Title
Insect ecology from the scale of plant-level interactions to continent wide phylogeography

Author
Madsen-Steigmeyer, Tara

Publication Date
2014

Peer reviewed|Thesis/dissertation

Insect ecology from the scale of plant-level interactions to continent wide phylogeography

by

Tara Madsen-Steigmeyer

A dissertation submitted in partial satisfaction of the
requirements for the degree of

Doctor of Philosophy

in

Environmental Science, Policy and Management

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Neil D. Tsutsui, Chair
Professor Stephen C. Welter
Professor Erica B. Rosenblum
Professor David D. Ackerly

Spring 2014

Abstract

Insect ecology from the scale of plant-level interactions to continent wide phylogeography

by

Tara Madsen-Steigmeyer

Doctor of Philosophy in Environmental Science, Policy and Management

University of California, Berkeley

Professor Neil D. Tsutsui, Chair

This dissertation considers insect ecology from two perspectives, at two scales of space and time, using very different methodologies. All three chapters, however, examine insect ecology within the framework of a plant-based environment. In the first chapter, I investigate the direct interactions between plants and insects through plant defenses against herbivory. In the second and third chapters, I investigate the geographic patterns of genomic diversity in an herbivorous insect that underwent a dramatic host switch to become an important invasive agricultural pest.

The first chapter explores the effect of plant exposure to smoke on herbivore growth and development. While previous research has examined the direct effects of fire on plant and insect communities, the potential for smoke itself to interact with these communities has received little attention. Plant defenses against herbivory rely on a complex network of environmental stimuli, plant signaling pathways, and physiological responses. Because smoke is chemically complex, and plant defenses are highly sensitive to airborne chemicals and abiotic stressors, smoke may play an important and previously unrecognized role in plant-insect interactions in the post-fire environment of both burned and adjacent un-burned communities. I used wild tobacco, a plant with both physiological responses to smoke exposure and herbivore defenses sensitive to volatile organic compounds, and its insect herbivore *Manduca sexta*, to test the hypothesis that smoke exposure would change insect growth by altering plant defenses. I grew *Nicotiana attenuata* in the greenhouse, and exposed treatment plants to 20 minutes of smoke fumigation. I conducted feeding experiments where I measured growth of *M. sexta* larvae reared on treated and untreated plants, and choice tests where I tested for insect preference for treated versus untreated plants. I found no significant difference between growth of *M. sexta* reared on smoke-exposed plants and of those reared on untreated plants. Larvae did not exhibit a preference for or against smoke exposed plants. Despite finding no effect in these tests, I believe that the mechanisms for potential impacts of smoke exposure on plant defenses are convincing enough to warrant further study. Examination of effects on the third trophic level, on a generalist rather than specialist herbivore, or on herbivores of a perennial rather than annual plant may reveal novel impacts of natural smoke in the environment. Such information would increase scientific understanding of the sensitivity of plant responses to environmental stimuli, and dramatically expand the spatial boundaries we must use when describing the role of fire in the environment.

The second chapter steps back from near-term, small scale, and direct plant-insect interactions at the individual level to take a long-term and large scale perspective on insect ecology through the examination of the population genetics of an introduced plant pest. The

Colorado potato beetle, *Leptinotarsa decemlineata*, is one of the most successful insect pests of agriculture. However, its evolutionary origins as an insect pest of potato remain unclear. I used genomic restriction-site associated DNA sequencing of 38 samples collected from across the native range of Mexico and from the East Coast of the USA in the introduced range to describe the geographic population structure of the beetle, and to examine the genomic diversity across the samples. I used a matrix of 4,542 SNPs to identify four genetic clusters that corresponded with the four geographic regions from which beetle samples were collected. In the native range, discriminant analysis of principle components divided samples from Southern lowland Mexico, Central highland Mexico, and Western Mexico into three unique genetic clusters. Samples collected from potato along the East Coast of the USA formed the fourth group. Genetic differentiation between Western and Southern Mexico was low, based on pairwise F_{ST} values, as was differentiation between Central Mexico and USA samples. Genomic similarity with Central Mexico suggests this is a likely source region for introduced beetles in the USA. Diversity measures were higher within the introduced region than across all regions of the native range, including expected and observed heterozygosities as well as the number of private alleles. High diversity in the introduced range is unexpected due to the demographic forces typically at work during the invasion process. One mechanism that can maintain diversity is multiple introductions, a possibility I tested in the third chapter.

In the final chapter, I take a phylogenomic approach to understanding the invasion history of Colorado potato beetle. Using 3,122 SNPs drawn from the previously identified RAD loci, I conducted a phylogenomic analysis on 38 beetles collected from the native and introduced range. The major features of the tree topology were consistent between different inference methods. The USA samples formed a monophyletic clade, sister to the similarly monophyletic Central Mexico group. In the native range, I identified a third monophyletic clade corresponding with the Western Mexico region, while the Southern Mexico samples formed a fourth but paraphyletic cluster. These results support Central Mexico as the most likely native source of the introduced beetles, and indicate that multiple introductions were unlikely. Together with the previous chapter, my research on the geographic diversity, genomic diversity, and historic origins of the Colorado potato beetle makes significant contributions to scientific knowledge of this important crop pest. The RAD loci and SNPs I identified will enable future efforts to determine the genomic basis for the success of this beetle, and to unravel details of the Colorado potato beetle's genomic response to selective forces encountered after switching host plants.

In summary, this dissertation presents a novel perspective on the influence of an understudied abiotic condition on plant-insect interactions, adds new evidence to the literature on the invasion history of the Colorado potato beetle, and provides a unique genomic resource for future work on this important agricultural pest.

Contents

List of Figures

List of Tables

Acknowledgements

This dissertation would not have been possible without the years of patience and support given me by my advisor and mentor, Steve Welter. He provided an environment in which I could learn, grow, and improve as a scientist, a teacher, and a human being. In characteristic Welter fashion, he also helped me forge the connections with other scientists that were essential to my research.

This dissertation would never have been completed without the unfaltering encouragement of Neil Tsutsui. He welcomed me into the discourse of his lab meetings long before adopting me into his lab. Without his advice and guidance, I would not have successfully navigated the critical transition in my dissertation from trophic ecology to molecular.

Yolanda Chen has been a mentor, an advocate, and a keen reader. Most significantly, she welcomed me as a collaborator in her research on the Colorado potato beetle. She provided essential scientific inspiration, as well as much appreciated moral support and encouragement. Her graduate student, Vic Izzo, was enormously generous with his CPB collections.

I owe special thanks to several faculty at UC Berkeley. E. Bree Rosenblum has been a role model and mentor. Not only is she an inspiring teacher, she also embraced my desire to participate in the inspiration when I worked with her as a graduate student instructor. David Ackerly always engaged me in stimulating discussions and asked thought provoking questions, regardless of how long it had been since I last showed up at his office. Max Moritz helped set me off on a fascinating research path through uncharted, smoky country. David and Max, together with Nick Mills and Perry de Valpine, helped make my qualifying exams an experience that was much more exciting and fascinating than stressful.

I am immensely grateful to all my Welter lab colleagues. Dani Casado answered endless questions about statistics and R programming, and always knew when I especially needed a hug. Steve Bayes provided years of camaraderie and feedback. E.J. Blitzer, Sara Emery, and Frances Cave helped make the lab the supportive, stimulating, and friendly place it was.

At Berkeley, I had the pleasure of teaching labs and discussions in ESPM as well as Integrative Biology. I am grateful to the many students who tolerated my instruction and helped me improve it. Marty Covington, Linda von Hoene, and Ellie Schindelman taught me a good deal of pedagogy, and in the ESPM & IB graduate group Teaching Counts I found other graduate students who shared my interest in effective teaching in the University environment.

Financial support for my dissertation was provided by the ESPM Department, the Julius M. Freitag Memorial Award, and the Edward Steinhaus Memorial Award. The Vermont Agricultural Experiment Station funded my lab work and genomic sequencing for the CPB.

The unconditional support of my parents gave me strength in this endeavor, as it has throughout my life. Irenka D. Pareto was my bosom companion through the ups and downs of the entire journey. Kirk H. and Frank C. bicycled thousands of miles with me, tolerated more griping than any friends should have to, and joined me in celebrating physical feats of endurance when those miles seemed my only forward progress, and those feats my only accomplishments.

