that *C. denikei* and *C. sexguttata* were significantly more ecologically divergent than expected by chance. Moreover, the observation that these two taxa are the most divergent in niche while still exhibiting low genetic differentiation is consistent with a peripatric ecological speciation scenario (Funk *et al* 1995). This supports the hypothesis that *C. denikei* does not simply represent a set of satellite populations

for *C. sexguttata*, but instead may have diverged and ecologically differentiated in its unique habitat, as suggested by previous authors (Kaulbars and Freitag 1993a, b).

Tests of niche overlap between *C. longilabris* subspecies yielded interesting and complicated results. The two montane subspecies, *C. l. perviridis* and *C. l. laurentii* (Figure 4.11) were not found to be significantly different in niche (Table 4.3). Furthermore, these broadly overlapped in predicted geographic distribution (Figure 4.12). In contrast, *C. l. longilabris* was found to be significantly ecologically differentiated from the other subspecies, although its predicted distribution also included a large area of geographic overlap with the others. All three subspecies were predicted to occur in the geographic area that Spanton (1988) described as a place of mixing ("hybridization") between the subspecies, resulting in what he termed "intergrades". It is tempting to conclude that the ENMs support this conclusion, but the results are not entirely consistent with this hypothesis. First, additional areas of overlap are predicted, yet phenotypic intergrades have not been observed in these (Figure 4.12). Second, molecular evidence does not support genetic breaks that match subspecies ranges (see Chapter II, Figs. 2.3, 2.5, 2.7). For additional discussion of subspecies validity see Chapters V, VI.

In the Meadow Group, mtDNA clades are tightly allopatric, despite the absence of obvious physiographic barriers to dispersal. ENMs revealed that their predicted distributions are more extensive than observed (Figure 4.13), and each was predicted to occur in areas overlapping with the other clades' distribution.

This was most pronounced between the Northwest and Southwest Clades, as contiguous suitable habitat is predicted for each clade throughout the Sierra Nevada range and adjacent Great Basin highlands (Figure 4.13C). This pattern coupled with the abrupt contact zone supports an extrinsic biological cause for range limits, such as competitive exclusion (Swenson 2006, Sanchez-Cordero *et al* 2008) or the effect of *Wolbachia* infections. If mtDNA clades are associated with different *Wolbachia* strains, then cytoplasmic incompatibility (Hoffman and Turelli 1997) could prevent a barrier to reproduction. Preliminary data are being gathered to address this last possibility.

Conclusions

ENMs were used to quantify niches and identify the factors responsible for limiting distributions in a group of North American tiger beetles. Results broadly supported prior researchers' hypotheses about the determinants of tiger beetle species ranges (Leffler 1979; Spanton 1988; Schultz *et al* 1992; Kaulbars and Freitag 1993b), with habitat, soil type, and temperature the primary factors. Hypotheses pertaining to species boundaries were addressed. In the case of *C. denikei*, niche models and subsequent tests of significance revealed that this taxon was more ecologically divergent from its putative parental species, *C. sexguttata*, than would be expected by chance. The validity of *C. denikei* as a

recently diverged separate species was further bolstered by Pleistocene habitat reconstructions that indicated no available habitat 18,000 years ago (Last Glacial Maximum). As such, the species appears to have evolved more recently *in situ* in its current alvar habitat, and this scenario supports incomplete lineage sorting as an explanation for the species lack of monophyly with respect to *C. sexguttata*. Pleistocene models were also used to address phylogeographic hypotheses pertaining to the "Meadow Group", one of the two ecological divisions within the North American *C. sylvatica* species group. Prior molecular studies revealed three deep genetic subdivisions within the Meadow Group and coupled with historical demographic analyses, isolation in glacial refugia was proposed as an explanation for the patterns. Paleoclimate ENMs revealed the presence of suitable habitat at the LGM and these areas included the hypothesized refugia for each clade.

In addition to species-level questions, ENMs were used to explain the distribution of subspecies and mtDNA clades. *C. longilabris* subspecies exhibited significant niche divergence between the boreal forest dwelling *C. l. longilabris* and more southerly-distributed montane subspecies, *C. l. perviridis* and *C. l. laurentii*. These last two were found to be more ecologically similar than expected, given their allopatry. For mtDNA clades, predicted niche overlap was considerable and coupled with their observed patterns of occurrence, extrinsic factors were most likely the cause of their distributions, either competitive interactions between them, or possibly a due to different Wolbachia strain associations resulting in cytoplasmic incompatibility (Hoffman and Turelli 1997).

Lastly, this study identified additional areas of potentially suitable habitat for *C. longilabris* in isolated mountain ranges to the south of the species known distribution. These areas warrant further investigation and if populations are found, they may represent a range extension, genetically distinct subdivisions, or possibly new species (Raxworthy *et al* 2003).

Acknowledgements

I would like to thank the following tiger beetle researchers and collectors who provided me with detailed species records: Dave Brzoska, Mike Kippenhan, Todd Lawton, Jason Schmidt, Barry Knisley, Richard Freitag, Rob Foster, Michael Oldham, John Stamatov, Richard Waldrep, Wayne Steffens, Kris Simpson, Jonathan Mawdsley, Doug Backlund, Scott Weins, Derek Sikes, David Melius. I would like to also thank Robert Guralnick and Eric Waltari for discussions and advice related to niche modeling theory and implementation. Rodger Gwiazdowski and Scott Egan reviewed parts of the chapter and provided helpful feedback on text and figures. Lastly, I would like to dedicate this chapter to the memory of Tim Spanton. It was his detailed and thorough research conducted over twenty years ago that served as a great inspiration for the entire project, and this chapter would not have been possible without his contributions.

References

Anderson, R. P., M. Gómez-Laverde, and A. T. Peterson. Geographical distributions of spiny pocket mice in South America: insights from predictive models. Glob. Ecol. Biogeogr. 11:131–141.

Braconnot, P. B. Otto-Bliesner, S. Harrison, S. Joussaume, *et al* 2007. Results of PMIP2 coupled simulations of the Mid-Holocene and Last Glacial Maximum - Part 1: experiments and large-scale features Climate of the Past, Volume 3, Number 2, 261-277.

Bonin, A., P. Taberlet, C. Miaud, F. Pompanon. 2006. Explorative genome scan to detect candidate loci for adaptation along a gradient of altitude in the common frog (Rana temporaria). Mol. Biol. Evol. 23(4): 773-783.

Cicero, C. 2004. Barriers to sympatry between avian sibling species (Paridae: Baeolophus) in local secondary contact. Evolution 58:1573-1587.

CLIMAP. 1981. Seasonal reconstructions of the Earth's surface at the last glacial maximum in Map Series, Technical Report MC-36. Boulder, Colorado: Geological Society of America.

Collins, W.D., M.L. Blackmon, G.B. Bonan, J.J. Hack, T.B. Henderson, J.T. Kiehl, W.G. Large, and D.S. McKenna. 2006. The Community Climate System Model version 3 (CCSM3), Journal of Climate 19: 2122–2143.

Connell, J.H. 1980. Diversity and the coevolution of competitors, or the ghost of competition past. Oikos 35, 131-138.

Dreisig, H. 1980. Daily activity, thermoregulation, and water loss in the tiger beetle *Cicindela hybrida*. Oecologia, 44: 376-389.

Egan, S.P., Nosil, P., Funk D.J. 2008. Selection and genomic differentiation during ecological speciation: Isolating the contributions of host association via a comparative genome scan of *Neochlamisus bebbianae* leaf beetles. Evolution 62:1162–1181.

Elith, J., C.H. Graham, R.P. Anderson, et. Al 2006. Novel methods improve prediction of species distributions from occurrence data. Ecography 29(2):129-151.

Fielding, A.H. and Bell, J.F. (1997) A review of methods for the assessment of prediction errors in conservation presence/absence models. Environmental Conservation, 24, 38–49.

Funk, D.J., D.J. Futuyma, G. Orti, and A. Meyer. 1995. A history of host associations and evolutionary diversification for Ophraella (Coleoptera: Chrysomelidae): new evidence from mitochondrial DNA. Evolution 49: 1008-1017.

Graves, R. C. 1981. Offshore flight in *Cicindela trifasciata*. Cicindela 13:45–47.

Hadley, N.F. and T.D. Schultz. 1987. Water loss in three species of tiger beetles (Cicindela): Correlations with epicuticular hydrocarbons. Journal of Insect Physiology 33(10): 677-682.

Hasumi, H. and S. Emori. 2004: K-1 coupled GCM (MIROC) description. K-1 Tech. Rep. 1: 1-39.

Hoffmann, A.A., and M. Turelli. 1997. Cytoplasmic incompatibility in insects. In: R.V. O'Neill, A.A. Hoffmann and J.H. Werren, Editors, *Influential Passengers*, Oxford University Press, Oxford, UK, pp. 42–80.

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. 2005a. Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology 25: 1965-1978.

Hijmans, R.J., Guarino, L., Jarvis, A., O'Brien, R., Mathur, P., Bussink, C., Cruz, M., Barrantes, I. and Rojas, E. 2005b. DIVA-GIS Version 5.2 manual Available at: www.diva-gis.org

Jiménez-Valverde, A. and Lobo, J.M. (2007) Threshold criteria for conversion of probability of species presence to either–or presence–absence. Acta Oecologica, 31, 361–369.

Kaulbars, M.M., and R. Freitag 1993a. Foraging behaviour of the tiger beetle *Cicindela denikei* Brown (Coleoptera: Cicindelidae). Can. Field-Nat. 107:53-58.

Kaulbars, M.M., and R. Freitag 1993b. Geographic variation, classification, reconstructed phylogeny, and geographic history of the *Cicindela sexguttata* group (Coleoptera: Cicindelidae). Can. Entomol. 125:267-316.

Kippenhan, M. G. 1994. The Tiger Beetles (Coleoptera: Cicindelidae) of Colorado. Trans. Amer. Entomological Society 120(1):1-86.

Knisley, C. B. and T. D. Schultz. 1997. The Biology of Tiger Beetles and a Guide to the Species of the South Atlantic States. Virginia Museum of Natural History, Martinsville, VA. 210 pp.

Kozak, K. H., C. H. Graham, and J. J. Wiens. 2008. Integrating GIS data into evolutionary studies. Trends in Ecology and Evolution 23:141–148.

Leffler, S.L. and D.L. Pearson. 1976. The tiger beetles of Washington. Cicindela 8:21-60.

Leffler, S. R. 1979. Tiger beetles of the Pacific Northwest (Coleoptera: Cicindelidae). Ph.D. Dissertation, University of Washington, Seattle. 731 pp.

Levine, R. S., A. T. Peterson, K. L. Yorita, D. Carroll, I. K. Damon, and M. G. Reynolds. 2007. Ecological niche and geographic distribution of human monkeypox in Africa. *PLoSONE* 2(1):e176. doi:10.1371/journalpone.0000176.

Liu, C., Berry P.M., Dawson T. P. Pearson, R.G. 2005. Selecting thresholds of occurrence in the prediction of species distributions. Ecography, 28, 385-393.

Lobo, J.M., A. Jiménez-Valverde, and R. Real 2008. AUC: a misleading measure of the performance of predictive distribution models. Global Ecology and Biogeography 17(2): 145-151.

Mayr, E. 1970. Populations, species, and evolution. Belknap, Cambridge, MA.

Menon, S., B.I. Choudhury, M.L. Khan, A.T. Peterson. 2010. Ecological niche modeling and local knowledge predict new populations of Gymnocladus assamicus, a critically endangered tree species. Endangered Species Research 11: 175-181.

Michel, A.P., S. Sim, T.H.Q. Powell, M.S. Taylor, P. Nosil, and J.L. Feder. 2010. Widespread genomic divergence during sympatric speciation. Proceedings of the National Academy of Sciences of the United States of America 107(21): 9724-9729.

Mix AC, Bard E, Schneider R. 2001. Environmental processes of the ice age: land, oceans, glaciers (EPILOG). Quaternary Science Reviews 20: 627-657.

Nix, H.A. 1986. A biogeographic analysis of Australian elapid snakes. In: Atlas of elapid snakes of Australia. Australian Flora and Fauna Series No. 7 (R. Longmore, Ed.), pp.4-15. Australian Government Publishing Service, Canberra.

Nosil, P., S.P. Egan, and D.J. Funk. 2007. Heterogeneous genomic differentiation between walking-stick ecotypes: “Isolation by adaptation” and multiple roles for divergent selection. Evolution 62(2): 316-336.

Nosil P, Funk D.J., Ortiz-Barrientos D. 2009. Divergent selection and heterogeneous genomic divergence. Molecular Ecology, 18, 375–402.

Pearson, D.L., and E.J. Mury 1979. Character divergence and convergence among tiger beetles (Coleoptera: Cicindelidae). Ecology 60:557-566.

Pearson, D.L. 1980. Patterns of limiting similarity in tropical forest tiger beetles (Coleoptera: Cicindelidae). Biotopica 12:195-204.

Pearson, D.L., and C.B. Knisley. 1985. Evidence for food as a limiting resource in the life cycle of tiger beetles (Coleoptera: Cicindelidae). Oikos 45:161-168.

Pearson, D.L. and S.A. Juliano. 1991. Mandible length ratios as a mechanism for co-occurrence: Evidence from a world-wide comparison of tiger beetle assemblages (Cicindelidae). Oikos 61:223-233.

Pearson, D. L., T. G. Barraclough, and A. P. Vogler. 1997. Distibutional maps for North American species of tiger beetles (Coleoptera: Cicindelidae). Cicindela 20:33-84.

Pearson, R.G., Thuiller, W., Araújo, M.B., Brotons, L., Martinez-Meyer, E., McClean, C., Miles, L., Segurado, P., Dawson, T.P., Lees, D. 2006. Model-based uncertainty in species’ range prediction. *Journal of Biogeography*, 33, 1704-1711.

Pearson, R.G. 2007. Species' distribution modeling for conservation educators and practitioners. Synthesis. American Museum of Natural History. Available at http://ncep.amnh.org

Pearson, R.G., Raxworthy, C.J., Nakamura, M. and Peterson, A.T. 2007. Predicting species’ distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. Journal of Biogeography 34: 102-117.

Olson, D. M, E. Dinerstein, E.D. Wikramanayake, N.D. Burgess, G.V.N. Powell, E.C. Underwood, J.A. D'amico, I. Itoua, H.E. Strand, J.C. Morrison, C.J. Loucks, T.F. Allnutt, T.H. Ricketts, Y. Kura, J.F. Lamoreux, W.W.Wettengel, P. Hedao, and K.R. Kassem. 2001. Terrestrial Ecoregions of the World: A New Map of Life on Earth. BioScience 51:933-938

Pearson, D.L., and E.J. Mury. 1979. Character divergence and convergence among tiger beetles (Coleoptera: Cicindelidae). Ecology 60:557-566.

Pearson, D.L. 1980. Patterns of limiting similarity in tropical forest tiger beetles (Coleoptera: Cicindelidae). Biotopica 12:195-204.

Pearson, D.L., and C.B. Knisley. 1985. Evidence for food as a limiting resource in the life cycle of tiger beetles (Coleoptera: Cicindelidae). Oikos 45:161-168.

Pearson, D.L. and S.A. Juliano. 1991. Mandible length ratios as a mechanism for co-occurrence: Evidence from a world-wide comparison of tiger beetle assemblages (Cicindelidae). Oikos 61:223-233.

Pearson, D. L., T. G. Barraclough, and A. P. Vogler. 1997. Distibutional maps for North American species of tiger beetles (Coleoptera: Cicindelidae). Cicindela 20:33-84.

Phillips, S.J., R.P. Anderson, and R.E. Schapire. 2006. Maximum entropy modeling of species geographic distributions. Ecological Modelling 190: 231-259.

Raxworthy, C.J., E. Martinez-Meyer, N. Horning, and R.A. Nussbaum. 2003. Predicting distributions of known and unknown reptile species in Madagascar. Nature 426, 837–841.

Remington, C.L. 1968. Suture-zones of hybrid interaction between recently joined biotas. pp. 321–428. In Evolutionary Biology (Eds.) T. Dobzhansky, M.K. Hecht, and W.C. Steere, Plenum Press, New York.

Richmond, G.M. and D.S. Fullerton, 1986, Summation of Quaternary glaciations in the United States of America. Quaternary Science Reviews. vol. 5, pp. 183-196.

Sanchez-Cordero, V., D. Stockwell, S. Sarkar, H. Liu, C.R. Stephens, J. Gimenez. 2008. Competitive interactions between felid species may limit the southern distribution of bobcats Lynx rufus. Ecography 31(6): 757-764.

Schincariol, L.A., and R. Freitag. 1991. Biological character analysis, classification, and history of the North American *Cicindela splendida* group taxa (Coleoptera: Cicindelidae). Can. Entomol. 123:1327-1353.

Schoener, T W. 1968. Anolis lizards of Bimini: resource partitioning in a complex fauna. Ecology 49:704–726.

Schultz, T. D. and N.F. Hadley. 1987. Structural colors of tiger beetles and their role in heat transfer through the integument. Physiological Zoology 60:737-745.

Schultz, T.D., M.C. Quinlan, and N.F. Hadley. 1992. Preferred body temperature, metabolic physiology, and water balance of adult *Cicindela longilabris*: A comparison of populations from boreal habitats and climatic refugia. Physiological Zoology 65(1): 226-242

Schultz, T.D. 1998. The utilization of patchy thermal microhabitats by the ectothermic insect predator, Cicindela sexguttata. Ecological Entomology 23: 444-450.

Sibrava, V., Bowen, D.Q, and Richmond, G.M., 1986, Quaternary Glaciations in the Northern Hemisphere, Quaternary Science Reviews. vol. 5, pp. 1-514.

Soberón, J., and A. T. Peterson. 2005. Interpretation of models of fundamental ecological niches and species' distributional areas. *Biodiversity Informatics*, 2:1-10.

Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture. U.S. General Soil Map (STATSGO2). http://soildatamart.nrcs.usda.gov (accessed July 2008)

Spanton, T.G. 1988. The *Cicindela sylvatica* group: Geographical variation and classification of the Nearctic taxa and reconstructed phylogeny and geographical history of the species (Coleoptera: Cicindelidae). Quaestiones Entomologicae 24: 51-161.

Stockwell, D.R.B., and I.R. Nobel 1992. Induction of sets of rules from animal distribution data: A robust and informative method of data analysis. Math. Comp. Simul. 32:249-254.

Svenning, J.C., S. Normand, F. Skov. 2006. Range filling in European trees. Journal of Biogeography 33, 2018–2021

Swenson, N.G. and D.J. Howard. 2004. Do suture zones exist? Evolution 58(11): 2391-2397.

Swenson, N.G. and D.J. Howard. 2005. Clustering of contact zones, hybrid zones, and phylogeographic breaks in North America. Am. Nat. 166: 581-591.

Swenson, N.G. 2006. GIS-based niche models reveal unifying climatic mechanisms that maintain the location of avian hybrid zones in a North American suture zone. Journal of Evolutionary Biology19: 717-725.

Vitt, L.J., P.A. Zani, and M.C. Esposito. 1999. Historical ecology of Amazonian lizards: implications for community ecology. Oikos 87: 286–294

Waltari, E., R. Hijmans , A. Peterson , Á. Nyári , S. Perkins, and R. Guralnick. 2007. Locating Pleistocene Refugia: Comparing Phylogeographic and Ecological Niche Model Predictions. PLoS ONE 2(7): e563.

Waltari, E. and R. Guralnick. 2009. Ecological niche modeling of Great Basin montane mammals: examining past and present connectivity of species across basin and ranges. Journal of Biogeography 36(1): 148-161

Warren, D.L., R. E. Glor, and M. Turelli. 2008. Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. Evolution 62:2868-2883.