And last, I am deeply grateful to my husband, Zak Steigmeyer, for his enduring patience, absolute faith, and grounding perspective. I could not have made this journey without him, nor would I have wanted to. Levi, also, brought tremendous joy to the last two and a half years of my dissertation.

Introduction

My dissertation contributes to our understanding of insect ecology by examining plant-insect interactions at two very different scales of time and space. Much of insect ecology either deals directly with plant-insect interactions, or else examines other aspects of insect ecology within the context of host plants. While some insects are carnivores, parasites, or detritivores, and thus live two or more steps removed from primary production, an enormous number of insects are herbivores, pollinators, predators of herbivores, or parasitoids of herbivores. For these insects, plants form the framework in which they exist. An understanding of insects often cannot readily be separated from an understanding of them within the context of plants. Similarly, the issues of scale, both temporal and spatial, are critical to all ecological investigations. In fact, much of the fascinating complexity of the world cannot be understood without attending to features within each nested level of spatial and temporal scale. Levin, in his seminal paper on pattern and scale in ecology (Levin 1992), called scale "the fundamental conceptual problem in ecology, if not in all of science." Not only are some ecological processes explicitly framed by questions of scale, for example disturbance ecology (Paine and Levin 1981), many processes driven by scale dependent interactions, for example community assembly (Emerson and Gillespie 2008, Fahimipour and Hein 2014), collective behavior (Gordon 2014), and ecosystem-level nutrient and energy budgets (Liu et al. 2009). Answering the ecological questions most relevant today depends on our ability to bridge multiple scales of space and time. Fortunately, current methods in ecology, from molecular biology to environmental sensing, allow us to incorporate information from these multiple scales into our understanding of systems and system interactions better than ever before (Chave 2013).

The study of plant-insect interactions incorporates processes acting across the range of temporal and spatial scales. A full understanding of insect herbivory and plant defenses against it integrates information from the nearly instantaneous molecular level, to short-term local interactions between individuals, to generations-long population and community dynamics occurring at the regional scale, to long-term evolutionary processes playing out across continental scales. In my dissertation, I begin at the local, immediate scale, using greenhouse experiments to look at direct plant defenses against herbivores. I then shift to a continent-wide geographic perspective in evolutionary time, using molecular techniques to investigate the invasion history of an insect in the context of a host-plant shift.

In Chapter One, I investigate the impact of smoke exposure on a direct plant-insect interaction. I use *Nicotiana attenuata* and its herbivore *Manduca sexta*, a model system in plant defense against herbivory, to ask if smoke affects plant-insect interactions independent from any direct impacts of fire on population numbers of spatial distribution. I use a variety of bioassays to measure the defense status of *N. attenuata* after exposure to smoke, and its effect on *M. sexta*. This hypothesis draws upon the molecular-level complexity of plant defenses to propose an individual-level response to a highly variable abiotic stressor.

In Chapter Two I use reduced whole genome sequencing to examine geographic patterns of genomic diversity in the Colorado potato beetle, *Leptinotarsa decemlineata*, in Mexico and the USA. The Colorado potato beetle has a dramatic invasion history, at the center of which is a critical host-shift that caused it to become a major agricultural pest. I use population genetics and discriminant analysis of principal components to describe the geographic and genomic diversity of samples collected from across the proposed native range and from introduced populations. I then conduct a phylogenomic analysis of these data in Chapter Three. Using these analyses, I evaluate the possibility of multiple introductions and propose a geographic region from which introduced beetles most likely originated. These chapters thus use molecular data to investigate landscape level patterns that have been playing out over hundreds of years.

The three chapters in this dissertation deal with insect ecology from different perspectives of temporal and spatial scale, but in all cases framed within the context of plant-insect interactions through herbivory. This body of work provides a theoretical framework for continued examination of smoke impacts on plant-insect interactions, and resources for future empirical research on Colorado potato beetle genomics in native and introduced populations.

Literature Cited

Chave, J. 2013. The problem of pattern and scale in ecology: what have we learned in 20 years? Ecology Letters 16 Suppl 1:4–16.

Emerson, B. C., and R. G. Gillespie. 2008. Phylogenetic analysis of community assembly and structure over space and time. Trends in ecology & evolution 23:619–30.

Fahimipour, A. K., and A. M. Hein. 2014. The dynamics of assembling food webs. Ecology Letters 17:606–13.

Gordon, D. M. 2014. The ecology of collective behavior. PLoS biology 12:e1001805.

Levin, S. A. 1992. The problem of pattern and scale in ecology. Ecology 73:1943–1967.

Liu, F., X. Ben Wu, E. Bai, T. W. Boutton, and S. R. Archer. 2009. Spatial scaling of ecosystem C and N in a subtropical savanna landscape. Global Change Biology 16:2213–2223.

Paine, R., and S. Levin. 1981. Intertidal landscapes: disturbance and the dynamics of pattern. Ecological Monographs 51:145–178.

Chapter One

Impacts of smoke exposure on plant-insect interactions

1.1 Introduction

Induced plant defenses against herbivory

Plants respond to insect herbivore attack through the induction of a diverse array of defense mechanisms. Inducible defenses can be either direct or indirect. Direct induced defenses include stimulation of defenses, either chemical or physical in nature, that directly influence herbivore growth, development, reproduction, or feeding patterns. Changes in a wide range of bioactive plant compounds have been documented following wounding, including cyanogenic and iridoid glucosides, glucosinolates, flavonoids, quinons, alkaloids, phenolics, tannins and proteinase inhibitors (reviewed by Edwards and Wratten 1983, Chen 2008). Responses to herbivory range in scale from local to systemic, and rates of induction vary from long-term to nearly instantaneous (reviewed by Chen 2008, Fürstenberg-Hägg et al. 2013). Physical characteristics that protect plants from herbivory, including leaf toughness, trichomes and waxes can also change in response to herbivory (Agrawal 1999, Tian et al. 2012, Frederickson et al. 2013, reviewed by Fürstenberg-Hägg et al. 2013). Indirect defenses are those that act by recruiting natural enemies of herbivores. Indirect defenses include both the provision of resources such as food or shelter, and the provision of information (Kessler and Heil 2011) through the production of C6 green leaf volatiles and herbivore-induced volatile organic compounds (HIVOCs) (Thaler 1999, Heil and Bueno 2007; reviewed by Dicke et al. 2003, Arimura et al. 2005, Kessler and Heil 2011), primarily terpenes (Holopainen 2004).

Various stimuli induce plant defense responses, and different stimuli provoke responses with varying specificity. Herbivores directly induce many responses. Oral secretions and regurgitant from various herbivores have been shown to elicit or modify plant herbivore-defense responses (reviewed by Howe and Jander 2008, Fürstenberg-Hägg et al. 2013). Oviposition secretions and insect eggs can produce a range of responses in plants (Hilker and Meiners 2011), both eliciting defenses (Kim et al. 2012, Beyaert et al. 2012, Geiselhardt et al. 2013), and suppressing them (Bruessow et al. 2010). Even the damage caused by the footsteps of insects, both from caterpillar crochettes digging into leaf tissue and from breakage of trichomes has been demonstrated to stimulate defenses (Bown et al. 2002, Peiffer et al. 2009). Mechanical wounding often produces a reduced or partial defense response (Baldwin 1988, Khan and Harborne 1991), but mechanical wounding that mimics the extent and duration of caterpillar feeding has been shown to elicit VOC emission from lima bean without the addition of herbivore-specific compounds (Mithöfer et al. 2005). Some plants respond differently to different herbivores, discriminating between attacks by closely related herbivore species (De Moraes et al. 1998), or even between different developmental stages of the same insect (Takabayashi et al. 1995).

Insect attack itself is not always required for the induction or priming of defenses to occur. Not only are HIVOCs an important component of induced defenses for their role in arresting carnivores at insect infested plants, VOCs from plants under herbivore attack also play an important role in inducing defenses in neighboring plants (Arimura et al. 2000, Karban et al. 2000) or different parts of the same plant (Frost et al. 2007). In many cases, exposure to HIVOCs primes these defenses rather than fully inducing them, so plants respond more quickly or with greater intensity when subsequently challenged by herbivores (reviewed by Frost et al. 2008, Kim and Felton 2013).