Warren, D.L., R. E. Glor, and M. Turelli. 2010. ENMTools: a toolbox for comparative studies of environmental niche models. Ecography 33(1): DOI: 10.1111/j.1600-0587.2009.06142.x

Wiens, J.J. and C.H. Graham. 2005. Niche conservatism: integrating evolution, ecology, and conservation biology. Annual Review of Ecology, Evolution, and Systematics 36:519–539.

Zink, R.M. 2004. The role of subspecies in obscuring avian biological diversity and misleading conservation policy. Proc. R. Soc. Lond. B 271: 561-564.

Table 4.1. Museum collections surveyed for species locality data.

Academy of Natural Sciences	Philadelphia, PA
American Museum of Natural History	New York, NY
Brigham Young University	Provo, UT
Clemson University	Clemson, SC
Colorado State University	Fort Collins, CO
Museum of Comparative Zoology	Cambridge, MA
Natural History Museum (UK)	London, UK
Rutgers University	New Brunswick, NJ
Smithsonian Museum of Natural History	Washington, D.C.
University of California at Davis	Davis, CA
University of Georgia	Athens, GA
University of Massachusetts at Amherst	Amherst, MA
University of Missouri at Columbia	Columbia, MO

Table 4.2. Number of occurrences for each taxon and ecological niche model support statistics. AUC (area under the receiver operating characteristic) is a measure of model performance (Fielding and Bell 2002, Phillips *et al.* 2006). AUC ranges from 0-1, with higher values indicative of a greater fit of the model to the data. An AUC of 0.5 would be consistent with a random prediction, and 1.0 indicates a perfect fit. Values over 0.9 have been considered high support for the fit of the model to the data (Fielding and Bell 2002). The Test AUC is the measure of how well the training model fit the test data that have been set aside to validate the model (random 25% of points). The test AUC is the best single indicator of ENM predictive power (Phillips et al 2006). Tests of significance were conducted in MAXENT and the null hypothesis that test points are predicted no better than random can be rejected in all cases (p <0.05).

Taxon	# of presences	Training AUC	Test AUC
Species			
C. longilabris	262	0.960	0.920
C. nebraskana	136	0.983	0.910
C. sexguttata	227	0.941	0.918
C. denikei	51	0.999	0.953
C. patruela	106	0.986	0.977
C. longilabris subspecies (no “intergrades”)			
C. l. longilabris	153	0.968	0.898
C. l. laurentii	44	0.996	0.994
C. l. perviridis	27	0.996	0.987
Clade			
Meadow Group	375	0.944	0.897
Continental Clade	60	0.944	0.939
Southwest Clade	13	0.996	0.981
Northwest Clade	25	0.993	0.995

Table 4.3. Measures of niche overlap and differentiation. Statistical tests were implemented using the software package ENMTools (Warren *et al.* 2010) to examine niche similarity using two measures, based on Schoener's (1968) *D*, and a derivative of Hellenger's Distance *I* (Warren *et al.* 2008). Identity tests evaluate the null hypothesis that species are identical in niche, and Background tests evaluate whether allopatric taxa are more different than would be expected by chance, given the underlying environmental differences between the regions in which they occur. No taxon comparisons displayed identical niches. In two comparisons, allopatric taxa were more similar than null expectations: the subspecies *C. l. laurentii* and *C. l. perviridis*, and the Southwest and Northwest mtDNA clades.

Taxon	Overlap (D)	Overlap (I)	Identity test	Background test
Species				
C. longilabris – C.nebraskana	0.378	0.578	$p<0.05$	-
C. sexguttata – C. patruela	0.583	0.716	$p<0.05$	-
C. sexguttata – C. denikei	0.047	0.342	$p<0.05$	$p<0.05$
C. longilabris subspecies				
C. l. longilabris – C. l. laurentii	0.197	0.470	$p<0.05$	$p<0.05$
C. l. longilabris – C. l. perviridis	0.165	0.449	$p<0.05$	$p<0.05$
C. l. laurentii – C. l. perviridis	0.580	0.709	$p<0.05$	NS
Meadow Group mtDNA clades				
Continental – Northwest	0.153	0.430	$p<0.05$	$p<0.05$
Continental – Southwest	0.208	0.474	$p<0.05$	$p<0.05$
Southwest – Northwest	0.559	0.707	$p<0.05$	NS

Figure 4.1. Jackknife analysis of variable contribution to *C. longilabris* ecological niche model. Environmental variables were selected because they are likely to limit species distributions (Leffler 1979; Spanton 1988; Hijmans *et al.* 2005a). The red bar represents the test AUC (area under the receiver operating characteristic), a measure of the model predictive performance (Phillips *et al.* 2006). Dark blue bars indicate how well each variable predicts the species niche when used alone. Light blue bars show model performance when the variable is removed. Because environmental variables are often correlated, removal of a single variable often has little effect on the total model (Phillips *et al.* 2006).

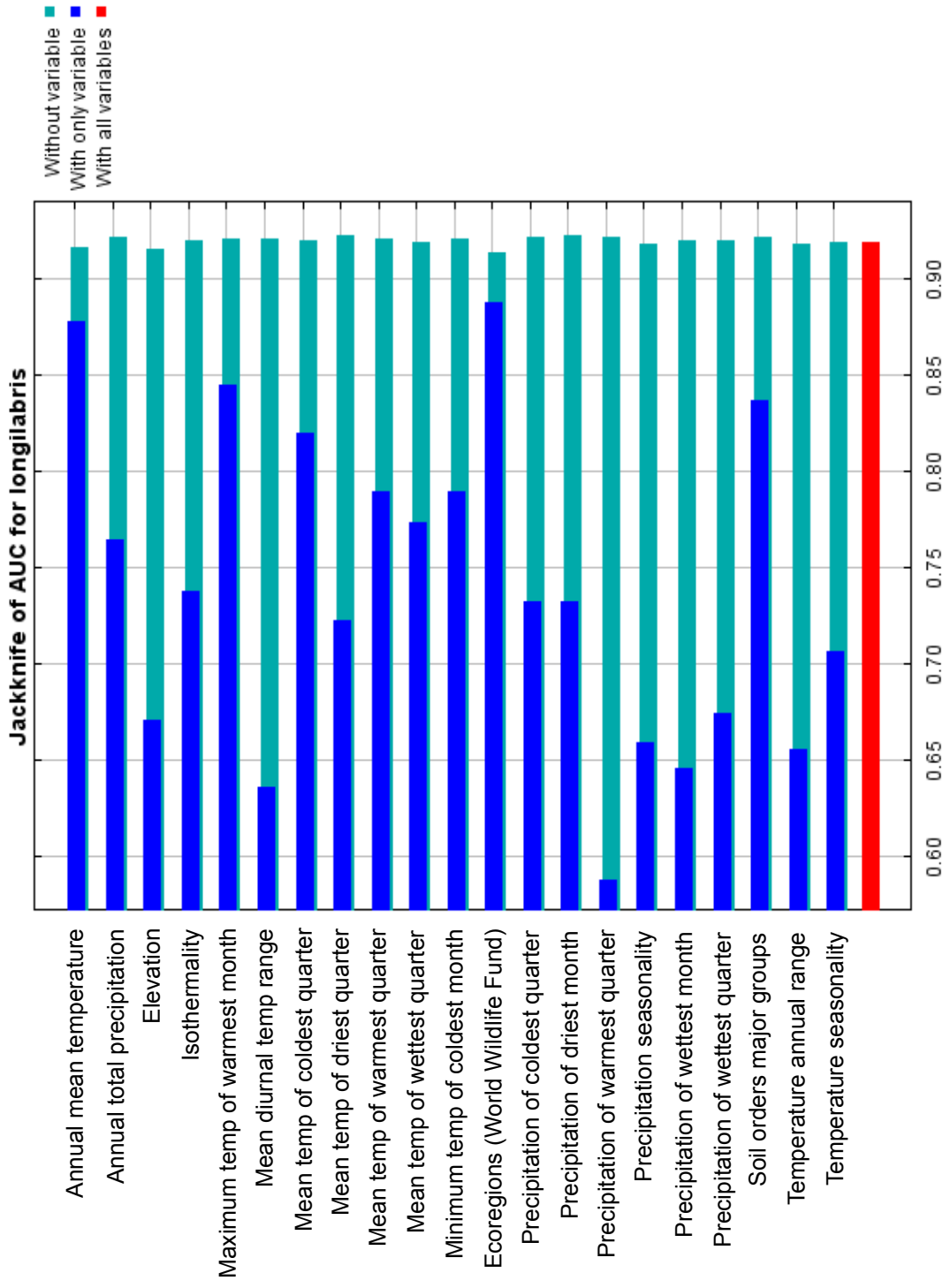

Jackknife of AUC for longilabris
Without variable
With only variable
With all variables
Annual mean temperature
Annual total precipitation
Elevation
Isothermality
Maximum temp of warmest month
Mean diurnal temp range
Mean temp of coldest quarter
Mean temp of driest quarter
Mean temp of warmest quarter
Mean temp of wettest quarter
Minimum temp of coldest month
Ecoregions (World Wildlife Fund)
Precipitation of coldest quarter
Precipitation of driest month
Precipitation of warmest quarter
Precipitation seasonality
Precipitation of wettest month
Precipitation of wettest quarter
Soil orders major groups
Temperature annual range
Temperature seasonality
0.60
0.65
0.70
0.75
0.80
0.85
0.90

Figure 4.2. Jackknife analysis of variable contribution to *C. nebraskana* ecological niche model. Environmental variables were selected because they are likely to limit species distributions (Leffler 1979; Spanton 1988; Hijmans *et al.* 2005a). The red bar represents the test AUC (area under the receiver operating characteristic), a measure of the model predictive performance (Phillips *et al.* 2006). Dark blue bars indicate how well each variable predicts the species niche when used alone. Light blue bars show model performance when the variable is removed. Because environmental variables are often correlated, removal of a single variable often has little effect on the total model (Phillips *et al.* 2006). Note that in some cases, the model actually performs slightly better when a variable is removed.

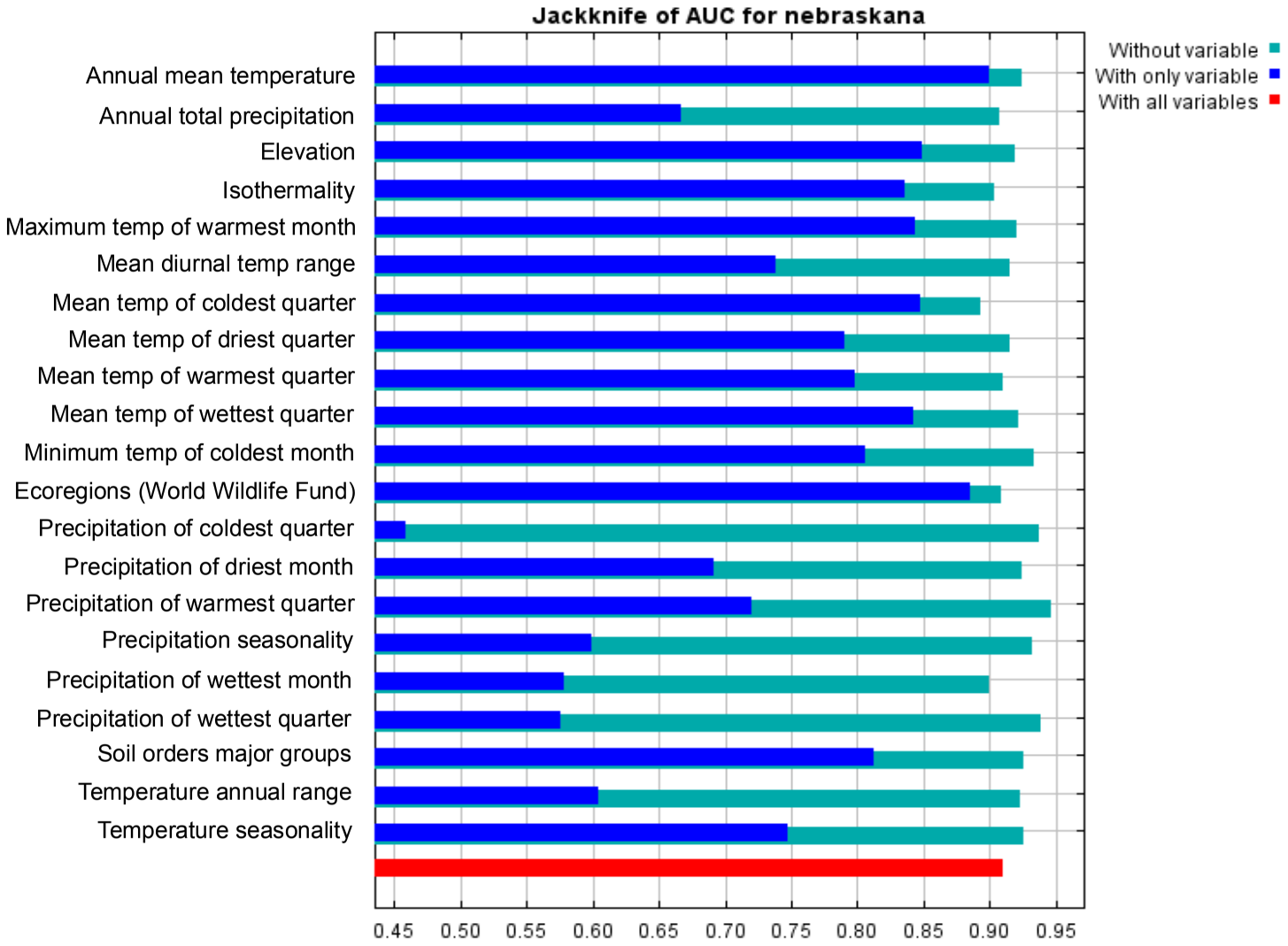

Jackknife of AUC for nebraskana
Annual mean temperature
Annual total precipitation
Elevation
Isothermality
Maximum temp of warmest month
Mean diurnal temp range
Mean temp of coldest quarter
Mean temp of driest quarter
Mean temp of warmest quarter
Mean temp of wettest quarter
Minimum temp of coldest month
Ecoregions (World Wildlife Fund)
Precipitation of coldest quarter
Precipitation of driest month
Precipitation of warmest quarter
Precipitation seasonality
Precipitation of wettest month
Precipitation of wettest quarter
Soil orders major groups
Temperature annual range
Temperature seasonality
Without variable
With only variable
With all variables
0.45
0.50
0.55
0.60
0.65
0.70
0.75
0.80
0.85
0.90
0.95

Figure 4.3. Jackknife analysis of variable contribution to *C. sexguttata* ecological niche model. Environmental variables were selected because they are likely to limit species distributions (Leffler 1979; Spanton 1988; Hijmans *et al.* 2005a). The red bar represents the test AUC (area under the receiver operating characteristic), a measure of the model predictive performance (Phillips *et al.* 2006). Dark blue bars indicate how well each variable predicts the species niche when used alone. Light blue bars show model performance when the variable is removed. Because environmental variables are often correlated, removal of a single variable often has little effect on the total model (Phillips *et al.* 2006).

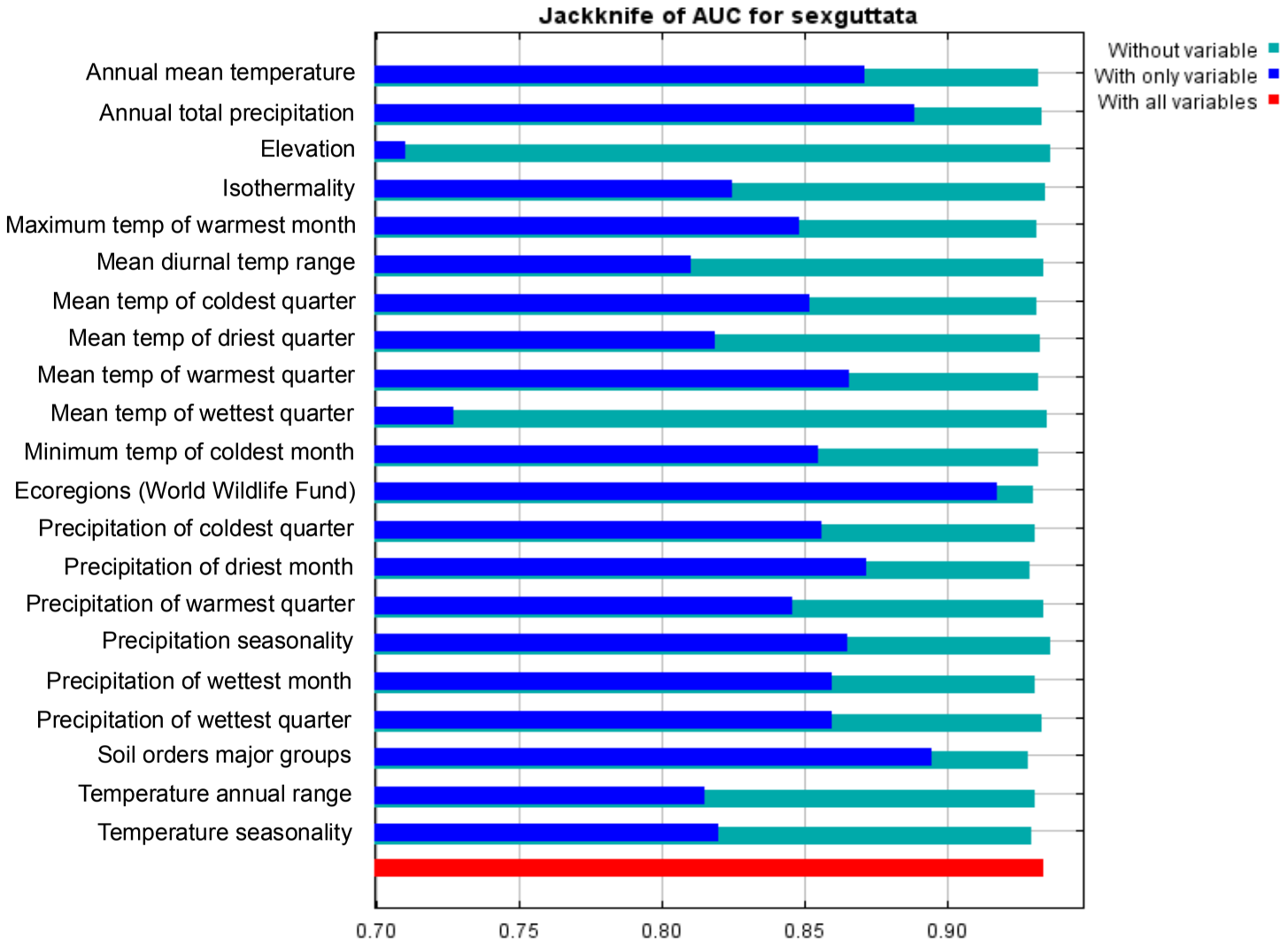

Jackknife of AUC for sexguttata
Without variable
With only variable
With all variables
Annual mean temperature
Annual total precipitation
Elevation
Isothermality
Maximum temp of warmest month
Mean diurnal temp range
Mean temp of coldest quarter
Mean temp of driest quarter
Mean temp of warmest quarter
Mean temp of wettest quarter
Minimum temp of coldest month
Ecoregions (World Wildlife Fund)
Precipitation of coldest quarter
Precipitation of driest month
Precipitation of warmest quarter
Precipitation seasonality
Precipitation of wettest month
Precipitation of wettest quarter
Soil orders major groups
Temperature annual range
Temperature seasonality
0.70
0.75
0.80
0.85
0.90

Figure 4.4. Jackknife analysis of variable contribution to *C. patruela* ecological niche model. Environmental variables were selected because they are likely to limit species distributions (Leffler 1979; Spanton 1988; Hijmans *et al.* 2005a). The red bar represents the test AUC (area under the receiver operating characteristic), a measure of the model predictive performance (Phillips *et al.* 2006). Dark blue bars indicate how well each variable predicts the species niche when used alone. Light blue bars show model performance when the variable is removed. Because environmental variables are often correlated, removal of a single variable often has little effect on the total model (Phillips *et al.* 2006).

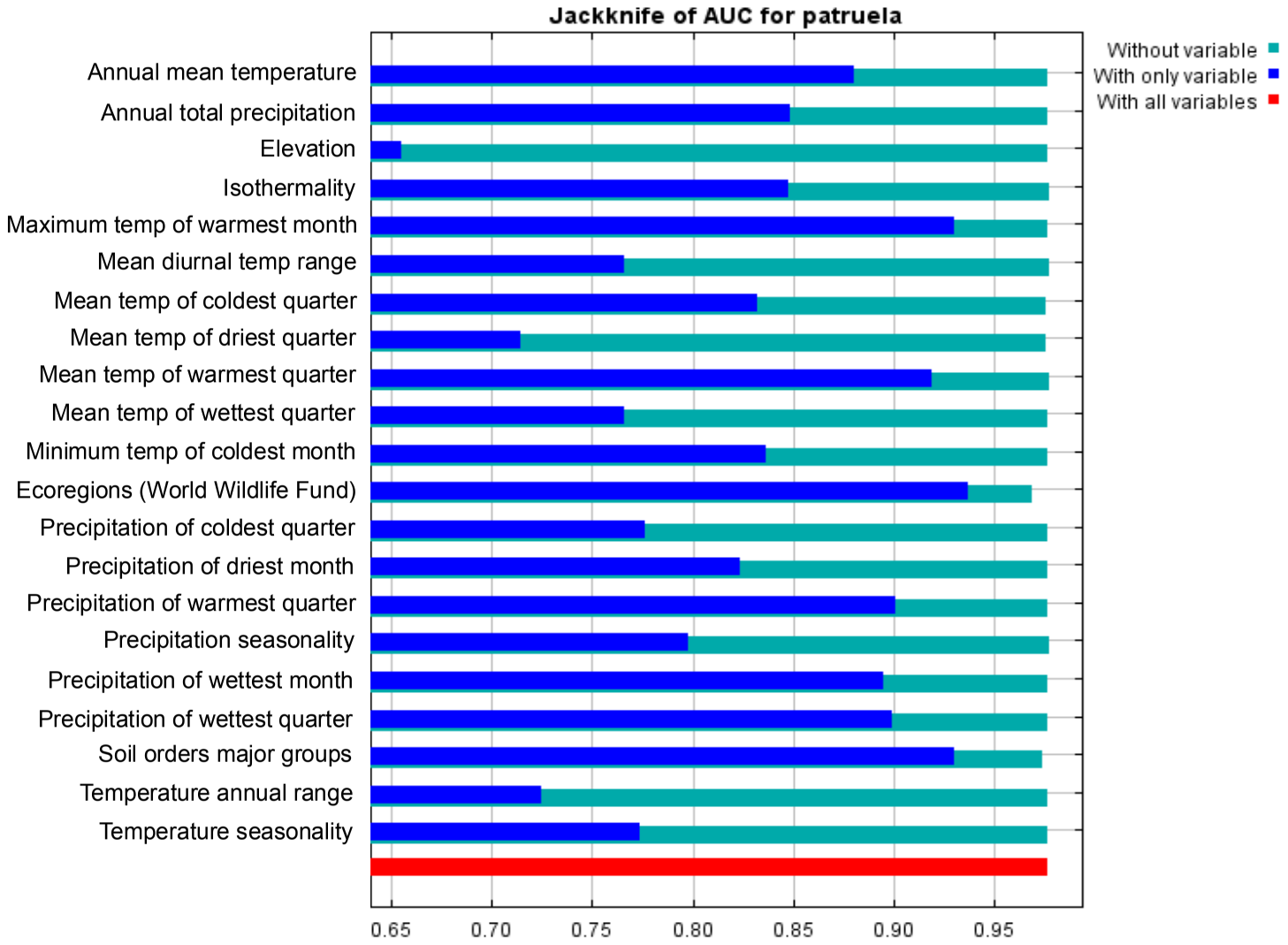

Jackknife of AUC for patruela
Annual mean temperature
Annual total precipitation
Elevation
Isothermality
Maximum temp of warmest month
Mean diurnal temp range
Mean temp of coldest quarter
Mean temp of driest quarter
Mean temp of warmest quarter
Mean temp of wettest quarter
Minimum temp of coldest month
Ecoregions (World Wildlife Fund)
Precipitation of coldest quarter
Precipitation of driest month
Precipitation of warmest quarter
Precipitation seasonality
Precipitation of wettest month
Precipitation of wettest quarter
Soil orders major groups
Temperature annual range
Temperature seasonality
0.65
0.70
0.75
0.80
0.85
0.90
0.95
Without variable
With only variable
With all variables

Figure 4.5. Jackknife analysis of variable contribution to *C. denikei* ecological niche model. Environmental variables were selected because they are likely to limit species distributions (Leffler 1979; Spanton 1988; Hijmans *et al.* 2005a). The red bar represents the test AUC (area under the receiver operating characteristic), a measure of the model predictive performance (Phillips *et al.* 2006). Dark blue bars indicate how well each variable predicts the species niche when used alone. Light blue bars show model performance when the variable is removed. Because environmental variables are often correlated, removal of a single variable often has little effect on the total model (Phillips *et al.* 2006).

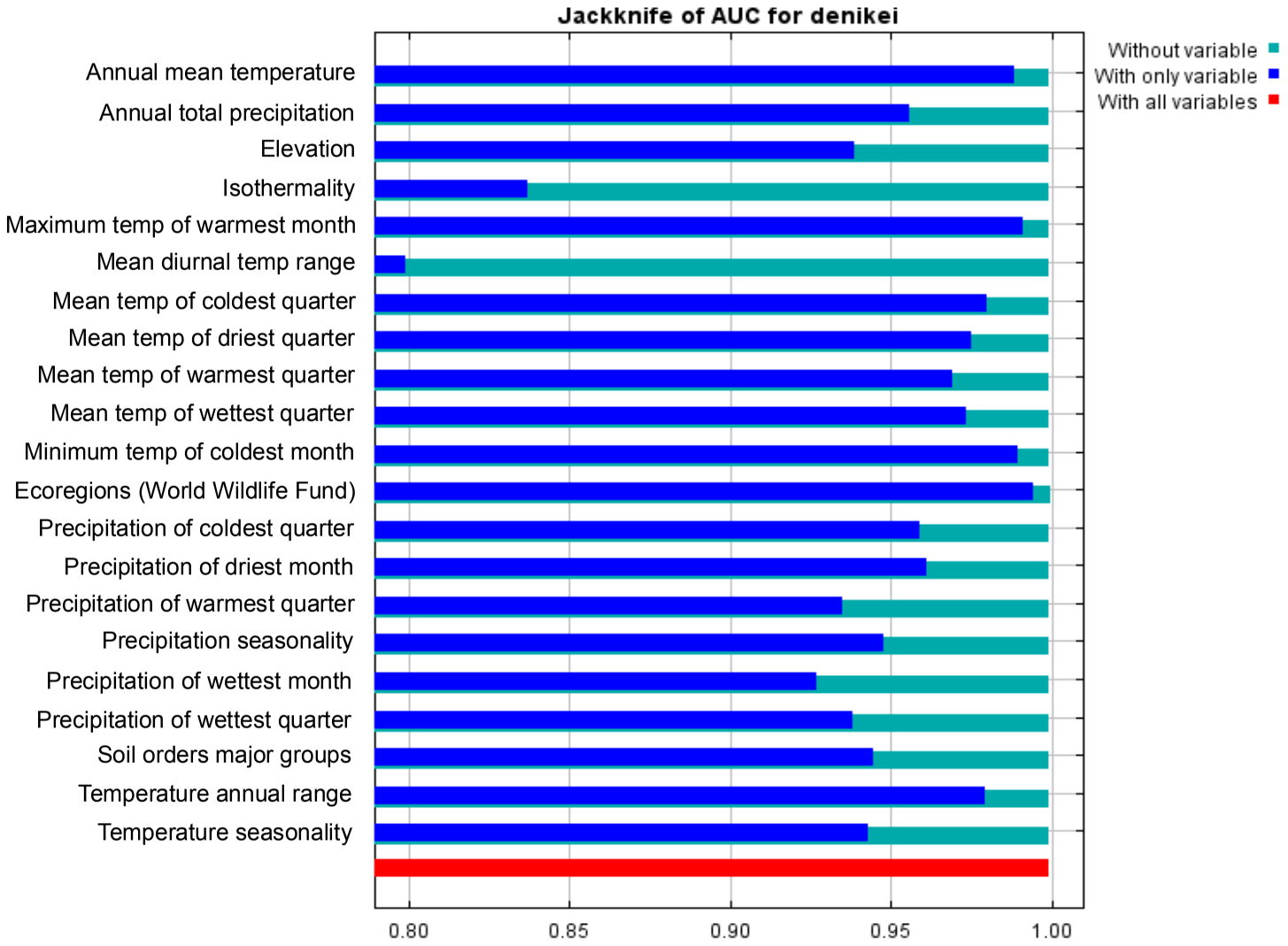

Jackknife of AUC for denikei
Annual mean temperature
Annual total precipitation
Elevation
Isothermality
Maximum temp of warmest month
Mean diurnal temp range
Mean temp of coldest quarter
Mean temp of driest quarter
Mean temp of warmest quarter
Mean temp of wettest quarter
Minimum temp of coldest month
Ecoregions (World Wildlife Fund)
Precipitation of coldest quarter
Precipitation of driest month
Precipitation of warmest quarter
Precipitation seasonality
Precipitation of wettest month
Precipitation of wettest quarter
Soil orders major groups
Temperature annual range
Temperature seasonality
0.80
0.85
0.90
0.95
1.00
Without variable
With only variable
With all variables

Figure 4.6. *Cicindela longilabris* ecological niche model. Black squares represent known species occurrences. Green areas indicate the predicted species distribution by MAXENT. Lighter green: predicted using the Lowest Presence Threshold (LPT). Under the LPT, the model includes all a training points (no omission errors). Darker green: predicted with more conservative M10 threshold,10% of points are considered outliers and omitted from the model. Additional areas in Yukon and Alaska are expected to be suitable habitat (Spanton 1988), but may contain environmental conditions outside the training model for the species. This is likely the result of collector sampling bias towards more accessible areas in the rest of the continent. The marked difference in predicted habitat between the two thresholds indicates the presence of outlier points in suboptimal habitat, or niche divergence in portions of the species range. Habitats vary across the range, but *C. longilabris* is typically associated with openings in boreal forests, and alpine meadows that are often at high elevation (Spanton 1988)

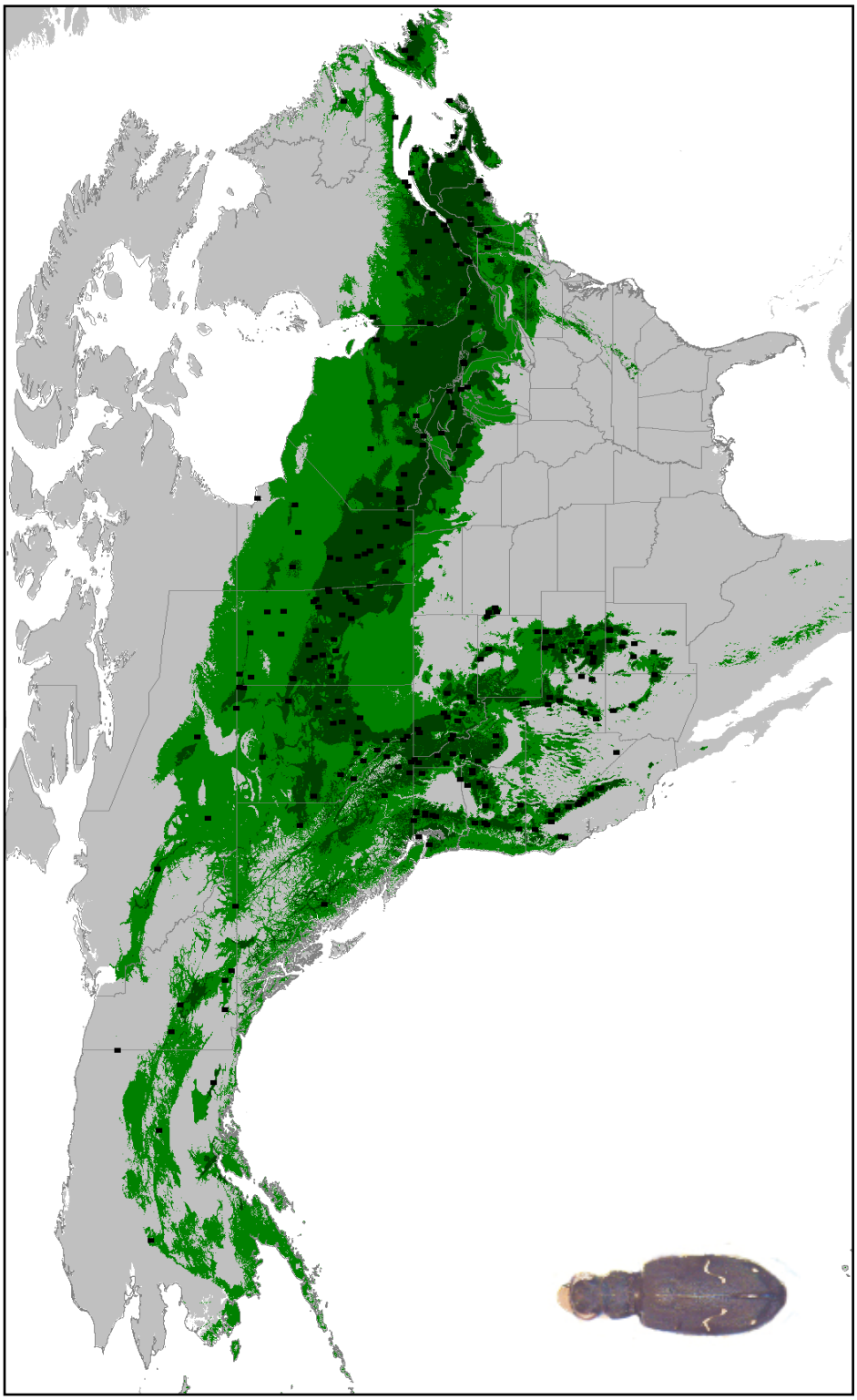


Figure 4.7. *Cicindela nebraskana* ecological niche model. Black squares represent known species occurrences. Green areas indicate the predicted species distribution by MAXENT. Lighter green: predicted using the Lowest Presence Threshold (LPT). Under the LPT, the model includes all a training points (no omission errors). Darker green: predicted with more conservative M10 threshold,10% of points are considered outliers and omitted from the model. High degree of concordance between thresholds indicates that outliers have little effect on the model. *C. nebraskana* is typically associated with bare clay openings in grasslands (Spanton 1988).

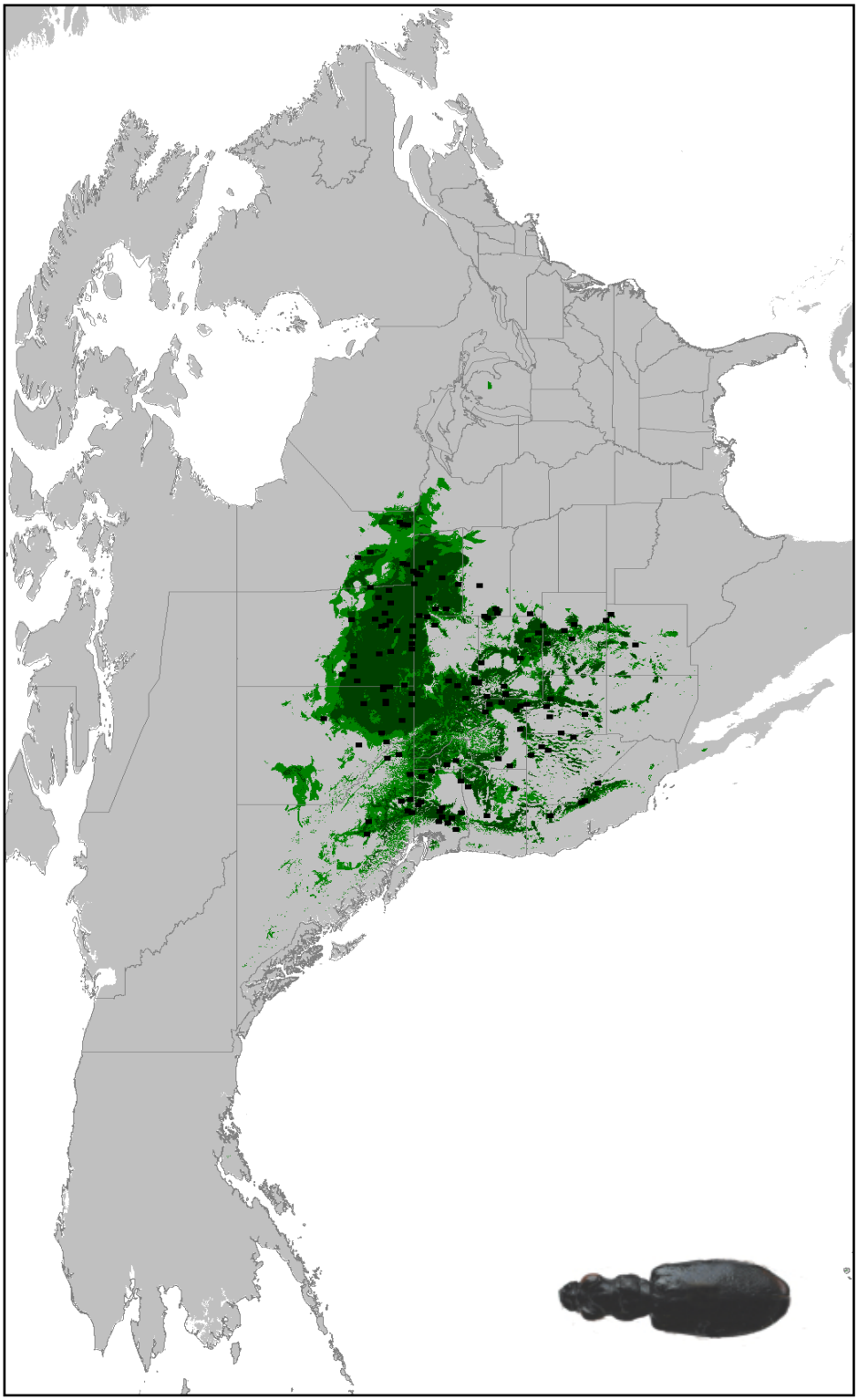


Figure 4.8. *Cicindela sexguttata* ecological niche model. Black squares represent known species occurrences. Green areas indicate the predicted species distribution by MAXENT. Lighter green: predicted using the Lowest Presence Threshold (LPT). Under the LPT, the model includes all a training points (no omission errors). Darker green: predicted with more conservative M10 threshold,10% of points are considered outliers and omitted from the model. The thresholds differ most in predicting the species west of the Mississippi River. *C. sexguttata* is typically associated with deciduous or mixed forest edges (Kaulbars and Freitag 1993b) and these habitats appear to be patchier in the Midwestern portion of their range.