Abiotic environmental stressors can also trigger induced defenses. Exposure to solar ultraviolet-B radiation (UVB) typically reduces herbivory or reduces insect growth (Caldwell et al. 2003). Additionally, much of this effect appears to be due to overlap in the response to herbivory and UVB (reviewed in (Caldwell et al. 2007). For example, UVB exposure reduced herbivory on *Arabidopsis* in the field, and made plants less attractive to diamondback moths for oviposition, an effect that was erased in plants with a mutation causing jasmonic acid-insensitivity (Caputo et al. 2006). *Nicotiana attenuata* and *N. longiflora* both produce responses to UVB exposure that are similar to plant responses to herbivory in terms of phenolics production and transcriptional activity (Izaguirre et al. 2003, 2007). Ozone is also very damaging to plant cells, and exposure to high doses can cause production of biogenic VOCs (BVOCs) and changes in defense signaling pathways (Schraudner et al. 1993, Eckey-Kaltenbach et al. 1994, Heiden et al. 1999, Rao et al. 2000, Vuorinen et al. 2004). Blande and co-workers found that, while high ozone concentrations degraded several HIVOCs from lima bean and reduced the active space of volatile signals, extra-floral nectar production was increased by direct ozone exposure (Blande et al. 2010). Other environmental conditions, such as water stress and nutrient availability can influence quality or intensity of induced defense (Takabayashi et al. 1994, Lou and Baldwin 2004). There also are often complicated interactions between plant defense responses. Responses to different herbivores, pathogen resistance and abiotic stress may interact either due to tradeoffs in prioritizing defenses, or due to interference between different defense systems. Additionally, plant response to multiple stressors differs from response to individual stressors, often in a non-additive fashion (Atkinson and Urwin 2012).

Physiological effects of smoke on plants

One important and understudied abiotic stressor is smoke. Natural smoke from wildland fire is a potential trigger of induced defenses that has been entirely overlooked. Not only is there is an impressive diversity of airborne chemicals to which plants are known to respond (De Moraes et al. 2004), but also natural smoke contains a diverse array of compounds, some of which are known to have physiological effects on plants. As fire regimes change, there is a possibility that altered smoke regimes in the environment will affect plant-insect dynamics.

The presence of smoke in the environment is highly variable between different ecosystems, biomes, and seasons. In some environments smoke is rare or sporadic, while other environments are exposed to smoke on a frequent basis. Local exposure to natural smoke can be decoupled from local exposure to fire, due to the often heterogeneous nature of both fire burn patterns and landscape vegetation matrices. Even a plant community with very long fire-return intervals may experience frequent smoke exposure, if prevailing weather patterns regularly transport smoke from nearby communities with shorter return intervals.

Human modification of the landscape and fire regimes has likely altered the natural pattern of exposure to smoke in the environment. Native Americans often used fire extensively to manage the landscape for wildlife and edible plants. In more recent times, fire suppression efforts may have reduced the frequency of smoke exposure in some regions. If smoke has an ecological effect on plant-insect interactions, these suppression efforts may have altered more than just fuel loads and stand characteristics. In contrast, increased population density is often associated with increased fire risk (Sofiev 2013). In regions where climate change is likely to result in hotter, more arid climates, changing fire regimes will change smoke regimes yet again (McKenzie et al. 2009). Any ecological impacts of smoke will increase under conditions of increased smoke exposure.

The chemical composition of natural smoke is highly variable, and dependent on both fuel type and fire characteristics (Ward and Hardy 1991, Sandberg et al. 2002, Radojevic 2003, Schultz et al. 2008). It contains a complex mixture of products, including a diverse array of air pollutants and greenhouse gasses (Andreae and Merlet 2001, Urbanski et al. 2009). Not only do the amounts of oxidized and reduced species vary with characteristics of the combustion process (Schultz et al. 2008), different fuels release numerous different organic aerosols and VOCs into the atmosphere (Bell and Adams 2009, Alves et al. 2011, Evtyugina et al. 2013).

Much of the research on the impact of air pollution on vegetation has focused on the impacts of industrial smoke. Industrial smoke has long been known to have impacts on plant health. Sulfur dioxide, nitrogen oxides, and ozone have all been examined as industrial air pollutants with impacts on plant health (Bytnerowicz 1996, Agrawal and Agrawal 1999). All these compounds are present in various amounts in wildfire smoke as well (Ward and Hardy 1991, Davies and Unam 1999). Recently, however, the contributions of biomass burning, and their effects on plants, have begun to receive attention. This new work has been recently reviewed by Bell and Adams (2009).

Due to the chemical complexity of smoke, there is great potential for smoke to contain compounds that impact plant physiology (Bell and Adams 2009), including responses to herbivory. Ecological effects of smoke have not been entirely overlooked. Smoke plays an important ecological role in fire-prone ecosystems by directly stimulating the germination of many fire-following plants (Dixon et al. 1995, Brown and van Staden 1997, Keeley and Fotheringham 1998, Ghebrehiwot and Kulkarni 2009). There is now extensive literature on smoke-induced seed germination. Particularly important in this interaction is a butenolide isolated from plant-derived smoke (van Staden et al. 2004). Exposure to this butenolide stimulates germination and seedling vigor, and enhances seedling root and shoot elongation. These effects have been shown in a wide range of seeds, from specifically fire-following plants to important crop plants and weeds from non fire-prone environments (Sparg et al. 2006, Ma et al. 2006, van Staden et al. 2006, Jain and van Staden 2006, Daws et al. 2006, Ghebrehiwot et al. 2008). While science has recognized this critical ecological role of smoke in the environment, with a few exceptions the broader diversity of potential impacts of smoke has gone largely unexamined.

In addition to research on the germination effects of smoke, there is a very small pool of research on other physiological effects of smoke on plants. Several studies have now investigated the impact of smoke exposure on photosynthesis. A decrease in photosynthesis is part of the classic response to both abiotic and biotic stress, and has been documented in the few direct tests of plant physiological response to smoke exposure. Particularly, temporary decreases in photosynthesis rates after smoke fumigation were seen in a variety of plants including the South

African shrub *Chrysanthemoides monilifera* (Gilbert and Ripley 2002), five North American tree species (Calder et al. 2010), some varieties of grapevine (Bell et al. 2013) and *N. tabacum* (Cowan 2010). In another example of physiological effects of smoke, reduced smoke exposure was implicated in the catastrophic loss of the Floridian tree *Torreya taxifolia* to disease, either through direct smoke toxicity to pathogens or through stimulation of plant resistance (Schwartz et al. 1995). Smoke-induced disease resistance is also being pursued as a mechanism to explain spatial variation in patterns of infection by the sudden oak death pathogen (*Phytophthora ramorum*) in California (Moritz et al. unpublished).

Potential impacts on plant-insect interactions

Given the sensitivity of induced defenses to abiotic stressors and chemical signals, the chemical complexity of smoke, and the known interactions of smoke with other aspects of plant physiology and growth, we hypothesized that smoke would impact plant defenses against herbivores. Impacts of smoke exposure on plant defense responses may be non-adaptive or adaptive. One possibility is that exposure to one or more of the chemicals in smoke may cause non-adaptive interference with the highly chemically sensitive process of defense induction. Alternatively, a stress response to smoke could interfere with herbivore defenses, for example either as a secondary response due to shared effects of stressors or by inducing a resource allocation tradeoff between competing stressors. Finally, plant responses to natural smoke may be adaptive if smoke is a reliable signal of imminent changes in herbivory.

Post-fire changes in herbivory have been documented in a number of different ecosystems. A diversity of insect responses to fire were reviewed by Swengel (2001). Changes in insect populations tend to be species-specific, or at least guild-specific, with responses occurring at various temporal and spatial scales. Insect community responses are often linked, in addition to direct fire-induced mortality, to changes in host plant quality, abundance, and apparency, as well as to changes in habitat structure. Immediate post-fire reductions in insect abundance are typical, although some species are specifically attracted to fire or smoke. Intermediate effects are more variable, with increases in populations not uncommon (Swengel 2001). In a Florida sandhill longleaf pine ecosystem, fire created a spatial gradient in insect abundance and herbivory, both of which were higher at the edge than in the interior of burned areas in the immediate post-burn season (Knight and Holt 2005). In a desert prairie system, forb-feeding grasshoppers increased dramatically in burned areas after a low-intensity fire, while other species suffered short-term declines (Parmenter et al. 2011). Significantly higher herbivory was observed in burned versus unburned trees in a neotropical savannah (Lopes and Vasconcelos 2011). The impact of herbivory on plants in the post-fire environment is also highly variable, ranging from temporary herbivore release to increased mortality due to insect feeding on fire-damaged plants (McCullough et al. 1998).