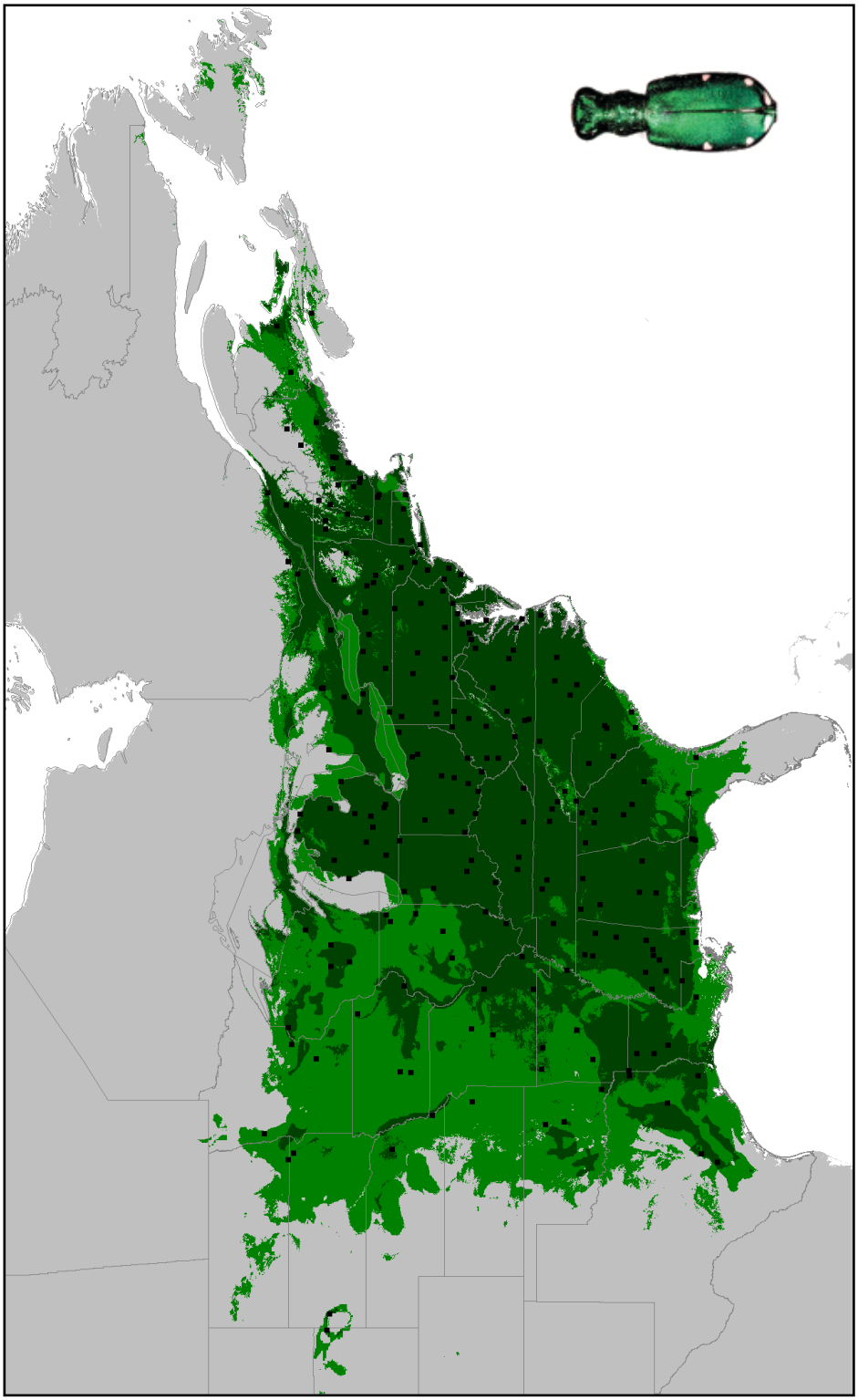


Figure 4.9. *Cicindela patruela* ecological niche model. Black squares represent known species occurrences. Green areas indicate the predicted species distribution by MAXENT. Lighter green: predicted using the Lowest Presence Threshold (LPT). Under the LPT, the model includes all a training points (no omission errors). Darker green: predicted with more conservative M10 threshold,10% of points are considered outliers and omitted from the model. *C. patruela* is associated with "barrens" habitat, areas of poor soil, often sand, but also shale, talus, and other substrates (Knisley *et al.* 1990, Kaulbars and Freitag 1993b).

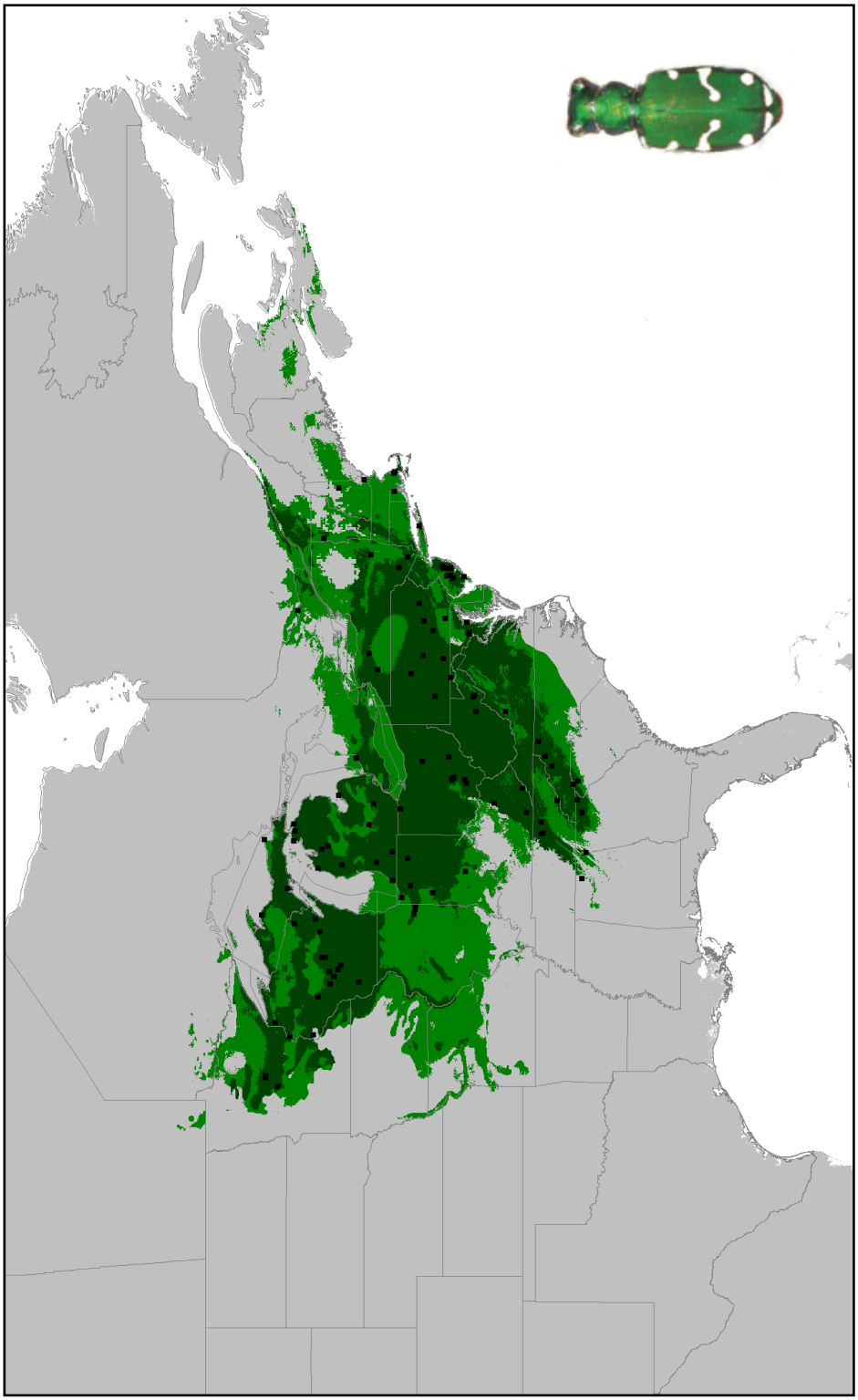
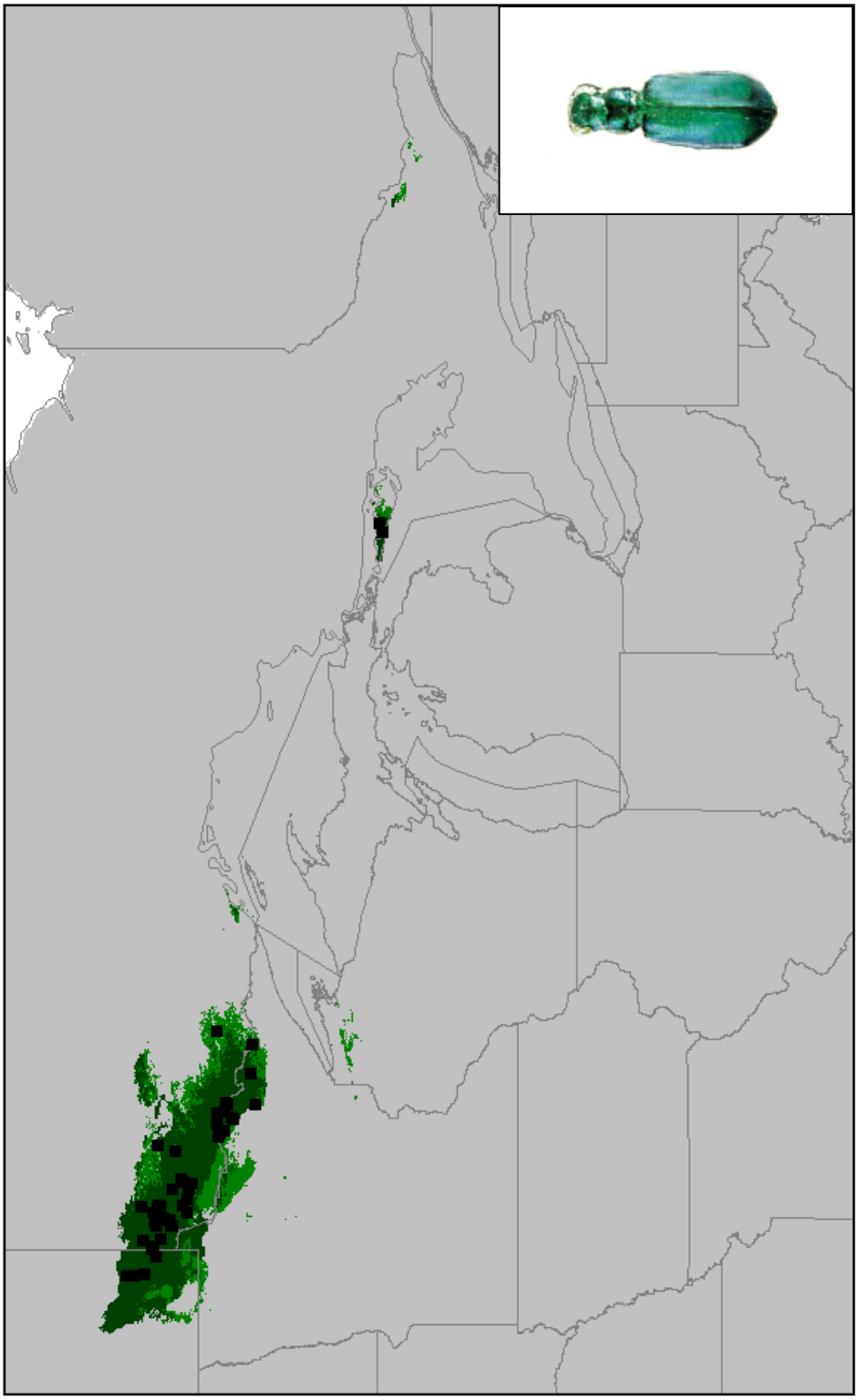


Figure 4.10. *Cicindela denikei* ecological niche model. Black squares represent known species occurrences. Green areas indicate the predicted species distribution by MAXENT. Lighter green: predicted using the Lowest Presence Threshold (LPT). Under the LPT, the model includes all a training points (no omission errors). Darker green: predicted with more conservative M10 threshold,10% of points are considered outliers and omitted from the model. Little addition habitat is predicted outside of known occurrence points. This could be reflective of the fact that no additional suitable habitat exists, or alternatively due to niche model limitations using species with very small geographic distributions (Pearson *et al.* 2007). *C. denikei* is associated with alvar habitat (Kaulbars and Freitag 1993a, b), areas of bare rock with little to no soil, exposed from glacial retreat.



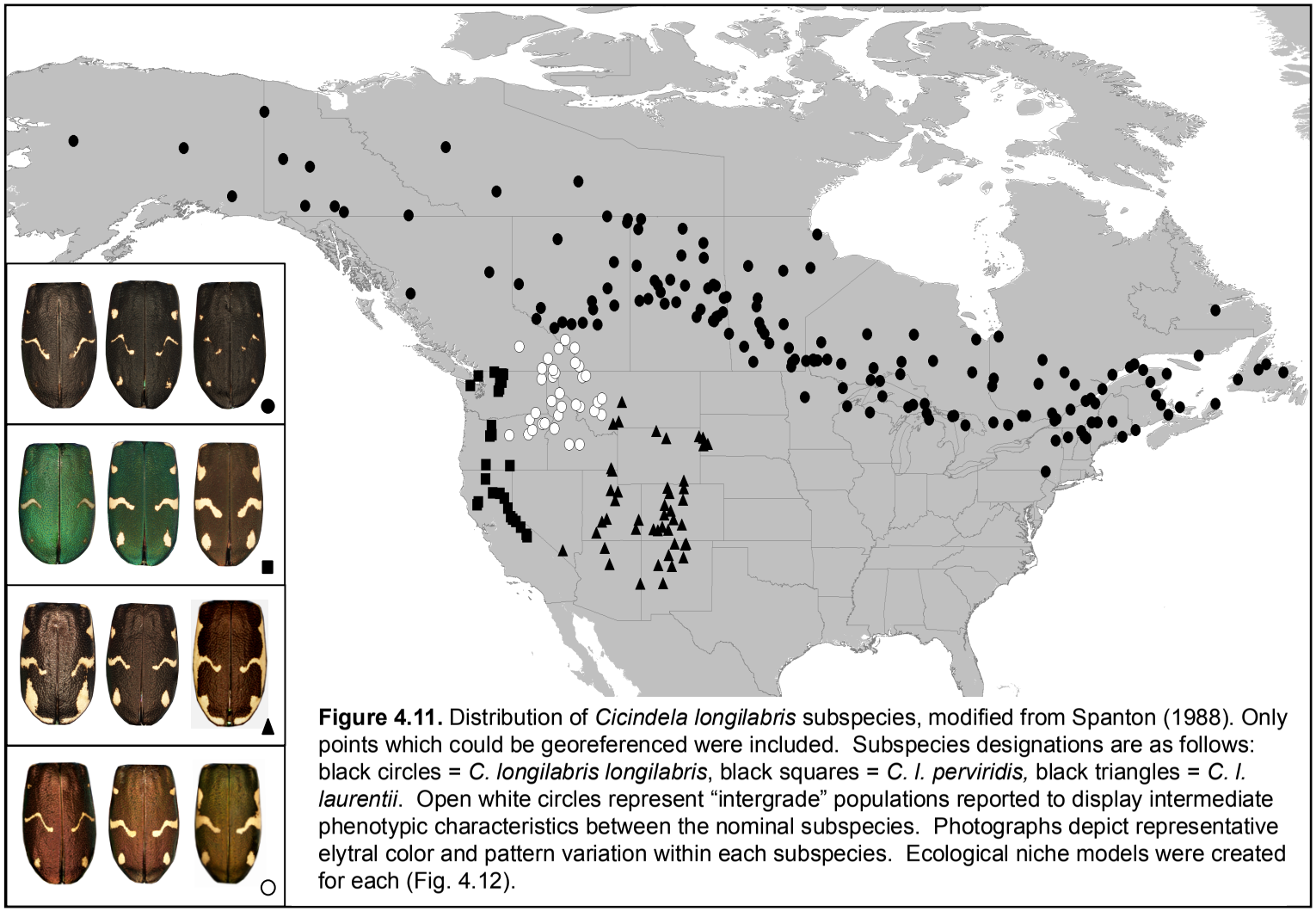


Figure 4.11. Distribution of *Cicindela longilabris* subspecies, modified from Spanton (1988). Only points which could be georeferenced were included. Subspecies designations are as follows: black circles = *C. longilabris longilabris*, black squares = *C. l. perviridis,* black triangles = *C. l. laurentii*. Open white circles represent "intergrade" populations reported to display intermediate phenotypic characteristics between the nominal subspecies. Photographs depict representative elytral color and pattern variation within each subspecies. Ecological niche models were created for each (Fig. 4.12).

Figure 4.12. Ecological niche models for *Cicindela longilabris* subspecies. **A)** *C. longilabris longilabris*, **B)** *C. l. perviridis*, **C)** *C. l. laurentii*. Black circles, squares, and triangles represent known occurrences respectively. Green areas indicate the predicted species distribution by MAXENT. Lighter green: predicted using the Lowest Presence Threshold (LPT). Under the LPT, the model includes all a training points (no omission errors). Darker green: predicted with more conservative M10 threshold,10% of points are considered outliers and omitted from the model. Note that each model cross-predicts a portion of the other subspecies ranges. The area circled in black represents the geographical distribution of "intergrade" populations that display intermediate phenotypic characteristics between the nominal subspecies. This area was strongly predicted by all subspecies ENMs. However there are additional adjacent areas of predicted overlap that do not contain "intergrades", and molecular data were inconsistent with subspecies boundaries (compare to Fig 2.3). Statistical tests performed with ENMTools (Warren *et al.* 2010) demonstrated that *C. l. longilabris* was significantly different in niche from the other subspecies. The subspecies*, C. l. laurentii* and *C. l. perviridis* were ecologically interchangeable and not significantly more different in niche than expected by chance. Regardless of the taxonomic validity of the nominal subspecies, the results indicate that there is niche divergence in different parts of the range.