While much has been learned in the past few years about the regulation of plant responses to environmental stressors, there is still much that we do not know regarding the diversity of signals to which plants respond and the maintenance of specificity in these signaling pathways. The impact of natural smoke on these systems remains entirely unexplored. If smoke affects mature plants, either as an environmental stressor like ozone or UVB radiation, or as a chemical signal like herbivore induced VOCs, the potential influence of smoke in the environment is both significant and largely unrecognized.

A system for examining smoke ecology

One strategy for exploring this question is through examination of a system in which induced defenses have been documented, and an evolutionary history of plant exposure to smoke is realistic. The fire-adapted wild tobacco provides just such a system. Wild tobacco, *Nicotiana attenuata* Torr. ex Watson (Solanaceae), is an annual plant native to the cold deserts of the intermontane western United States, predominately in the Great Basin region of California, Nevada and Utah (Goodspeed 1954, Wells 1959). *N. attenuata* is an early-successional ruderal plant found within the plant communities of pinyon-juniper woodland, sagebrush scrub and occasionally ponderosa pine forest (Wells 1959). While isolated populations in dry washes and along disturbed roadsides are sometimes persistent (>3yrs), populations of *N. attenuata* primarily occur ephemerally after fire in sagebrush and pinyon-juniper ecosystems, where they typically persist for no more than three years (Barney and Frischknecht 1974, Young and Evans 1978, Koniak 1985, Krock et al. 2002, Bahulikar et al. 2004). *N. attenuata* flourishes in the reduced competition and nutrient rich soils of post-fire conditions. Exposure to chemicals in plant-derived smoke, along with the elimination of leaf litter from which germination inhibitors leach, synchronizes germination into this favorable environment (Baldwin et al. 1994, Preston et al. 2002, Krock et al. 2002, Schwachtje and Baldwin 2004).

Since *N. attenuata* seeds respond to smoke, it is also possible that *N. attenuata* plants are sensitive to biologically active compounds in plant-derived smoke and provide a mechanism for fire to have indirect ecological effects on plant-insect interactions. Because of the high mobility of smoke and the patchy, mosaic burn patterns that are common in Great Basin habitats, a population of *N. attenuata* growing in a fire scar from a previous year may be exposed to smoke from a neighboring fire in the current season without suffering direct effects of the fire.

Nicotiana attenuata has been well studied in the context of plant-insect interactions, including inducible defenses and volatile signaling. Early work by Karban and others (Karban et al. 2000, 2003, Karban 2001, Karban and Maron 2002) suggested that *N. attenuata* exposed to VOCs from experimentally damaged sagebrush *Artemisia tridentata* experienced less herbivory than unexposed *N. attenuata*. It was later demonstrated that, rather than direct induction of defenses, exposure to sagebrush VOCs most likely primes the insect defense responses of *N. attenuata* (Kessler et al. 2006). Exposure of *N. attenuata* and *N. longiflora* to ultraviolet B radiation have been shown to stimulate insect defense mechanisms (Izaguirre et al. 2003, 2007). Particularly interesting, given that ozone is often formed in smoke, ozone exposure also was shown to stimulate insect defenses in *N. tabacum* (Heiden et al. 1999) and extra-floral nectar production in lima beans (Blande et al. 2010). Combined with the known importance of smoke in affecting germination physiology of *N. attenuata*, the sensitivity of this species to stimulation of induced defenses by biogenic VOCs and abiotic stressors makes it seem an ideal candidate for the study of ecological effects of smoke exposure on plant-insect interactions.

Plant defenses are known to respond to airborne compounds including VOCs and O_3, both of which are present in smoke along with a diversity of other compounds. Additionally, herbivory has been shown to change in the post-fire environment. Consequently, both informative and disruptive influences of smoke on plant defenses against herbivores are possible. Therefore, we conducted a series of experiments to test the hypothesis that plant exposure to smoke would alter insect herbivore success through changes in direct defenses. We conducted larval growth assays with *Manduca sexta* reared on smoke exposed or untreated control plants to look for changes in growth and development of insects to plant fumigation, and choice tests to

determine if *M. sexta* exhibited behavioral responses. We also collected pilot data on diurnal emissions of biogenic VOCs stimulated by smoke exposure.

1.2 Methods

Plants

Second-generation inbred *N. attenuata* seeds (original seed from NCSU Dept. of Crop Research germplasm facility) were soaked for one hour in 5 ml of 1:50 liquid smoke (Wright's) dilution supplemented with 50 µl of 0.1 M gibberellic acid (GA_3) (Krügel et al. 2002). Seeds were then rinsed 3 times with deionized water and transferred to a petri dish lined with damp filter paper. Food-grade liquid smoke, commonly used to germinate smoke-sensitive seeds, is produced by passing smoke from burning wood chips through a condenser and diluting it with water, and thus contains the majority of water soluble compounds that would have been found in the originating smoke vapor (Shallwani 2009). Petri dishes were sealed with Parafilm to maintain humidity, and incubated in a growth chamber at 26 °C / 16 hours light, 24 °C / 8 hours dark until germination. After 5-7 days, germinated seeds were individually transferred to 5cm pots containing a blend of Sunshine potting soil, coarse sand and perlite, and placed in a greenhouse (21-27 °C, 16 L 8 D) to establish. Established seedlings with 3-5 leaves were subsequently transplanted to 15 cm pots of potting soil, and grown for 1-2 months until they reached large rosette size. For choice tests, seeds were soaked for 1 hour in 5 ml DI water with 50 µl of 0.1 M GA_3, then rinsed and scatter-sown in a flat of vermiculite watered with dilute smoke solution (10% Liquid Smoke) to germinate. These seedlings were directly transplanted to 10 cm pots and grown to small rosettes.

Insects

Eggs of *Manduca sexta* L. (Lepidoptera: Sphingidae) were obtained from Carolina Biological Supply Company (Burlington, N.C.), and maintained in a growth room in 500 ml polyethylene containers lined with paper until they hatched. As larvae began to hatch, arenas were provided with artificial diet (Carolina Biological) that was raised above the paper on wire-mesh stands.

Smoke exposure treatment

Bigleaf sagebrush (*Artemisia tridentata*) litter was used as the fuel source for smoke production due to the prevalence of this plant in the sagebrush scrub and pinyon-juniper habitats in which *N. attenuata* typically grows. Sage was purchased as a smudge stick, and leaves and small stems were chopped to a uniform coarse litter. Plants in the treatment group were enclosed in a glass chamber 40 x 40 x 76 cm and fumigated for 20 minutes with sage smoke. Combustion occurred in a glass crucible outside the fumigation chamber, with smoke drawn into the chamber by vacuum pump, via 6.35 mm OD (outside diameter) stainless steel tubing that first passed through a cold water bath. A 6.4 cm computer fan on the floor of the chamber was run for approximately 10 seconds every two minutes to provide occasional mixing of smoke inside the chamber.

Bioassays

Insect growth

To test if host plant exposure to smoke affected *M. sexta* larval growth, we monitored the mass of larvae reared on smoke exposed or untreated plants repeatedly for 14 days in four replicated studies. In all studies, each individual larvae was confined to an individual plant for the duration of the study, with samples replicated as one plant plus one larvae. Trials were initiated with 10-20 samples in each treatment. A total of five control and nine treatment samples were lost during the course of the trials due to death of *M. sexta* from viral infection, larval abandonment subsequent to complete defoliation of host plant, and occasional escape of larvae from apparently healthy host plants. Samples with incomplete data were excluded from analysis, leaving a total of 50 control and 46 treatment samples for analysis (Trial 1: N= 9 control, 8 smoke; Trial 2: N= 13, 13; Trial 3: N= 10, 9; Trial 4: N= 18, 16). An additional trial was conducted where larvae were weighed only on days 0, 7, and 14 (N= 12 control, 13 smoke).