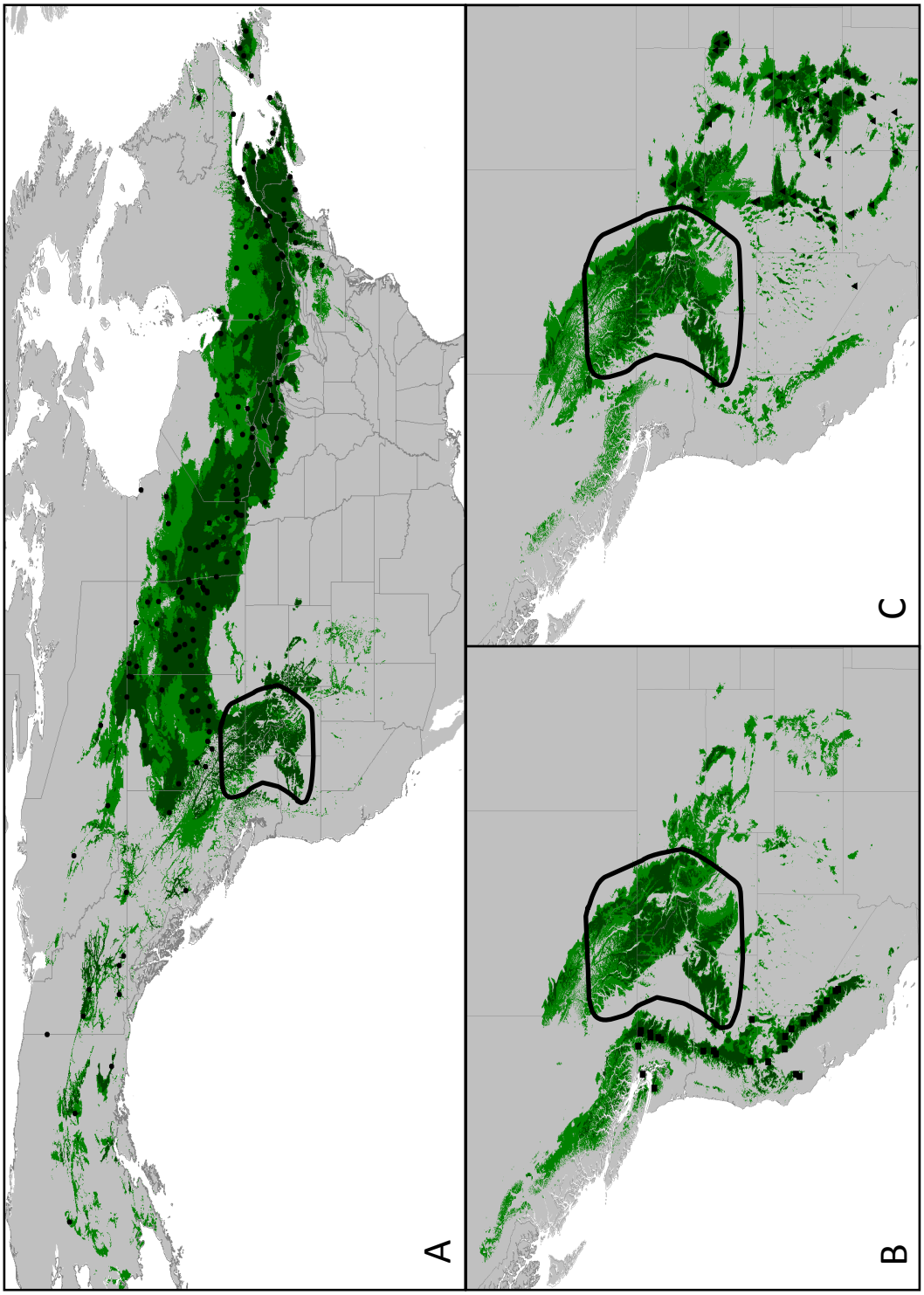

A
B
C

Figure 4.13. Ecological niche model prediction for Meadow Group mtDNA Clades (see Fig. 2.3). **A)** Continental Clade, **B)** Northwest Clade, **C)** Southwest Clade, **D)** Close-up of mtDNA clade contact zone in the northern Sierra Nevada range. Blue and yellow circles represent Northwest and Southwest Clade populations respectively. Green areas are those predicted as habitat for both clades. Note that contiguous suitable habitat is predicted for each clade throughout the Sierra Nevada range and adjacent Great Basin highlands. This pattern coupled with the abrupt contact zone supports an extrinsic biological cause for range limits, such as competitive exclusion (Swenson 2006, Sanchez-Cordero *et al.* 2008) or the effect of *Wolbachia* infections. If mtDNA clades are associated with different *Wolbachia* strains, then cytoplasmic incompatibility (Hoffman and Turelli 1997) could prevent survival of migrants across the contact zones

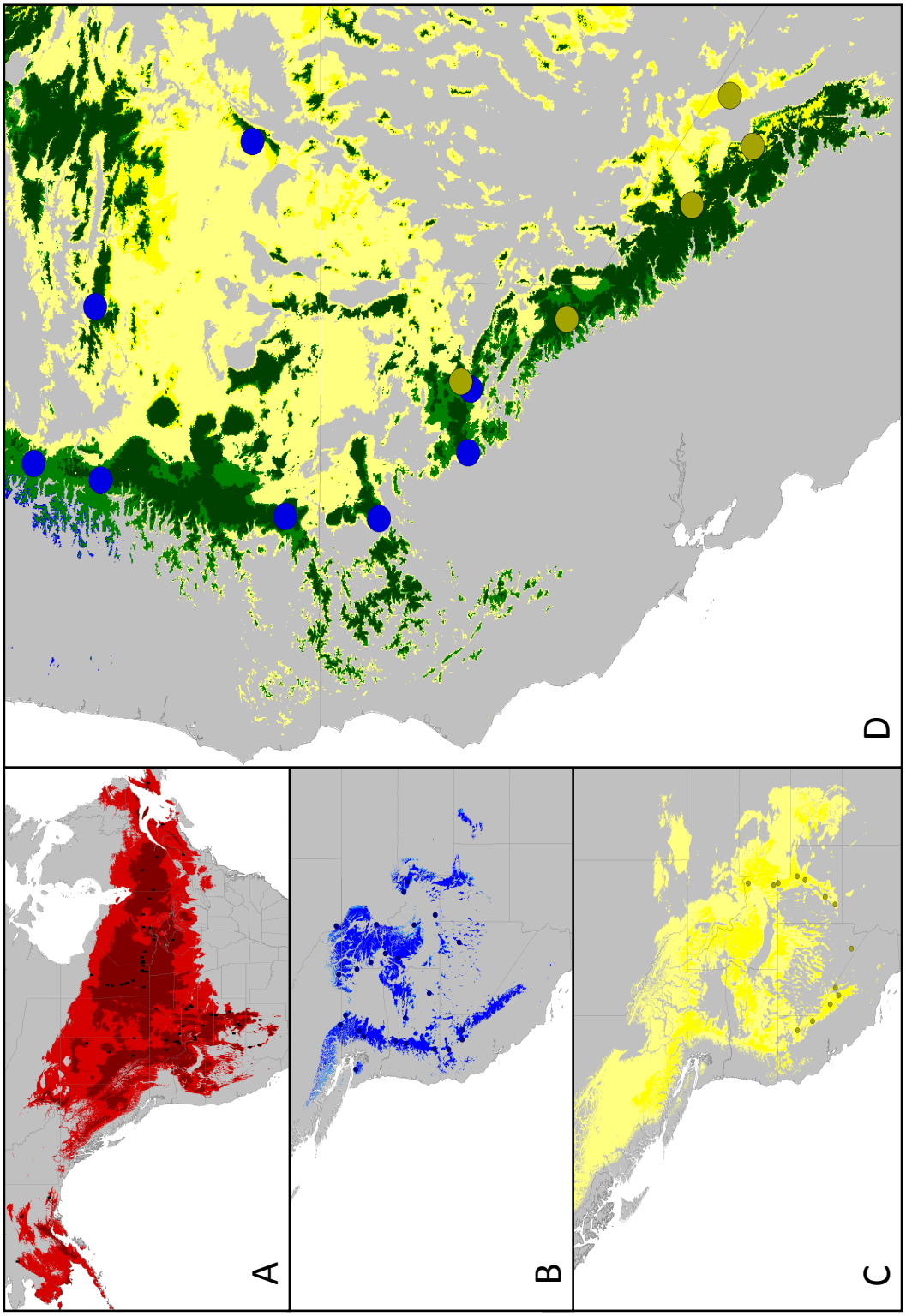

A
B
C
D

Figure 4.14. A) Meadow Group ecological niche model. This model treats *C. longilabris* and *C. nebraskana* as a single species (see Chapter II Discussion). Green areas indicate the predicted species distribution by MAXENT. Lighter green: predicted using the Lowest Presence Threshold (LPT). Under the LPT, the model includes all a training points (no omission errors). Darker green: predicted with more conservative M10 threshold, 10% of points are considered outliers and omitted from the model. **B,C)** Ecological niche model prediction for Meadow Group distribution during the Last Glacial Maximum (~18,000 years ago). Panel C shows greater detail of the focal area. The CCSM and MIROC models (Collins *et al.* 2004, Hasumi and Emori 2004) were used to generate the LGM model predictions, shown in green. Darker colors indicate higher probability of occurrence. Lightest green: predicted as suitable under the LPT by either of the models. Light green: predicted under LPT from both models. Green: predicted under M10 threshold by one of the models. Darkest green: predicted under M10 threshold by both models. Black circles indicate the approximate location of hypothesized glacial refugia based on phylogeography and mismatch distribution analyses (Chapters II, III). Highly predicted areas occur in each refugium and post-glacial dispersal routes exist as each are broadly connected to suitable present-day habitat. Blue-grey hatching indicates the approximate location of ice sheets during the LGM, and the present-day coastlines are outlined in darker grey.

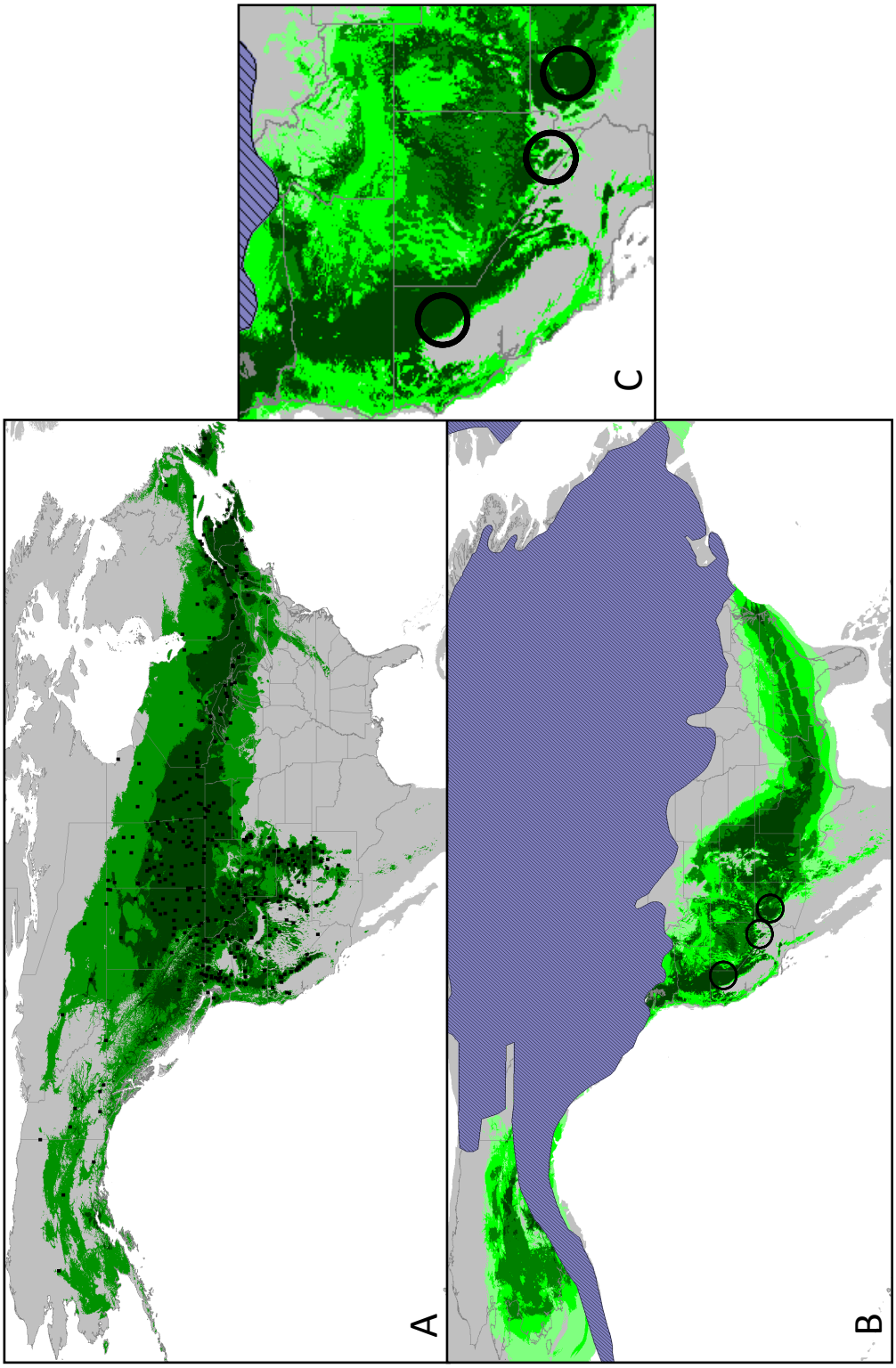
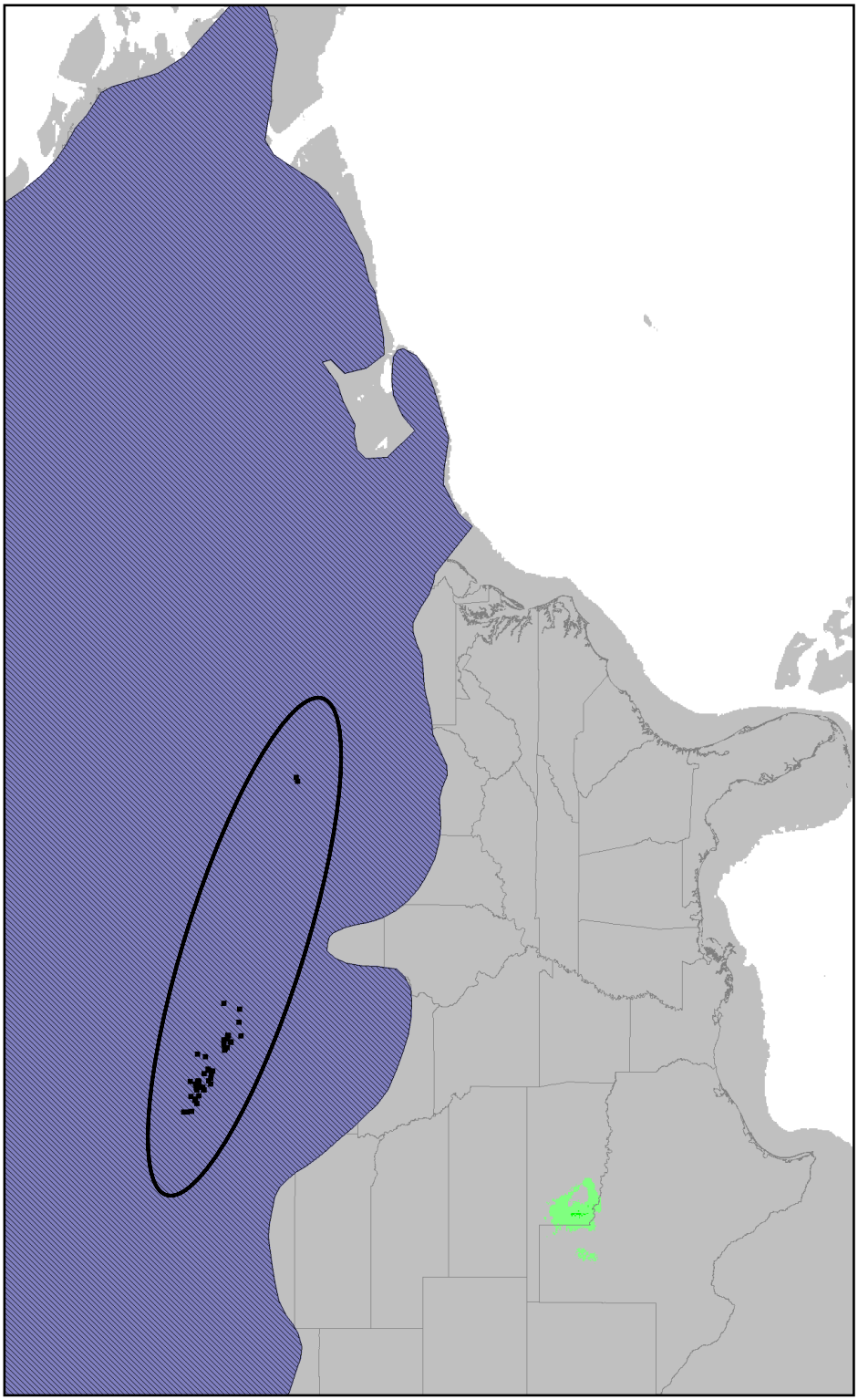

C
A
B

Figure 4.15. Ecological niche model prediction for *C. denikei* during the Last Glacial Maximum (LGM), ~18,000 years ago. The CCSM and MIROC models (Collins *et al.* 2004, Hasumi and Emori 2004) were used to generate the LGM model predictions, shown in green. Darker colors indicate higher probability of occurrence. The lightest green indicates the suitable habitat predicted by the CCSM model using the lowest threshold, LPT. The light green indicates the area predicted by the more stringent M10 threshold under the CCSM model. No areas are predicted as suitable by any threshold under the MIROC model. The weak support for the existence of suitable habitat during the LGM is consistent with the hypothesis that if *C. denikei* is a distinct species, that it evolved *in situ* after glacial ice retreated and created the species current alvar habitat (Kaulbars and Freitag 1993a,b).



CHAPTER V

AN ASSESSMENT OF ENVIRONMENTAL INFLUENCES ON COLOR PATTERN IN *CICINDELA LONGILABRIS*

Introduction

One of the most pervasive questions in biology is the nature of the relationships between phenotypic variation, genetics, and the environment. Well before the use of molecular genetic data, naturalists had noted the existence of geographical patterns of phenotypic variation in plant and animal species. These geographic races or subspecies were commonly construed to be populations with reduced gene flow, or even incipient species that may become completely reproductively isolated in the future (*e.g.* Mayr 1942, 1963). In more recent years, it has become possible to test for concordance between morphologically recognized taxonomic subspecies and molecular data, as phylogeographic studies make use of neutrally evolving molecular genetic markers to uncover the patterns of population subdivision (*e.g.* Avise *et al.* 1987; Avise 2000; Zink 2004). One of the surprising outcomes of many such studies is the identification of deep genealogical splits which do not correspond to any obvious phenotypic characters (*e.g.* Avise *et al* 1979; Vogler and DeSalle 1993, 1994; Bernatchez

1997). However, other workers have found higher levels of agreement between molecules and subspecies boundaries (Phillimore and Owens 2006). Moreover, some species also exhibit phenotypic variation within populations, and this could be a result of neutral allelic variation (*e.g.* Tan 1945), developmental plasticity (*e.g.* Nice and Fordyce 2006), or local adaptation (*e.g.* Dudley 1996). So the question remains as to whether the geographic distribution of polymorphism is largely the result of historical sundering of species, local adaptation, environmental developmental plasticity, or some combination of these. More thorough and integrative studies are needed to assess the general trends concerning the causes of phenotypic variation.