For these insect growth trials, *N. attenuata* plants in 15 cm pots were grown until mid-size rosettes began bolting, whereupon the central stem of all plants was pruned to a few inches height (typically between 2nd and 3rd stem leaves) and side branches were allowed to grow until sufficient leaf density was present to sustain two weeks of larval feeding. We presumed that consistency in plant manipulation would cause any potential impacts of pruning on defenses to be distributed equally across plants. A test of this assumption is presented in the following section. To reduce the potential for pruning effects to obscure treatment effects, we required that a minimum of one week pass between any pruning and initiation of an experiment. After smoke exposure, treated and untreated plants were randomly located in a grid on the greenhouse bench in individual saucers. Plants were watered from below via these saucers for the duration of the experiment. Acetate sleeves 15-25.5 cm height and 15 cm diameter were placed in pots to contain *M. sexta* to their assigned plant.

Depending on the larvae available when plants were large enough to start a trial, late first instar, second instar, or new 3rd instar larvae were used for the trial. All larvae within a trial hatched within 2.5 days of each other. For each trial, larvae were randomly selected from the available and apparently healthy specimens and weighed on a Metler scale. Larvae which were more than one standard deviation from the mean mass for that trial were exchanged for alternates. Larvae were placed into individual portion cups for transport to the greenhouse, randomly allocated to control or treatment plants, and transferred from portion cups to leaves with a fine paintbrush by late afternoon on the day of plant fumigations. Larvae were weighed every two to three days. This involved collecting larvae in portion cups, transporting them to the lab to be weighed and returning them to their plants. Plants were checked periodically for the presence of larvae and removed from the study if no larvae were found. Fourteen days of growth produced larvae ranging from 3rd to 5th (final) instars.

Pruning test

We conducted a two way factorial trial to test for interactions between pruning and smoke exposure. We set up the trial with 10 replicates in each factorial combination, and lost six samples across the trial (N= 7, no smoke + unpruned; 9, no smoke + pruned; 9, smoke + unpruned; 9, smoke + pruned). Plants were treated as described above, allowing a gap of one week between pruning of central stems and smoke exposure. Larvae were weighed on days 0, 7, and 14. Analysis of variance on this trial showed no effect of smoke exposure and no interaction

between smoke and pruning, but a significant reduction of larval mass due to pruning ($p= 8.69x10^{-05}$) (Figure 1.1). Since we found no interaction between pruning and smoke exposure, we were justified in presuming that pruning effects would be consistent across treatments.

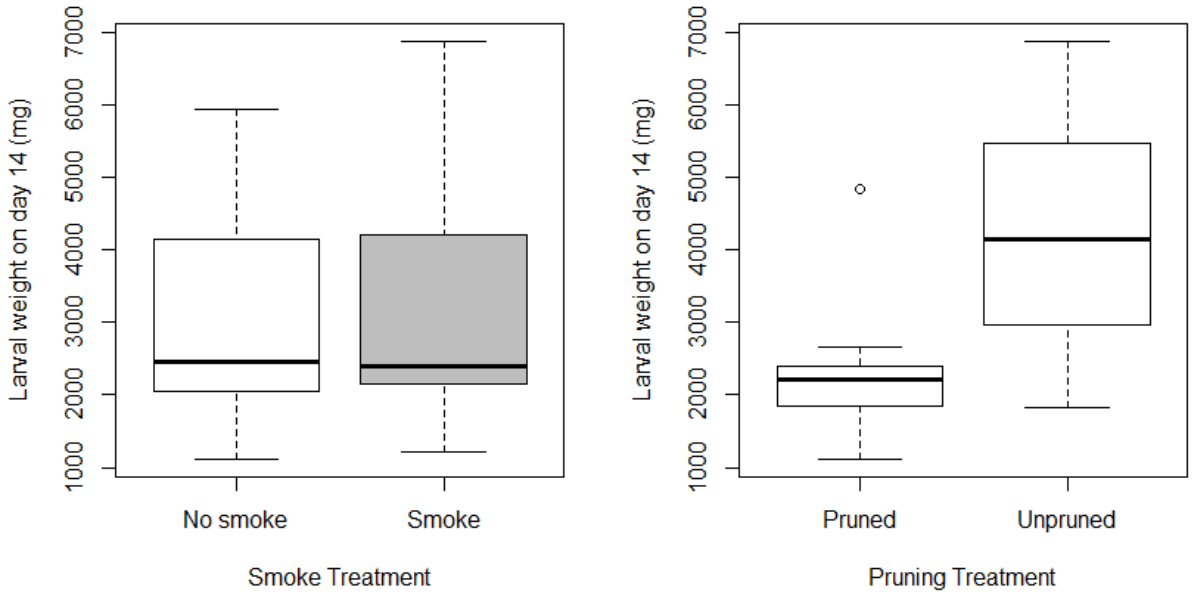

Figure 1.1: Test for interaction between plant pruning and smoke exposure.

Rate of development trial

Because the larvae seemed to have highly variable growth rates, we did not want to discount the possibility that the intrinsic variability of the larvae was masking a small but real effect of smoke exposure. In an effort to reduce the variation and clarify any potential differences between treatments, another trial was conducted to determine if smoke exposure of host plant affected the time spent at different developmental stages. In this trial we tightly controlled the initial age of the larvae. Neonate larvae (2-6hrs old) were placed individually on random smoke exposed or untreated plants. Larvae at this age weigh very little, and initial mass differences were assumed to be negligible. Larvae were observed daily for developmental markers including active feeding, head-capsule slippage, pre-molt stasis and ecdysis. To minimize any potential impacts due to repeated handling, larvae were observed but left undisturbed until day 7, when they were removed from plants, weighed and returned to plants. Surviving larvae were collected for a final weighing on day 14 then freeze killed and dry-weighed a week later (N = 19 control plant + larvae, 19 treatment). For this experiment, large rosette plants just initiating elongation of the central stem were used. Plants assigned to the treatment group were exposed to 20 minutes of smoke fumigation. After fumigation, a sheet of filter paper was placed over the soil surface of each 15cm pot, and cut to accommodate the plant. Control and treatment plants were randomly placed in a grid of saucers on greenhouse benches, and acetate collars around plants prevented *M. sexta* from migrating between plants.

Choice test

A choice test was conducted to determine if *M. sexta* had a preference for smoke exposed or untreated plants. Small rosette stage plants in 10 cm pots were used for choice tests. After fumigation, treated plants were size-matched with untreated plants. One leaf on each plant of approximately the same size and position was selected, and the two blades were enclosed together in a plastic clip-cage with a 75 mm diameter. Clip cages had solid bottoms and sides, and open tops. Plant pairs were randomly arrayed on a bench in the greenhouse. Digital photographs of choice arenas (Fujifilm Finepix Z) document the original size and condition of all leaf pairs. Ten first-instar larvae aged 24-30 hrs were selected and placed into clip-cages on the arena bottom, equidistant from each leaf. Freely moving larvae were continuously observed for the initial 1.5 hrs until all larvae had located and established on a leaf. Bioassays were photographed again after two days and after four days of feeding.

Biogenic Volatile Organic Compound Emissions

In addition to the bioassays with *M. sexta*, we conducted a small pilot study to investigate possible changes in VOC emissions after smoke exposure. Immediately after *N. attenuata* exposure to smoke, a mid-sized rosette stage plant was enclosed in a dynamic enclosure, designed in such a way to monitor BVOC emissions, plant physiology (transpiration, photosynthesis) and chamber environmental parameters including temperature, photosynthetically active radiation (PAR) and relative humidity. Another identical chamber enclosed a non smoke-exposed plant simultaneously. Each pair of plants was monitored over one day, and measurements were conducted over 5 days with 5 different plant pairs. Details on the apparatus and the analytical systems used to quantify and qualify the emissions are described in Ormeño et al. (2010) and Fares et al. (2010). Terpene emissions were analyzed by means of Proton-Transfer-Reaction Mass Spectrometry (PTR-MS) and Gas-Chromatography Mass Spectrometry (GC-MS). Isoprene (C_5H_{12}) is the terpene most released by plants (Guenther et al. 1995), followed by a variety of monoterpenes ($C_{10}H_{16}$) and sesquiterpenes ($C_{15}H_{24}$).

Statistical Analysis

All statistical analyses were conducted using the *R* statistical software system. To test for differences in larval growth between smoke-exposed and untreated plants, we used a linear mixed effects model. Treatment was modeled as a fixed effect, and blocking by trial was treated as a random effect. Measurement days across trials were treated as repeated measures, also with random effects. Box-Cox transformation of larval mass indicated the log transformation was appropriate for our data. We also conducted an ANCOVA on final mass across all five growth trials, with treatment and trial as categorical variables and initial mass as a continuous variable.

We used the built-in power analysis for t-tests in *R* to conduct a power analysis on the difference between final masses of larvae reared on treated and untreated plants. We used the data from these trials to estimate effect size and standard deviation for the power analysis.

We used several test to analyze the rate of development trial. We conducted an analysis of variance for the final dry mass fitted against treatment, and against treatment plus developmental stage at various days. To compare age structure of both populations, we examined the frequency of larvae in different developmental stages for all dates. We used the Wilcoxon rank-sum test as a failure-time analysis of the time to reach two different development stages.

We also used the built-in *R* power-t-test to conduct a power analysis for this data of the difference between the final masses of larvae reared on treated and untreated plants.

For larval choice-tests, we used *Adobe Photoshop* to scale arena images to a uniform size, extract leaves from image background, and individually re-scale leaf images for growth-corrected leaf-area analysis. We used *Assess* 2.0 image analysis software (Lamari 2002) to measure leaf area, both with and without growth correction, in order to determine the amount of leaf area consumed. We conducted paired t-tests comparing the amounts of leaf area consumed.

1.3 Results

Larval growth

Plant exposure to smoke did not induce significant differences in *M. sexta* growth. LME using repeated-measures blocked across the four growth trials suggested a non-significant 0.076 reduction in log mass due to smoke exposure (se=0.0467, df=102, p=0.106) (Figure 1.2). While there were no significant differences between treatments, final mass for larvae reared on smoke-exposed plants was consistently slightly less than for those from control plants (Figure 1.3). To conduct a power analysis for this data, we used the by-trial average reduction in final mass due to smoke treatment (297 mg) as an estimate of effect size, and the average of the standard deviations from the five trials (946) to estimate the standard deviation. Power analysis for a one-sided t-test indicated that, in the case of a true effect, a minimum of 127 samples in each treatment would be needed to have an 80% chance of correctly rejecting the null hypothesis. In an ANCOVA on final mass across all five growth trials, with treatment and trial as categorical variables and initial mass as a continuous variable, exposure to smoke had virtually no effect (-164 mg, p=0.5) (Figure 1.4). Initial mass and trial had the most effect on final larval mass.

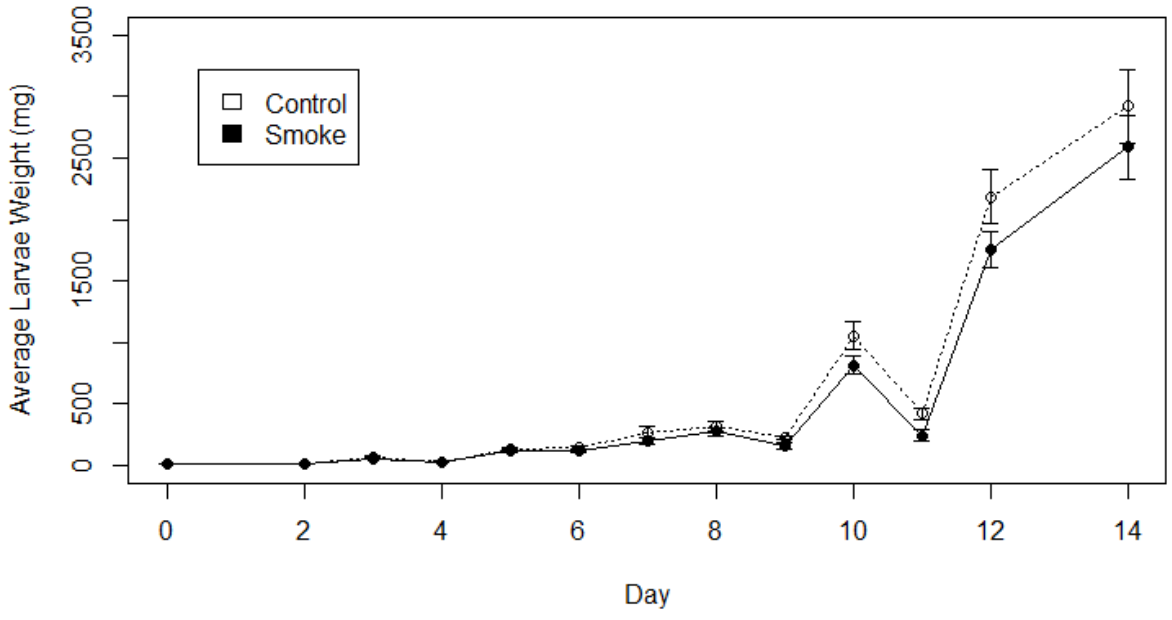

Figure 1.2: Larval growth was not significantly different between smoke exposed and control plants across four trials.

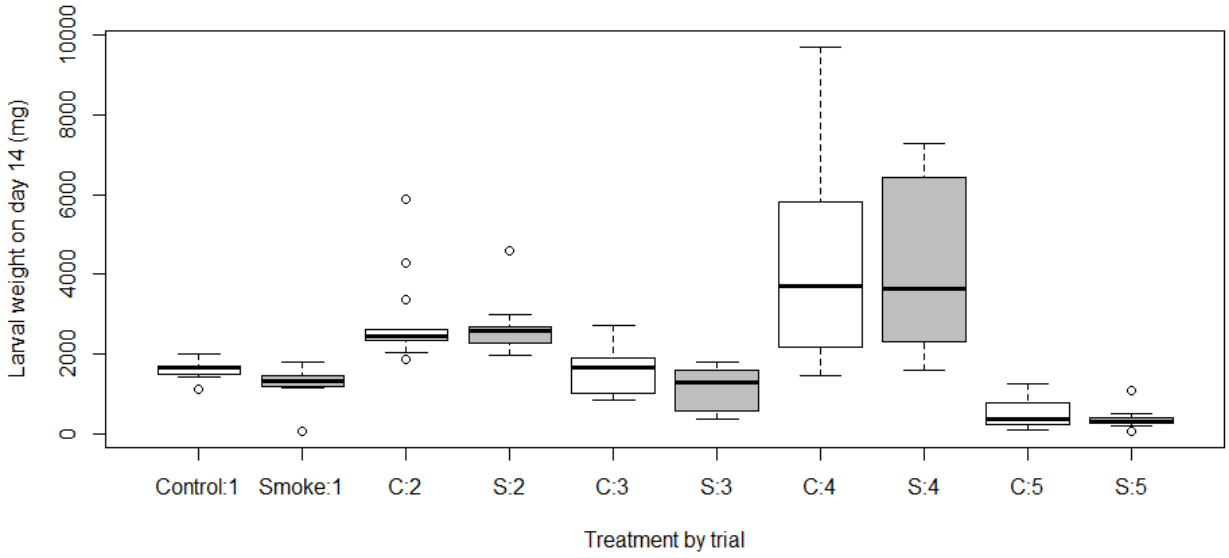

Figure 1.3: Final larval mass for all growth trials. Sample sizes ranged from 8 to 18 larvae across trials (Trial 1: N= 9 control, 8 smoke; Trial 2: N= 13, 13; Trial 3: N= 10, 9; Trial 4: N= 18, 16; Trial 5: N= 12, 13).