Tiger beetles in the genus *Cicindela* often display striking variation in their color patterns within and among species (Shelford 1917; Willis 1967; Acorn 1992). These can vary both in ground color and in the extent and shapes of dorsal unpigmented areas (maculations). Some species of *Cicindela* exhibit considerable variation in dorsal color across their geographic range or within a single population and this may include a continuum of color forms (e.g. populations of *C. circumpicta johnsoni* ranging from blue to red), or in some cases two or more distinct color forms (e.g. *C. scutellaris rugifrons* black and green morphs). All dorsal colors in tiger beetles are “structural colors” created by the optical properties of the cuticle, and not the result of different pigments (Schultz and Rankin 1983a, b). The cuticle is laminated with alternating layers of melanin pigment and translucent epicuticle, and the distance between these layers largely determines the color reflected. In addition to the wavelengths

reflected through these layers, the surface microsculpturing can also create a mosaic of microscopic patches that reflect different wavelengths, which pointillistically mix to create many of the brown and olive colors in *Cicindela* (Schultz and Bernard 1989). Some studies have demonstrated that color may be important in predator avoidance as a result of crypsis by resemblance to the substrates they frequent (Willis 1967; Schultz 1986, 1991). Other studies have demonstrated that extent of maculation in *Cicindela* can have significant adaptive consequences for thermoregulation by facilitating heat transfer through the integument (Schulz and Hadley 1987; Acorn 1992; Hadley *et al.* 1992). Consequently, the potential exists for natural selection to be operating on different phenotypes within populations of *C. longilabris.* However, the structure of the cuticular layers and the resulting colors have been found in some cases to at least partly reflect phenotypic plasticity in development as a consequence of such abiotic factors as varying levels of humidity or temperature (Shelford 1917; Schultz 1983). Shelford (1917) reared the larvae of five species of *Cicindela* under two different temperature regimes (21^{o}C and 35^{o}C) while also subjecting each of these to low and high humidity treatments. He found differences in color associated with treatments; hotter and drier conditions resulted in brighter and more green/blue (as opposed to reddish) colors. He concluded that local climate was therefore responsible for much of the geographical color variation observed in tiger beetles, although his interpretations were criticized by Schultz (1983) as overgeneralizations. Schultz repeated many of Shelford's observations but also examined the developmental physiology of the exoskeleton, and determined that

hot and dry conditions caused shrinking of the laminated layers of melanin in the cuticle, resulting in shorter wavelength colors (*e.g.* shifted towards the blue/violet end of the spectrum). Because these experiments were carried out on only a handful of species under a small set of possible conditions, it has not been possible to say how representative these patterns are. Despite these issues, the taxonomy of *Cicindela* subspecies (and occasionally species) is often based entirely on subtle differences in color and pattern (*e.g.* Willis 1968; Graves *et al.* 1988; Schincariol and Freitag 1991; Kritsky and Horner 1998), characters that are either potentially under strong selection, or possibly the result of developmental plasticity. In both cases, these characters may be inappropriate for inferences about systematic relationships as such traits are prone to patterns of convergence and can result in 'polytopic subspecies' (Wilson and Brown 1953) that are not representative of evolutionarily meaningful entities. Yet multiple tiger beetle subspecies have been included in state and federal endangered species lists (NatureServe. 2009, accessed June 2010). An improved understanding of the underlying causes for phenotypic variation would be valuable for accurate taxonomy in this popular and conservationally important group of insects.

Within the tiger beetles, *C. longilabris* is ideal for studying the nature of polymorphism and underlying causes of phenotypic variation. This species exhibits an extreme amount of phenotypic variation in color pattern across its continent-wide range, which has led to the description of numerous subspecies (reviewed in Spanton 1988), some of which were given full-species rank by previous authors. Spanton revised the taxonomy, synonymizing many of the

previous names, but recognizing three subspecies based on linear discriminant function analyses of morphometric data (Sneath and Sokal 1973). Although there appears to be a geographical component to the phenotypic variation, there also exists considerable variation within each of these subspecies (Figure 5.1). Additionally, there is often dramatic color pattern variation within populations in a number of geographical areas (*i.e.* ID, MT, AZ, WA), which may be a result of elevated variation in heritable traits, developmental plasticity in particularly variable environments, or an interaction of these factors.

Part of my dissertation research was focused on the systematics of the North American *C. sylvatica* group (Chapter II), including *C. longilabris*. Utilizing a thorough 'congeneric phylogeography' approach I sampled intensively from all *C. longilabris* subspecies and intergrades. Although mtDNA sequence data and AFLP clustering both revealed the existence of deep genetic subdivisions in the group, these were not concordant with any subspecies boundaries. Following the results of Shelford (1917) and Schulz (1983), I sought to investigate the degree to which environment might explain variation in color and maculation. To this aim, I photographed dried beetle specimens from throughout the range of *C. longilabris*, using high-resolution photographic equipment, and quantified color and maculation with photo-processing software. I performed multiple regression analyses of key environmental variables against these measures of phenotype and interpreted the resulting patterns below.

Materials and Methods

Sampling

Populations were chosen to represent each of the subspecies and intergrades, broadly encompassing the range of each group. The dataset includes 151 individuals and 31 sampling sites (Figure 5.2), and each specimen was the morphological voucher of an individual used in prior molecular genetic analyses. Beetles were dried for at least 15 mins on pieces of filter paper that had been placed in individually labeled petri dishes. Dried beetles were then placed on a microscope stage containing an indented piece of modeling clay that held each beetle in place and in the same plane and orientation. I photographed each specimen using a Nikon DF-Fi1 camera, NI-150 light source, and SMZ 1500 microscope. In order to ensure that colors were consistently captured, each photograph was taken using the same settings for lighting and exposure: 50% power on the light source, and 300 ms exposure, Gain 1.00, medium contrast for the camera. In addition, five color standard strips were included in each photograph, corresponding to white, red, blue, yellow, and green. Photographs were saved as uncompressed TIFF files (2560 x 1920 resolution) and imported to Photoshop CS5.

Phenotype scoring

Color was scored using two separate systems for comparison. First, each color was scored according to the ordered color states used in Spanton (1988). This scale ranged from 1-7, as follows: 1) black, 2) dark brown, 3) medium brown-bronze, 4) olive green, 5) green, 6) blue green, 7) dark blue. Although this method necessarily simplifies color into seven binned categories, there were two justifications for its inclusion. First, I have observed that colors in *C. longilabris* (and some other species) do follow this progression of ordered states, and populations will often include two or more of these states and they are always consecutive (*i.e.* CA: Duck Lake has states 5-7, ID: Snowhaven has 2-3, CO: Poudre Canyon has states 1-3, *etc*). Moreover, an individual specimen may shift its color towards the higher number end of the continuum through rapid drying under a heat source (*pers. obs*.; Schimdt, Lawton, and Kippenhan, *pers. comm*., 2005). Second, this scale was the same one used in Spanton's linear discriminant function analyses, and therefore my analyses would be performed using the same metric for color comparison. In addition to this scale, color was also measured using the Hue, Saturation, Brightness scale in Photoshop CS5. The HSB scale was chosen over Red, Green, Blue (RGB), because hue is a single chromatic value that has been shown to capture color variation effectively (McKenna *et al.* 1999; Robertson and Zamudio 2009). Color saturation was scored as a phenotype and may be interpreted as the "pureness" of color based on the Munsell (1914) system. Saturation represents the degree to which a color

differs from a neutral grey. Brightness was also scored and represents the light "intensity" of color. To record HSB values, I used the color picker tool in Photoshop to take a 101x101 pixel average from the same part of the elytra disk for each beetle (1/3 from anterior end, near the elytral suture). Maculations were scored using the "lasso" selection tool. First, each elytron was traced and the number of pixels recorded. Next, the elytral maculations were outlined and combined and the total number of pixels was recorded in an adjacent column. Percent maculation was calculated as the pixel area of the maculations divided by the total elytral area.

Selection of environmental variables

To investigate the effect of environment on phenotype, I chose environmental variables that met two criteria: 1) having sufficient resolution of the data so that values precisely represented the conditions encountered for each sampling locality, 2) representing some *a priori* expectation of association with phenotype based on previous detailed studies by Shelford (1917) and Schultz (1983b). The only environmental layers with sufficient resolution across the entire North American continent were the BIOCLIM variables by Hijmans (2005a) and these were also used in the ecological niche modeling part of my dissertation (see Chapter IV Materials and Methods). From this total set, I selected the variables that pertained to temperature, precipitation (as a potential surrogate for humidity), and elevation. The program DIVA-GIS (Hijmans *et al.*

2005b) was used to extract point values for each layer at the latitude and longitude coordinates for the 31 sampling localities.

Statistical analyses

Multiple regression analyses were conducted using JMP 8.0 (SAS Corporation, Cary, NC). Dependent phenotype variables (Ys) were regressed on independent environmental variables (Xs) using a Standard Least Squares model. Because the environmental variables chosen are likely to show some degree of covariance (Hijmans *et al* 2005a), I generated a matrix of variable correlation (Table 5.1). Normality of the data was tested by plotting residuals for each Y, in order to assure that the data did not violate assumptions of the model. Statistical testing of normality was performed with a Shapiro-Wilk Goodness of Fit test. Leverage plots were created to visually inspect for the presence of outlier points that may be affecting the models. Linearity and homoscedasticity were assessed by examining the shape of plots of the residuals versus the predicted values. For each significant regression, t ratios were generated to assess the statistical significance of specific variables towards the total model. In addition, standardized (beta) coefficients were generated in order to compare directly the magnitude of the effects of each independent variable on the model.

Results

Shapiro-Wilk tests were non-significant ($p > 0.05$) and the null hypothesis that data were normally distributed could be accepted in all cases. Plots of residuals versus the predicted values revealed evenly distributed points with no U-shaped tendencies (no violation of linearity or homoscedasticity). Therefore, no transformation was necessary for any of the phenotype data. Not surprisingly, there was a high degree of correlation between some environmental variables (*i.e.* annual mean temperature and mean temperature of the warmest month are 97.2% correlated) (Table 5.1), however this has no bearing on the overall significance of the results, only that it should be kept in mind when assessing the effect of particular variables in the model.

Multiple regression results indicated a highly significant ($p < 0.0001$) correlation between environmental variables and each measure of phenotype (Tables 5.2 – 5.6). Spanton's (1988) color scale displayed the highest coefficient of determination for the overall model ($R^2 = 0.751$), signifying that 75% of the variation in color can be explained by the environmental factors in the model. Hue was the second best explained by the model ($R^2 = 0.623$), followed by brightness ($R^2 = 0.530$), maculation percentage ($R^2 = 0.403$), and color saturation ($R^2 = 0.388$). Different sets of environmental variables were significantly predictive ($p \leq 0.05$) for each measure of phenotype. Mean temperature of the warmest quarter, and maximum temperature of the warmest month were

significantly predictive of color (Spanton's states and hue), but not brightness or maculation percentage. Elevation was not predictive of Spanton's color states or hue, but was significantly associated with saturation and brightness and nearly significantly associated with maculation ($p = 0.06$). Most environmental variables displayed more idiosyncratic patterns of association with the different measures of phenotype, and interpretations will be discussed below.

Discussion

Interestingly, the multiple regression model was more predictive of Spanton's color series ($R^2 = 0.751$) than hue ($R^2 = 0.623$). This result demonstrates that his scale was not arbitrary, but in fact an accurate representation of the continuum of color transformations within the species. Both measures of color were significantly predicted by environment, and the high coefficient of determination values demonstrate the predictive power of the models in explaining color ($R^2 = 0.751, 0.623$). T ratios are indicative of the significance of variables in the model, and for both chromatic measures, the three most statistically significant variables were related to temperature (Tables 5.2 – 5.3). Standarized coefficients showed that these same variables were among the most important in their effect on the model. It should be noted that these variables are also correlated with each other (Table 5.1). These results

are consistent with prior studies that demonstrated that temperature during pupation is at least partly responsible for color development in some tiger beetle species (Shelford 1917; Schultz 1983). Furthermore, the positive predictive association between temperature and color corroborates the specific findings of the above authors who found that higher temperatures during pupation resulted in shorter wavelength (*i.e.* blue-shifted) elytral colors. My statistical analyses suggest that this same phenomenon may be a partial explanation for color variation in *C. longilabris*. Unfortunately, I was unable to include a direct measure of humidity into the model, in order to compare to prior studies. Relative humidity can vary considerably over extremely small distances (personal observation), especially where streams or lakes are nearby, and no such layers were available at even $1km^2$ resolution throughout the continent. It is possible that precipitation measures may be correlated with relative humidity on this rough scale, although this is not known. Some measures of precipitation were significantly predictive of Spanton's colors (*e.g.* precipitation of the warmest quarter and annual total precipitation). However, these same factors were not significant predictors of hue, while other precipitation measures were (*e.g.* precipitation of the wettest quarter and wettest month). As such, it is difficult to make generalizations about the association between precipitation and color. Elevation was not found to be predictive of color using Spanton's scale or hue (p = 0.656, 0.662 respectively). Other variables were found in some cases to be significant (*e.g.* mean diurnal temperature range significantly predicted Spanton's

color) although biological interpretations of these additional associations are less obvious.

Color saturation and brightness were both found to be significantly predicted by the environmental models. Nonetheless, these two measures may or may not be comparable to Shelford's (1917) and Schultz's (1983) definitions of "dull" and "bright" colors. Schultz and Rankin (1983a, b) determined that microscultpuring of the elytral surface can cause diffraction of light that results in less saturated, duller colors. In addition, they found that more humid conditions during development resulted in an accumulation of epicuticular waxes that also created duller elytra. The *C. longilabris* multiple regression analyses indicate a significant correlation between elytral brightness and precipitation (all variables), whereas measures of mean and maximum temperature were not significant (Table 5.5). Interestingly, measures of temperature variance were highly significant (*e.g.* mean diurnal temperature range, temperature annual range) in predicting brightness. It is less clear what the biological significance of this last pattern may mean. Saturation was significantly predicted by elevation and mean temperature (annual mean, and the warmest quarter) (Figure 5.4), although again the biological interpretations are not apparent.

Maculation percentage was found to be significantly correlated with environmental variables when the whole model was examined ($R^2 = 0.403$, $p < 0.0001$) (Table 5.6), but no single environmental variable was significantly predictive of this phenotypic characteristic based on the t Ratio values. Elevation was nearly significant, as well as two measures of precipitation and the annual

mean temperature. Examination of leverage plots revealed that the four beetles from AZ: Coconino Co (North Rim area) were outliers; when removed both elevation and annual mean temperature became significantly predictive of maculation (p = 0.0041 and 0.0167 respectively). However, the standardized coefficient for elevation was low (0.5) compared to annual mean temperature (3.8) and precipitation of the wettest month (4.7). These two variables were nearly statistically significant predictors of maculation (p = 0.07 for both). Interestingly, no previous tiger beetle studies had demonstrated a connection between the extent of maculations and elevation, although Shelford (1917) had postulated that the association may exist. The findings suggest that further examination of this pattern is warranted, and larger datasets may be informative. If additional outlier populations are also identified then it may be possible to make additional inferences about the nature of these exceptions to the association between maculation and elevation, in order to generalize upon the underlying biological significance of the pattern. Interpretation of the correlation between annual mean temperature and maculation appears to be more straightforward. Multiple studies have shown that maculations are important in thermoregulation (Schulz and Hadley 1987; Acorn 1992; Hadley *et al.* 1992), and therefore localized natural selection could result in different degrees of maculation as a response to the different temperature regimes these populations face. Maculations could also be affected by plasticity in development, but this pattern was not observed in Schultz's (1983) experiments, and only weakly observed in Shelford's (1917).

Additional evolutionary ecology considerations

Although it is not possible at the present to make conclusions on the underlying genetic basis for each of these phenotypic traits, these results do suggest that color may be at least partly due to the effects of developmental plasticity, consistent with earlier tiger beetle research. Moreover, the adaptive significance of color in this species is somewhat suspect. If visual predators were consistently exerting strong directional selection on color then it would be difficult to explain the presence of both brightly colored (green – blue) and cryptic (dark brown - bronze) populations throughout each of the subspecies and intergrades. Furthermore, some populations exhibit a range of colors (*e.g.* MT: Stevensville area, Figure 5.1 bottom row, #5-8) making the argument for cryptic coloration in the face of visual predators even more improbable. Although other workers have hypothesized that crypsis can be important in anti-predator defenses (Willis 1967; Schultz 1986, 1991), those species examined were found in more open habitats at low elevation. It may be that those cryptic species simply encounter more frequent and consistent selection from predators, compared to *C. longilabris*. In light of these questions, it appears that additional studies are necessary to fully characterize the possible adaptive significance of phenotypic characters and the genetic architecture underlying them.

Acknowledgments

I would like to thank Jonas King and Julian Hillyer (Vanderbilt University) for allowing me to use their photographic equipment and for providing excellent technical advice. Jonas King and Dan Funk assisted with the photographing of specimens used in the analyses. Jason Schmidt, Todd Lawton, and Mike Kippenhan all discussed aspects of tiger beetle color plasticity with me. Mike Kippenhan discovered the "intergrade" population of *C. longilabris* from Mt. Pisgah, OR and generously provided specimens. Rodger Gwiazdowski and Scott Egan provided feedback on parts of this manuscript. This work was funded by a Vanderbilt University Discovery Grant to Dan Funk and by a Theodore Roosevelt Memorial Grant from the American Museum of Natural History.

References

Acorn, J. A. 1992. The historical development of geographic color variation among dune *Cicindela* in western Canada. In The Biogeography of Ground Beetles of Mountains and Islands. Noonan, G. E., G. E. Ball, and N. E. Stork (Eds.). Intercept Press, Andover UK.

Avise, J.C., Giblin-Davidson, C., Laerm, J., Patton, J.C. and Lansman, R.A. 1979. Mitochondrial DNA clones and matriarchal phylogeny within and among geographic populations of the pocket gopher, Geomys pinetis. PNAS 76: 6694-6698.

Avise, J.C. 1987. Identification and interpretation of mitochondrial DNA stocks in marine species. In H. Kumpf and E.L. Nakamura (eds.) Proceeding of the Stock Identification Workshop. pp. 105-136. Pub. National Oceanographic and Atmospheric Administration. Panama City, Florida.

Avise, J.C. 2000. Phylogeography: The history and formation of species. Harvard University Press., Cambridge, Mass.

Bernatchez, L. 1997. Mitochondrial DNA analysis confirms the existence of two glacial races of rainbow smelt Osmerus mordax and their reproductive isolation in the St. Lawrence River estuary (Quebec, Canada) Mol. Ecol. 6:73-83.

Dudley, S.A. 1996. The response to differing selection on plant physiological traits: Evidence for local adaptation. Evolution 50: 103-110.

Graves, R.C. 1988. Geographicl distribution of the North Ameican tiger beetle *Cicindela hirticollis* Say. Cicindela 20(1):1-21.

Hadley, N. F., A. Savill, and T. D. Schultz. 1992. Coloration and its thermal consequences in the New Zealand tiger beetle *Neocicindela perhispida*. J. Thermal Biology . v. 17 p. 55-61

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. 2005a. Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology 25: 1965-1978.

Hijmans, R.J., Guarino, L., Jarvis, A., O'Brien, R., Mathur, P., Bussink, C., Cruz, M., Barrantes, I. and Rojas, E. 2005b. DIVA-GIS Version 5.2 manual. Available at: www.diva-gis.org

Kritsky, G. and L. Horner. 1998. Geographic variation in *Cicindela tranquebarica* Herbst (Coleoptera: Cicindelidae). Cicindela 30(3-4): 13-32.

Mayr, E. 1942. Systematics and the origin of species. Columbia Univ. Press, New York, NY.

Mayr, E. 1963. Animal species and evolution. Harvard University Press. Cambridge, Mass.

McKenna SJ, Raja Y, Gong S. 1999. Tracking colour objects using adaptive mixture models. Image Vis Comput. 17:225–231.

NatureServe. 2009. NatureServe Explorer: An online encyclopedia of life [web application]. Version 7.1. NatureServe, Arlington, Virginia. Available http://www.natureserve.org/explorer. (Accessed: June, 2010).

Nice, C. C. and J. A. Fordyce. 2006. How caterpillars avoid overheating: behavioral and phenotypic plasticity of pipevine swallowtail larvae. *Oecologia* 146:541-548

Phillimore and Owens. 2006. Are subspecies useful in evolutionary and conservation biology? Proc Royal Soc B 273:1049-1053.

Robertson, J.M. and K.R. Zamudio. 2009. Genetic diversification, vicariance, and selection in a polytypic frog. Journal of Heredity 100(6): 715-731.

Schultz, T.D. 1983. The ultrastructure, physiology, and ecology of epicuticular interference reflectors in tiger beetles (Cicindela). Ph.D. Dissertation. The University of Texas at Austin. 214 pp.