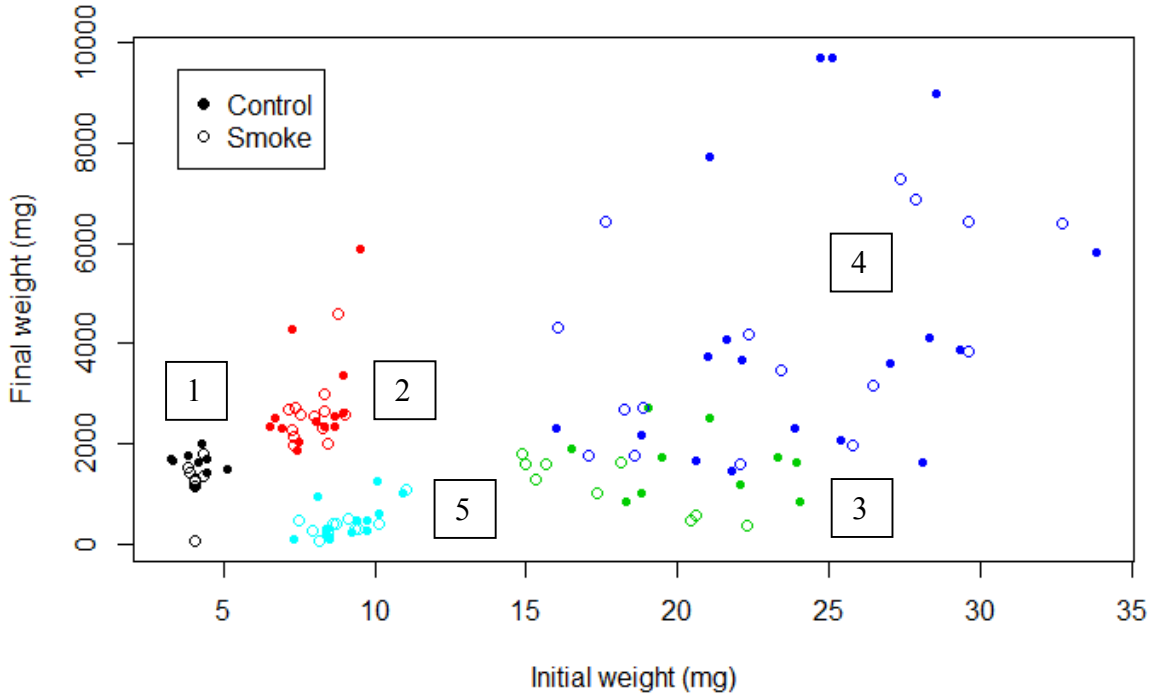

Figure 1.4: Final mass of *M. sexta* was correlated with initial mass and trial, but treatment had no effect. Symbol colors indicate growth trials 1-5.

Rate of development

In this trial, plant treatment had no effect on final dry mass of 14-day-old larvae (ranging from 3rd to 5th instar). Model testing showed that variation in dry mass was best explained by the developmental stage 2 days before the end of the experiment (day 12), with no significant treatment effect (Figures 1.5, 1.6). Failure-time analysis of the number of days taken to reach 4th and 5th instars did not vary with treatment (Wilcoxon rank sum test on time to 4th instar: W = 174.5, p-value* = 0.3535; time to 5th instar: W = 181, p-value* = 0.5912 (*inexact p-value due to ties)). We found no effects of smoke exposure on the population age structure at any day, although the age structure of the control population was slightly ahead of that of larvae on smoke exposed plants. Our power analysis indicated that, even with half the standard deviation of the samples and ten times the effect size, it would take more than 60 samples in each treatment to achieve a power of 80%.

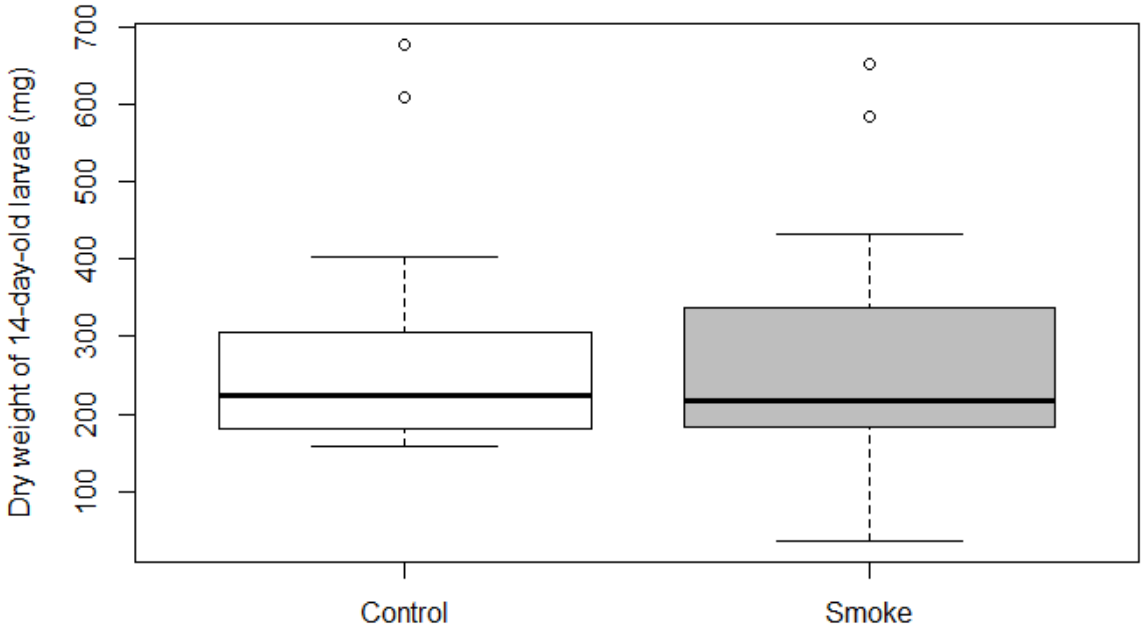

Figure 1.5: Dry mass of 14-day old *M. sexta* did not vary with plant treatment in development rate trial. Sample size was N=19 (control) and N=19 (smoke-reared).

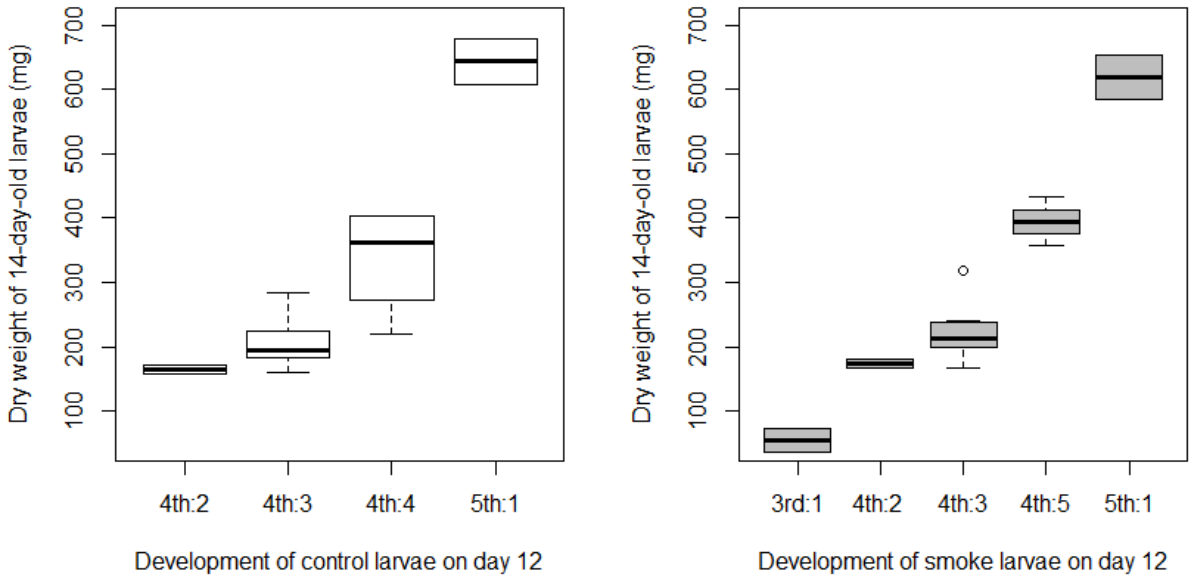

Figure 1.6: Dry mass of *M. sexta* larva was correlated with developmental stage at day 12, on which day larvae exhibited the greatest range of identifiable developmental markers. Development codes indicate instar followed by steps 1-5: 1 = freshly molted; 2 = actively feeding; 3 = pre-molt, head visibly small for body; 4 = early molting with head capsule separating; 5 = actively molting.

Feeding preference

M. sexta larvae showed no feeding preference between treated and control plants (p-value for growth-corrected paired t-test: 0.86, n=10).The percent change in leaf area over 4 days was not consistently different between leaf pairs.

VOC emissions

Preliminary results from our pilot study suggest that terpene emission rates from control and smoke-exposed plants were similar (Figure 1.7), and environmental conditions and plant physiological responses did not differ substantially between the two chambers (Figure 1.2). Results were similar among all pairs of plants, but due to incomplete measurements across the diurnal periods only one representative pair for which all data were available are shown here. Analyses through GC-MS allowed us to determine terpene speciation. Monoterpenes detected in both treatments were sabinene, d3-carene, and limonene. Camphor and γ-terpinene were only observed for smoke-exposed plants. Sesquiterpenes detected were β-longipinene, α-trans-Bergamotene and 2-cis-farnesal. The sesquiterpene aristolochene was only observed for smoke-exposed plants. These results suggest that smoke exposure did not lead to quantitative changes in plant terpenoid metabolism although some induced emissions may have been triggered.