Schultz, T.D. and M.A. Rankin. 1983a. The ultrastructure of the epicuticular interference reflectors of tiger beetles (*Cicindela*). J. Exp. Biol. 117: 87-110.

Schultz, T.D. and M.A. Rankin. 1983b. Development changes in the interference reflectors and coloration of tiger beetles (*Cicindela*). J. Exp. Biol. 117:111-117.

Schultz, T.D. and G.D. Bernard 1989. Pointillistic mixing of interference colours in crypyic tiger beetles. Nature 337:72-73.

Schultz, T.D. 1986. Role of sctructural colors in predator avoidance by tiger beetles of the genus Cicindela (Coleoptera: Cicindelidae). Bull. Entomol. Soc. Am. 32:142-146.

Schultz, T. D. and N.F. Hadley. 1987. Structural colors of tiger beetles and their role in heat transfer through the integument. Physiological Zoology 60:737-745.

Schultz, T.D. 1991. Tiger Hunt. Natural History 100: 38-44.

Shelford, V.E. 1917. Color and color pattern mechanism of tiger beetles. Illinois Biological Monographs 3:395-532.

Sneath, P.H.A. and R.R. Sokal. 1973. Numerical taxonomy. The principles and practice of numerical classification. W.H. Freeman and Company. San Francisco. 573 pp.

Spanton, T.G. 1988. The *Cicindela sylvatica* group: Geographical variation and classification of the Nearctic taxa and reconstructed phylogeny and

geographical history of the species (Coleoptera: Cicindelidae). Quaestiones Entomologicae 24: 51-161.

Tan, C.C. 1945. Mosaic dominance in the inheritance of color patterns in the lady-bird beetle, *Harmonia axyridis*. Genetics 31:195-210.

Vogler, A. P., and R. DeSalle. 1993. Phylogeographic patterns in coastal North American tiger beetles (*Cicindela dorsalis* Say) inferred from mitochondrial DNA sequences. Evolution 47: 1192-1202.

Vogler, A.P. and R. DeSalle. 1994. Evolution and phylogenetic information content of the ITS-1 region in the tiger beetle, *Cicindela dorsalis*. Mol. Biol. Evol. 11(3): 393-405.

Willis, H.L. 1967. Bionomics and zoogeography of tiger beetles of saline habitats in the central United States (Coleoptera: Cicindelidae). Univer. Kans. Sci. Bull. 47:145-313.

Willis, H.L. 1968. Artificial key to the species of Cicindela of North America north of Mexico (Coleoptera: Cicindelidae). J. Kansas Entomol. Soc. 41:303-317.

Wilson, E. O., and W. L. Brown. 1953. The subspecies concept and its taxonomic application. Syst. Zool. 2(3):97-111.

Zink, R.M. 2004. The role of subspecies in obscuring avian biological diversity and misleading conservation policy. Proc. R. Soc. Lond. B 271: 561-564.

Table 5.1. Correlation of environmental variables used in the multiple regression analysis. For these bioclimatic variables, correlation is expected and inevitable (Hijmans *et al* 2005a). The matrix shows the degree to which each factor covaries and whether the relationship is positive or negative.

	Elev.	Ann Precip	Isotherm -ality	Max Temp	Mean Diurnal Range	Mean Temp Warm Quart	Mean Temp Wet Quart	Precip Warm Quart	Precip Wet Month	Precip Wet Quart	Temp Ann Range	Temp Season-ality	Ann Mean Temp
Elevation	*1.0000*	-0.1335	-0.3402	0.2123	-0.1086	-0.2009	-0.3678	-0.0109	0.2708	-0.2114	0.1064	0.0134	0.2195
Ann Precip	-0.1335	*1.0000*	0.1055	0.1095	0.2753	-0.1589	0.5296	-0.7344	-0.4719	0.1538	-0.3712	0.3590	0.1539
Isothermality	-0.3402	0.1055	*1.0000*	0.0770	-0.6372	0.2643	0.3090	0.2771	-0.2060	0.1779	0.2550	-0.1918	-0.3691
Max Temp	0.2123	0.1095	0.0770	*1.0000*	0.3108	-0.7996	0.3534	0.3579	0.0095	-0.0654	-0.7099	0.7661	0.6414
Mean Diurnal Range	-0.1086	0.2753	-0.6372	0.3108	*1.0000*	-0.5925	0.0996	-0.2308	-0.1996	0.1138	-0.8478	0.7465	0.6195
Mean Temp Warm Quart	-0.2009	-0.1589	0.2643	-0.7996	-0.5925	*1.0000*	-0.2569	-0.0899	0.0481	0.0265	0.8111	-0.9258	0.9720
Mean Temp Wet Quart	-0.3678	0.5296	0.3090	0.3534	0.0996	-0.2569	*1.0000*	-0.2431	-0.3266	0.1392	-0.3605	0.3453	0.1647
Precip Warm Quart	-0.0109	-0.7344	0.2771	0.3579	-0.2308	-0.0899	-0.2431	*1.0000*	0.3979	-0.1874	0.0294	-0.0188	-0.0150
Precip Wet Month	0.2708	-0.4719	-0.2060	0.0095	-0.1996	0.0481	-0.3266	0.3979	*1.0000*	-0.9411	0.2433	-0.2310	-0.0537
Precip Wet Quart	-0.2114	0.1538	0.1779	-0.0654	0.1138	0.0265	0.1392	-0.1874	-0.9411	*1.0000*	-0.1130	0.1037	-0.0172
Temp Ann Range	0.1064	-0.3712	0.2550	-0.7099	-0.8478	0.8111	-0.3605	0.0294	0.2433	-0.1130	*1.0000*	-0.9581	-0.7491
Temp Seasonality	0.0134	0.3590	-0.1918	0.7661	0.7465	-0.9258	0.3453	-0.0188	-0.2310	0.1037	-0.9581	*1.0000*	0.8819
Ann Mean Temp	0.2195	0.1539	-0.3691	0.6414	0.6195	0.9720	0.1647	-0.0150	-0.0537	-0.0172	-0.7491	0.8819	*1.0000*

Table 5.2. Multiple regression of color (*sensu* Spanton 1988) by environmental variables. Color states are ordered from 1-7: 1) black, 2) dark brown, 3) medium brown-bronze, 4) olive green, 5) green, 6) blue-green, 7) dark blue. Overall regression is highly significant ($P < 0.0001$), demonstrating that the environmental variables in the model are predictive of color phenotype. The R^2 indicates that over 75% of the variation in this phenotypic character can be explained by the set of variables.

Color (Spanton 1988) response:
R^2 = 0.751
Observations = 151

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Model	13	246.0496	18.9269	27.7403	**<0.0001***
Error	137	93.4736	0.6823		
Total	150	339.5232			

Parameter Estimates

Term	Partial Reg Coeff	Std Error	Std Coeff	t Ratio	Prob>|t|
Mean temperature wettest quarter	0.0027	0.0005	0.508	5.1369	<0.0001*
Max temperature warm month	0.0328	0.0068	2.966	4.8171	<0.0001*
Mean temperature warmest quarter	0.1052	0.0219	8.027	4.8153	<0.0001*
Precipitation of warmest quarter	-0.0054	0.0011	-0.907	-4.6654	<0.0001*
Temperature seasonality	0.0018	0.0004	10.665	4.1807	0.0001*
Annual mean temperature	0.0664	0.0168	5.424	3.9402	0.0001*
Temperature annual range	-0.0407	0.0105	-6.828	-3.8869	0.0002*
Annual total precipitation	0.0018	0.0005	2.171	3.4141	0.0008*
Mean diurnal temperature range	0.0350	0.0112	2.066	3.1075	0.0023*
Isothermality	-0.0511	0.0324	-0.895	-1.5773	0.1170
Precipitation of wettest quarter	-0.0034	0.0026	-2.034	-1.3166	0.1902
Elevation	0.0000	0.0000	0.107	0.6285	0.5307
Precipitation of wettest month	0.0037	0.0082	0.766	0.4472	0.6555

Table 5.3. Multiple regression of color (hue) by environmental variables. Hue was determined with Photoshop CS5 by using the color picker tool and selecting a 101x101 pixel average of the elytra disk. Overall regression is highly significant ($P < 0.0001$), demonstrating that the environmental variables in the model are predictive of color phenotype. The R^2 indicates that over 62% of the variation in this phenotypic character can be explained by the set of variables.

Color (hue) response:
$R^2 = 0.623$
Observations = 151

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Model	13	108762.62	8366.36	17.4511	**<0.0001***
Error	137	65680.08	479.42		
Total	150	174442.70			

Parameter Estimates

Term	Partial Reg Coeff	Std Error	Std Coeff	t Ratio	Prob>|t|
Max temperature warm month	3.8109	0.6895	4.184	5.5267	<0.0001*
Mean temperature warmest quarter	11.4862	2.2097	10.655	5.1981	<0.0001*
Annual mean temperature	7.4513	1.7039	7.403	4.3732	<0.0001*
Temperature seasonality	0.1695	0.0447	11.890	3.7904	0.0002*
Mean temperature wettest quarter	0.1725	0.0536	0.391	3.2155	0.0016*
Precipitation of wettest quarter	-0.8403	0.2636	-6.056	-3.1874	0.0018*
Temperature annual range	-3.3709	1.0577	-6.884	-3.1870	0.0018*
Precipitation of wettest month	2.4974	0.8326	6.316	2.9993	0.0032*
Mean diurnal temperature range	2.1060	1.1375	1.514	1.8515	0.0662
Isothermality	-4.5551	3.2756	-0.970	-1.3906	0.1666
Elevation	0.0010	0.0022	0.092	0.4381	0.6620
Precipitation of warmest quarter	-0.0494	0.1160	-0.102	-0.4261	0.6707
Annual total precipitation	0.0171	0.0526	0.254	0.3254	0.7454

Table 5.4. Multiple regression of color saturation by environmental variables. Saturation was determined with Photoshop CS5 by using the color picker tool and selecting a 101x101 pixel average of the elytra disk. Overall regression is highly significant (*P* <0.0001), demonstrating that the environmental variables in the model are predictive of color phenotype. The R^2 indicates that over 38% of the variation in this phenotypic character can be explained by the set of variables.

Color saturation response:
R^2 = 0.388
Observations = 151

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Model	13	5555.299	427.331	6.6821	**<0.0001***
Error	137	8761.310	63.951		
Total	150	14316.609			

Parameter Estimates

Term	Partial Reg Coeff	Std Error	Std Coeff	t Ratio	Prob>|t|
Elevation	0.0024	0.0008	0.795	2.9810	0.0034*
Annual mean temperature	-1.6531	0.6223	-5.733	-2.6565	0.0088*
Mean temperature warmest quarter	2.0566	0.8071	6.660	2.5483	0.0119*
Temperature seasonality	-0.0316	0.0163	-7.728	-1.9323	0.0554
Annual total precipitation	0.0298	0.0192	1.548	1.5526	0.1228
Temperature annual range	0.5963	0.3863	4.251	1.5435	0.1250
Max temperature warmest month	-0.3886	0.2518	-1.490	-1.5430	0.1251
Precipitation of wettest month	-0.4683	0.3041	-4.135	-1.5400	0.1259
Precipitation of warmest quarter	-0.0580	0.0424	-0.417	-1.3690	0.1732
Precipitation of wettest quarter	0.1264	0.0963	3.180	1.3128	0.1914
Mean diurnal temperature range	-0.4333	0.4154	-1.087	-1.0429	0.2988
Mean temperature wettest quarter	-0.0096	0.0196	-0.076	-0.4901	0.6248
Isothermality	0.4845	1.1963	0.360	0.4050	0.6861

an temperature, Prob>|t| = 0.0167*.

Table 5.5. Multiple regression of color brightness by environmental variables. Brightness was determined with Photoshop CS5 by using the color picker tool and selecting a 101x101 pixel average of the elytra disk. Overall regression is highly significant (*P* <0.0001), demonstrating that the environmental variables in the model are predictive of color phenotype. The R^2 indicates that over 53% of the variation in this phenotypic character can be explained by the set of variables.

Color brightness response:
R^2 = 0.530
Observations = 151

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Model	13	5451.341	419.334	11.8604	**<0.0001***
Error	137	4843.746	35.356		
Total	150	10295.086			

Parameter Estimates

Term	Partial Reg Coeff	Std Error	Std Coeff	t Ratio	Prob>|t|
Mean diurnal temperature range	1.5885	0.3089	4.700	5.1427	<0.0001*
Isothermality	-4.3314	0.8895	-3.797	-4.8693	<0.0001*
Temperature annual range	-1.0185	0.2872	-8.562	-3.5458	0.0005*
Precipitation of warmest quarter	-0.0968	0.0315	-0.821	-3.0731	0.0026*
Precipitation of wettest month	-0.6584	0.2261	-6.855	-2.9117	0.0042*
Precipitation of wettest quarter	0.1870	0.0716	5.548	2.6121	0.0100*
Elevation	0.0015	0.0006	0.593	2.5348	0.0124*
Annual total precipitation	0.0340	0.0143	2.078	2.3780	0.0188*
Temperature seasonality	0.0244	0.0121	7.058	2.0126	0.0461*
Max temperature warm month	0.2499	0.1873	1.130	1.3345	0.1842
Mean temperature warm quarter	-0.5315	0.6001	-2.029	-0.8857	0.3773
Mean temperature wettest quarter	0.0100	0.0146	0.093	0.6838	0.4952
Annual mean temperature	0.1968	0.4627	0.805	0.4252	0.6713

Table 5.6. Multiple regression of maculation percentage (unpigmented areas) by environmental variables. Percent maculation was determined with Photoshop CS5 by using the lasso tool to select the maculated areas, recording the total number of pixels, then dividing this by the total pixel count for the elytra. Overall regression is highly significant (P <0.0001), demonstrating that the environmental variables in the model are predictive of color phenotype. The R^2 indicates that over 40% of the variation this phenotypic character can be explained by the set of variables.

Maculation percentage response:
R^2 = 0.403
Observations = 151

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Model	13	1567.030	120.541	7.1138	**<0.0001***
Error	137	2321.415	16.945		
Total	150	3888.445			

Parameter Estimates

Term	Partial Reg Coeff	Std Error	Std Coeff	t Ratio	Prob>|t|
Elevation	0.0008	0.0004	0.499	1.8940	0.0603+
Annual total precipitation	-0.0185	0.0099	-1.844	-1.8730	0.0632
Annual mean temperature	0.5771	0.3203	3.840	1.8016	0.0738+
Precipitation of wettest month	0.2774	0.1565	4.699	1.7718	0.0786
Mean temperature warmest quarter	-0.5886	0.4154	-3.657	-1.4168	0.1588
Precipitation of warmest quarter	0.0300	0.0218	0.414	1.3756	0.1712
Precipitation of wettest quarter	-0.0674	0.0496	-3.253	-1.3600	0.1761
Mean temperature wettest quarter	-0.0105	0.0101	-0.160	-1.0454	0.2977
Temperature seasonality	0.0082	0.0084	3.859	0.9769	0.3304
Isothermality	0.3806	0.6158	0.543	0.6180	0.5376
Temperature annual range	-0.0587	0.1989	-0.803	-0.2951	0.7684
Max temperature warmest month	0.0269	0.1296	0.198	0.2073	0.8361
Mean diurnal temperature range	-0.0309	0.2138	-0.149	-0.1443	0.8855

+Leverage plot indicated AZ: Coconino Co population as outliers. When removed elevation was found to be significant, Prob>|t| = 0.0041*, as well as annual mean temperature, Prob>|t| = 0.0167*.

Figure 5.1. Elytral color patterns of *C. longilabris* subspecies and "intergrades" (*sensu* Spanton 1988). (See Fig. 4.11 for geographic range of each) Specimens shown are a small subsample of those scored for multivariate analyses of phenotype and environment, and illustrate the range of variation within each. **Top row)** *C. longilabris longilabris* 1) AK: Valdez, 2) AK: Koyukuk, 3) BC: Castlegar, 4) MB: Orr Creek, 5) ON: Terrace Bay, 6) ON: Cochrane, 7) QC: Duparquet, 8) QC: LaTuque. **Second row)** *C. l. perviridis* 1-3) WA: Olympic National Park, 4) OR: Lake of the Woods, 5) CA: Susanville area, 6-8) Tioga Pass. **Third Row)** *C. l. laurentii* 1) WY: Bighorn Mts, 2-4) CO: Poudre Canyon, 5) CO: San Juan Mts, 6) AZ: North Rim area, 7) UT: Cedar Breaks area, 8) NV: Lee Canyon. **Bottom Row)** "Intergrade" populations that are purported to be the result of mixing between the nominal subspecies. 1-2) OR: Mt Pisgah, 3-4) ID: Snowhaven, 5-8) MT: Stevensville area

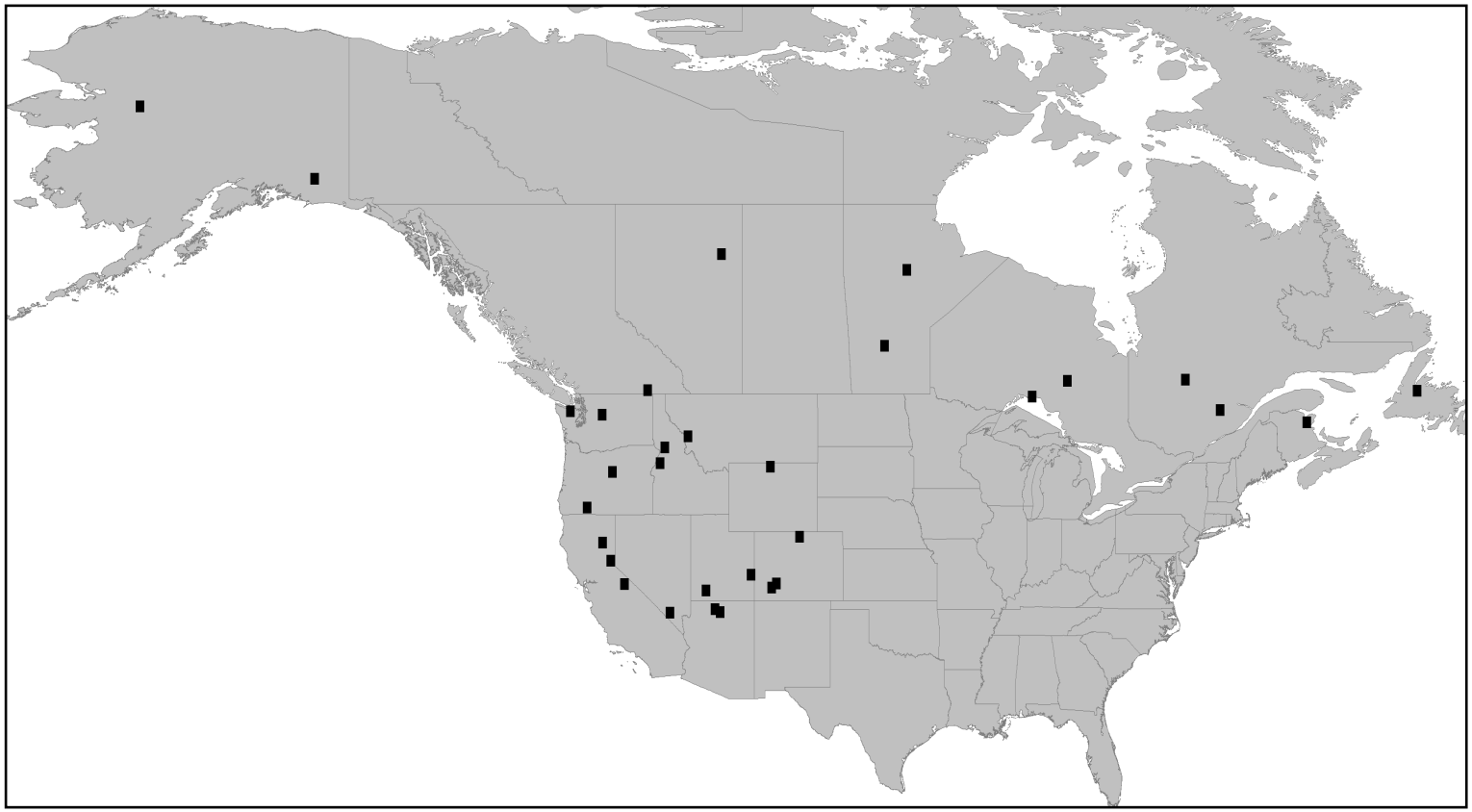


Figure 5.2. Sampling localities for *C. longilabris* used in multivariate analyses of phenotype and environmental variables. A total of 31 sites and 151 individuals were included. These represent each subspecies (*sensu* Spanton 1988) and "intergrades" where the subspecies are reported to mix. See Figure 4.11 for more detail on the distribution of *C. longilabris* subspecies.

CHAPTER VI

CONCLUSIONS

In this dissertation, I used multiple phylogenetic, population genetic, and coalescent-based molecular analyses in conjunction with GIS-based environmental niche modeling and statistical analyses of phenotypic characters to uncover the evolutionary history and ecology of a group of North American tiger beetles. This integrative multidisciplinary approach can serve as a model for future studies. Each series of investigations reciprocally informed the others, leading to a more robust set of conclusions about the causes of differentiation and speciation in these taxa. I addressed three fundamental questions with this approach: 1) What factors promote population differentiation and speciation? 2) What are the determinants of species ranges? and 3) What are the principal causes of phenotypic variation? The results of this dissertation also have important consequences for systematic research in general, especially for the disciplines of molecular systematics, "molecular taxonomy" (Blaxter and Floyd 2003; Tautz *et al* 2003; Blaxter 2004), and "DNA barcoding" (Hebert *et al* 2003 a, b). In addition to these larger issues, there are more specific systematic and taxonomic implications for the North American *Cicindela sylvatica* group and consequences for future tiger beetle research.

Factors promoting population differentiation and speciation

Using a 'congeneric phylogeography' approach (Funk and Omland 2003), I sampled intensively from all the nominal species of the North American *C. sylvatica* group including all recognized subspecies (Spanton 1988; Freitag 1999) and as many notable variants as possible. Phylogenetic methods (Chapter II) recovered two genetically and ecologically distinct clades: The "Meadow Group" (*C. longilabris* and *C. nebraskana*) often associated with alpine meadows and grasslands (Leffler and Pearson 1976; Spanton 1988), and the "Forest Group" (*C. sexguttata, C. patruela,* and *C. denikei*), typically found in forested areas and adjacent ecotones (Kaulbars and Freitag 1993a, b). My phylogeographic tests demonstrated that the nominal taxa of the Meadow Group were not separate species, but appear to represent phenotypic polymorphism within a single species. Nonetheless, deep genetic breaks were discovered within the Meadow Group, and coalescent analyses supported fragmentation during the Quaternary Ice Ages with evidence for recent demographic expansion following glacial retreat (Chapter III). Given the estimates of divergence times (1.3-2.4 mya, early-mid Pleistocene) and the allopatric distribution of clades, historical effects of isolation due to glacial cycles (Sibrava *et al* 1986) are the most likely cause for genetic differentiation in the Meadow Group. Although the nominal species, *C. longilabris* and *C. nebraskana*, were not supported by any phylogenetic or clustering analysis based on mtDNA (973 bp COI/COII) or AFLPs (1252 loci), they could represent populations containing specific genes under strong

selection against an otherwise undifferentiated genomic background. It is possible that particular areas of the genome may represent "genomic islands" associated with ecological and morphological traits (Nosil *et al* 2009; Michel *et al* 2010) under divergent selection. As such, I analyzed these nominal species with Ecological Niche Models (ENMs) treating them as separate taxa. Statistical tests of niche overlap and differentiation (Warren *et al* 2008, 2010) demonstrated that although *C. longilabris* and *C. nebraskana* were ecologically similar there was significant differentiation (Chapter IV). Future genome-scan analyses may reveal specific loci under divergent selection.

Markedly different divergence patterns were observed within the Forest Group. Although none of the three nominal species was monophyletic with respect to mtDNA (Chapter II), combined multilocus analyses demonstrated that the polyphyly between *C. sexguttata* and *C. patruela* could best be explained by occasional and ongoing hybridization and mtDNA introgression in some areas of geographic contact. Niche modeling and statistical testing of overlap and differentiation revealed that the two species are ecologically similar but significantly differentiated (Chapter IV). The putative divergence between nominal species *C. sexguttata* and *C. denikei* was substantially more complex to infer. The latter taxon exists as a "peripatric" (Mayr 1942) set of satellite populations with respect to the geographically more widespread *C. sexguttata*. Phylogenetic and clustering analyses did not recover monophyletic groups or clusters corresponding to those species boundaries. However coalescent simulations indicated that insufficient evolutionary time has passed to allow for

the purging of ancestral genetic polymorphisms (*i.e.* incomplete lineage sorting), even if speciation had occurred (Chapter II). Coalescent simulations indicated that <10,000 years would be required for complete lineage sorting. I used ecological niche models to assess *C. denikei*'s niche and subsequently to retrodict where the species may have occurred during the Last Glacial Maximum, ~18,000 years ago (CLIMAP 1981). Results demonstrated that there was low probability of any suitable habitat at the LGM. If there had been evidence of *C. denikei*'s existence during this time, it would have indicated that incomplete lineage sorting was an unlikely explanation for the polyphyly. Lastly, tests of ecological differentiation indicated that *C. denikei* and *C. sexguttata* displayed little niche overlap and were more differentiated than could be expected by chance, even accounting for inherent differences in environmental conditions due to allopatry. Given the combined pattern of low genetic differentiation and high ecological differentiation in context with the geographic distributions, this pattern suggested "peripatric ecological speciation" (Funk *et al* 1995).

Determinants of species ranges

I used ENMs and jackknife tests to identify and quantify the factors that could best explain species geographic distributions. Based on the general concordance with known ranges and the high estimates of predictive power for the ENMs, intrinsic physiological tolerances to environmental conditions are most supported as the general explanation for the species distributions (Chapter IV).

None of the five nominal species were identical in ecological niche, nor were any identical for the factors that were inferred to be most limiting. Previous workers had hypothesized about the factors believed to be most critical in delineating *Cicindela* ranges (Leffler 1979; Schincariol and Freitag 1991), including the North American *C. sylvatica* group (Spanton 1988; Schultz *et al* 1992; Kaulbars and Freitag 1993b) and the ENM jackknife results were consistent with these predictions. Temperature (maximum and mean) was critical for predicting the distribution of *C. longilabris*, as well as ecoregion (*i.e.* general landscape and habitat) and soil order. Similar results were found in *C. nebraskana*, although the relative contributions were not identical. *Cicindela sexguttata* and *C. patruela* distributions were best predicted by ecoregion and soil type. The next most important limiting factors appeared to be annual total precipitation for *C. sexguttata*, and maximum temperature for *C. patruela*. It is difficult to make conclusions regarding the limiting factors for *C. denikei*, because although most variables were equally highly predictive, this is likely a modeling artifact due to the species' apparently small geographic range (Phillips *et al* 2006).

Because Meadow Group mtDNA clades were found to be tightly allopatric and non-overlapping even in close geographic proximity, they were analyzed using the ENM approach to identify the possible explanations for the *current* limits of their distributions. Results were not consistent with intrinsic physiological tolerances limiting their distributions, but instead showed a pattern highly suggestive of an extrinsic biological cause. Possible explanations involve competitive exclusion between the clades, or perhaps the result of cytoplasmic

incompatibilities as a result of *Wolbachia* bacterial infections (Hoffman and Turelli 1997).

Causes of phenotypic variation

Cicindela longilabris exhibits substantial variation in color and pattern throughout its range, and as a result prior workers have described numerous subspecies and forms (reviewed in Spanton 1988). Spanton revised the North American *C. sylvatica* group recognizing three subspecies, yet considerable variation exists even within each of these. My mtDNA phylogeography results did not support the existence of any of these phenotypically and geographically defined subspecies (Chapter II). Moreover, genome-wide AFLP cluster analyses revealed no patterns of population structure consistent with the nominal subspecies. Although subspecies were not supported by the genetic results, I sought to investigate the association between phenotype and environment, by performing multivariate analyses using environmental layers and quantitative measures of phenotype. I scored characters for color (four characteristics) and the percentage of the elytra that are maculated (unpigmented) for populations distributed throughout the range of *C. longilabris*. Results demonstrated that environment was significantly predictive of these phenotypic characteristics, and the high coefficients of determination showed that a relatively high percentage (~40-75%) of the variation in these phenotypic characters is predicted by the environmental model. Although it is not possible to make conclusions on the

underlying genetic basis for each of these phenotypic traits, these results do suggest that color may be at least partly due to the effects of developmental plasticity, consistent with earlier tiger beetle research on other *Cicindela* species (Shelford 1917; Schultz 1983). Future studies may build upon these findings by employing a genome-scan approach (*e.g.* Egan *et al* 2008) to identify outlier AFLP loci that are segregating according to specific phenotypic characters for color or pattern.

Implications for molecular taxonomy, DNA barcoding

Species delineation and species identification are increasingly based on molecular data (*e.g.* Tautz *et al* 2002; Blaxter and Floyd 2003; Hebert *et al* 2003a, b; Blaxter 2004; Pons *et al* 2006), and species-level polyphyly is especially a relevant issue as it relates to the increasingly popular yet controversial method of "DNA barcoding" (Hebert *et al* 2003a, b). DNA barcoding aims to identify specimens solely through the use of a 600bp fragment of mitochondrial DNA. The methodology is founded on the assumption that individuals within species will form monophyletic groups to the exclusion of individuals from all other species. If this assumption is incorrect for a particular taxon then one possible result is a polyphyletic pattern, resulting in misidentification of the specimens of interest. Furthermore, it would be impossible to empirically observe a pattern of polyphyly unless multiple individuals of at least two species are examined. The 'congeneric

phylogeography' approach suggested by Funk and Omland (2003) involves intensive geographic sampling of all species that may potentially share alleles. I adopted this approach here by thoroughly sampling across all species, subspecies, and notable variants within the five nominal species of the North American *C. sylvatica* group. My results showed that no taxonomic species was monophyletic and further hypothesis testing uncovered the underlying evolutionary causes of the species polyphyly. Moreover, I discovered that even well established species were polyphyletic with respect to each other (*C. sexguttata* vs. *C. patruela*) and the importance of this result is underscored by the fact that *C. patruela* is a candidate for the endangered species list (NatureServe 2009, accessed May 2010). As such, a barcoding approach would have failed to identify a conservationally important taxon.

Implications for tiger beetle taxonomy and future studies

Even though tiger beetles are one of the most well-studied non-pest insect groups (Knisley and Schultz 1997) and multiple taxonomic revisions and dozens of natural history accounts have been published on the group, I was able to make powerful new inferences about the species limits and evolutionary history as a result of the integrative multidisciplinary approach taken in my dissertation. Systematic relationships and tiger beetle taxonomy will have to be re-evaluated as a result of these findings. The first conclusion is that *C. longilabris* and *C. nebraskana* are not separate species, as supported by multiple lines of evidence

from phylogenetic, coalescent, and population genetic analyses. Given that Say described *C. longilabris* in 1817, and Casey described *C. nebraskana* in 1909, any future formal taxonomic change would justify treating *C. nebraskana* as a junior synonym. These findings suggest that the latter form may be akin to a "variant", or "ecomorph". Although *C. longilabris* subspecies were not supported as a whole, ecological niche modeling demonstrated significant niche differentiation between the nominate subspecies *C. l. longilabris* and the other two subspecies, *C. l. laurentii*, and *C. l. perviridis*. Given the evidence of niche differentiation (*e.g.* local adaptation *sensu* Mayr 1970), and general allopatric distribution between the nominate subspecies and the other forms, one could make an argument for the future retention of two subspecies within *C. longilabris*. *C. patruela* and *C. sexguttata* are distinct species as has been accepted by most tiger beetle workers over the past fifty years (Wallis 1961; Willis 1968; Boyd 1988; Freitag 1999; Wiesner 1999), although before this study it was not known that they can and do occasionally hybridize. The collective evidence of genetic and ecological niche model analyses suggests that *C. denikei* is indeed a separate species from *C. sexguttata*. Recent checklists reflect this opinion (Freitag 1999; Wiesner 1999), although prior to Kaulbars and Freitag's work (1993a, b) *C. denikei* was generally treated as a subspecies or variant of *C. sexguttata*.

Future tiger beetle systematics and taxonomy, especially at the species-level could benefit greatly from a more multidisciplinary approach shown here. The integration of genetics, ecology, and morphology allows for more rigorous

hypothesis testing of the systematic and ecological relationships in this group of tiger beetles.

References

Blaxter, M.L., and R. Floyd. 2003. Molecular taxonomics for biodiversity surveys: already a reality. Trends Ecol Evol 18:268-269.

Blaxter, M.L. 2004. The promise of a DNA taxonomy. Philos Trans R Soc Lond B Biol Sci 359(1444): 669-679.

Boyd, H. P. and Associates. 1982. Checklist of Cicindelidae: The Tiger Beetles. Plexus Publishing Inc., New Jersey. 31 pp.

CLIMAP. 1981. Seasonal reconstructions of the Earth's surface at the last glacial maximum in Map Series, Technical Report MC-36. Boulder, Colorado: Geological Society of America.

Egan, S.P., Nosil, P., Funk D.J. 2008. Selection and genomic differentiation during ecological speciation: Isolating the contributions of host association via a comparative genome scan of *Neochlamisus bebbianae* leaf beetles. Evolution 62:1162–1181.

Freitag, R. 1999. Catalogue of the Tiger Beetles of Canada and the United States. NRC Research Press, Ottawa, Ontario, Canada.

Funk, D.J., D.J. Futuyma, G. Orti, and A. Meyer. 1995. A history of host associations and evolutionary diversification for *Ophraella* (Coleoptera: Chrysomelidae): new evidence from mitochondrial DNA. Evolution 49: 1008-1017.

Funk, D. J. and K. E. Omland. 2003. Species-level paraphyly and polyphyly: Frequency, causes, and consequences, with insights from animal mitochondrial DNA. AREES 34:397-423.

Hebert, P.D.N., A. Cywinska, S.L. Ball, and J.R. deWaard. 2003. Biological identifications through DNA barcodes. Proc Biol Sci, 270(1512), 313-321.

Hebert, P.D.N., S. Ratnasingham, and J.R. deWaard. 2003. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. Proc Biol Sci, 270 Suppl 1, S96-99.

Hoffmann, A.A., and M. Turelli. 1997. Cytoplasmic incompatibility in insects. In: R.V. O'Neill, A.A. Hoffmann and J.H. Werren, Editors, *Influential Passengers*, Oxford University Press, Oxford, UK, pp. 42–80.

Kaulbars, M.M., and R. Freitag. 1993a. Foraging behaviour of the tiger beetle *Cicindela denikei* Brown (Coleoptera: Cicindelidae). Can. Field-Nat. 107:53-58.

Kaulbars, M.M., and R. Freitag. 1993b. Geographic variation, classification, reconstructed phylogeny, and geographic history of the *Cicindela sexguttata* group (Coleoptera: Cicindelidae). Can. Entomol. 125: 267-316.

Knisley, C.B. & T.D. Schultz. 1997. The Biology of Tiger Beetles and a Guide to the Species of the South Atlantic States. Virginia Museum of Natural History, Martinsville, VA. 210 pp.

Leffler, S. R. 1979. Tiger beetles of the Pacific Northwest (Coleoptera: Cicindelidae). Ph.D. Dissertation, University of Washington, Seattle. 731 pp.

Michel, A.P., S. Sim, T.H.Q. Powell, M.S. Taylor, P. Nosil, and J.L. Feder. 2010. Widespread genomic divergence during sympatric speciation. Proceedings of the National Academy of Sciences of the United States of America 107(21): 9724-9729.

Phillips, S.J., R.P. Anderson, and R.E. Schapire. 2006. Maximum entropy modeling of species geographic distributions. Ecological Modelling 190: 231-259.

NatureServe. 2009. NatureServe Explorer: An online encyclopedia of life [web application]. Version 7.1. NatureServe, Arlington, Virginia. Available http://www.natureserve.org/explorer. (Accessed: May 2010)

Nosil P, Funk D.J., Ortiz-Barrientos D. 2009. Divergent selection and heterogeneous genomic divergence. Molecular Ecology, 18, 375–402.

Pons, J., T.G. Barraclough, J. Gomez-Zurita, A. Cardoso, D. Duran, S. Hazell, S. Kamoun, W.D. Sumlin, A.P. Vogler. 2006. Evolutionary species delineation for the DNA taxonomy of undescribed insects. Systematic Biology 55(4): 595-609.

Tautz, D., P. Arctander, A. Minelli, R. H. Thomas, and A. P. Vogler. 2003. DNA points the way ahead in taxonomy. Nature 418:479.

Schincariol, L.A., and R. Freitag. 1991. Biological character analysis, classification, and history of the North American *Cicindela splendida* group taxa (Coleoptera: Cicindelidae). Can. Entomol. 123:1327-1353.

Schultz, T.D. 1983. The ultrastructure, physiology, and ecology of epicuticular interference reflectors in tiger beetles (Cicindela). Ph.D. Dissertation. The University of Texas at Austin. 214 pp

Schultz, T.D., M.C. Quinlan, and N.F. Hadley. 1992. Preferred body temperature, metabolic physiology, and water balance of adult *Cicindela longilabris*: A comparison of populations from boreal habitats and climatic refugia. Physiological Zoology 65(1): 226-242

Shelford, V.E. 1917. Color and color pattern mechanism of tiger beetles. Illinois Biological Monographs 3:395-532.

Sibrava, V., Bowen, D.Q, & Richmond, G.M., 1986, Quaternary Glaciations in the Northern Hemisphere, Quaternary Science Reviews. Vol. 5, pp. 1-514.

Spanton, T.G. 1988. The *Cicindela sylvatica* group: Geographical variation and classification of the Nearctic taxa and reconstructed phylogeny and geographical history of the species (Coleoptera: Cicindelidae). Quaestiones Entomologicae 24: 51-161.

Vogler, A.P. and A. Welsh. 1997. Phylogeny of North American *Cicindela* tiger beetles inferred from multiple mitochondrial DNA sequences. Mol. Phylogenet. Evol. 6: 321-338.

Wallis, J.B.1961. The Cicindelidae of Canada. University of Toronto Press, Toronto, Canada. 74 pp.

Warren, D.L., R. E. Glor, and M. Turelli. 2008. Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. Evolution 62:2868-2883.

Warren, D.L., R. E. Glor, and M. Turelli. 2010. ENMTools: a toolbox for comparative studies of environmental niche models. Ecography 33(1): DOI: 10.1111/j.1600-0587.2009.06142.x

Wiesner, J. 1999. Verzeichnis der Sandlaufkäfer der Welt, Checklist of the Tiger Beetles of the World,. Verlag Erna Bauer, Keltern, pp.364.

Willis, H.L. 1968. Artificial key to the species of Cicindela of North America north of Mexico (Coleoptera: Cicindelidae). J. Kansas Entomol. Soc. 41:303-317.