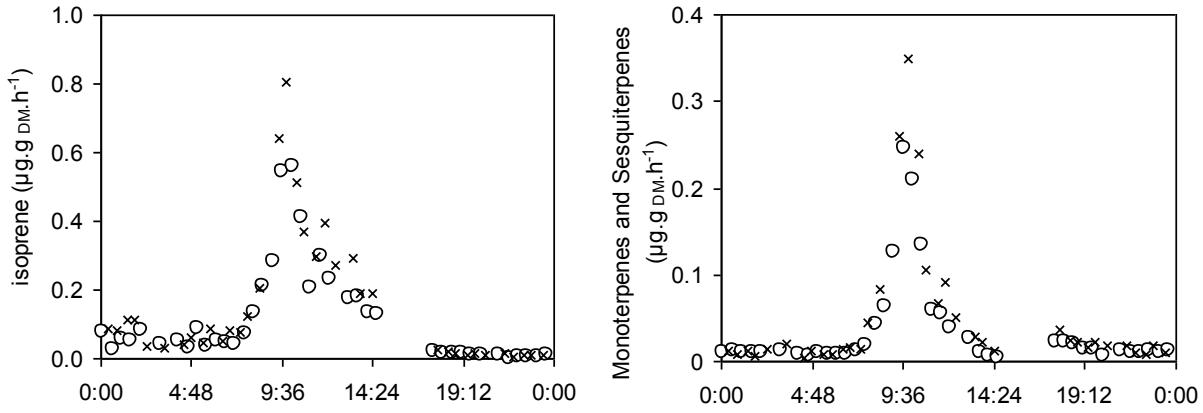

Figure 1.7: Diurnal cycle of terpenoid emission rate from a representative pair of control (o) and smoke-exposed (x) plants, with isoprene compounds shown in the left figure and mono- and sesquiterpenes on the right.

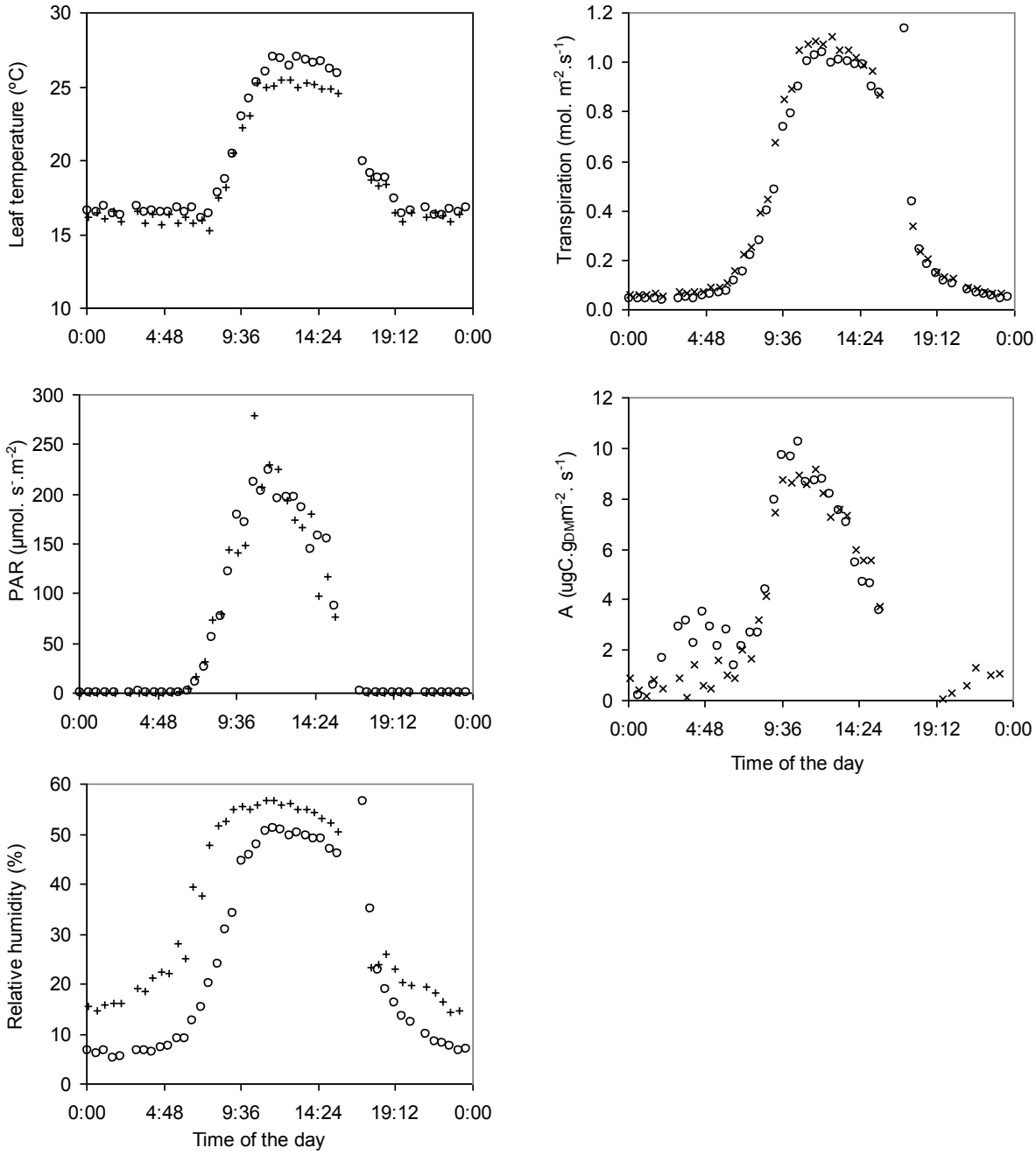

Figure 1.8: Environmental conditions (leaf surface temperature, PAR, relative humidity; left graphs) and plant physiology (transpiration and photosynthesis; right graphs) on one measurement day are shown for one representative pair of control (o) and smoke-exposed plants.

1.4 Discussion

We found no direct effect of host plant exposure to smoke on the herbivore in this system. *Manduca sexta* larvae weights and development were not significantly different between those reared on smoke exposed and untreated *N. attenuata*. Repeated-measures ANOVA with a blocking effect for trial suggested that smoke exposed plants produced larvae with a non-significant reduction in mass compared to untreated *N. attenuata*. However, analysis of final mass with initial larval mass as a covariate clearly shows that initial mass and trial effects, but not treatment effects, influenced *M. sexta* growth. Several factors may have contributed to the importance of trial in this model. For example, seasonal variations in temperature, light or fertilizing regime may have systematically affected plant growth and therefore insect weight. *Manduca sexta* did not evidence a feeding preference between treated and untreated plants. This is consistent with the lack of differences in larval growth. Preliminary analysis of BVOC emissions showed no quantitative differences in emissions due to smoke exposure. This was also consistent with our growth and behavior tests.

Although we did not document any impacts of smoke exposure in this study, and we feel confident that in this case it was not just due to low statistical power for detecting a small ecological effect, we believe that this question is worthy of future investigation. One hypothesis for the lack of effect in our study is that *M. sexta* are specialist herbivores on *N. attenuata* and voracious feeders, and may be indiscriminate as long as the host plant meets their minimum requirements. Effects of smoke exposure may be more apparent if *N. attenuata* defenses were assayed using a less adapted herbivore. Alternatively, it is possible that we did not detect any changes in plant defenses because we were testing at the wrong trophic level. *Nicotiana attenuata* are capable of recognizing damage caused by *M. sexta* from chemical cues present in the larval saliva, and *M. sexta* feeding elicits a specific plant defense response (Halitschke et al. 2001). Rather than increasing production of nicotine as *N. attenuata* does when attacked by non-specialist herbivores, when attacked by *M. sexta* the plant downregulates nicotine production (Kahl et al. 2000, Winz and Baldwin 2001; but see Musser et al. (2002) for a study that suggests herbivore saliva can act offensively to suppress plant defenses) and begins producing volatile chemicals (Kahl et al. 2000) that have been shown to attract predators of *M. sexta* (Kessler and Baldwin 2001). If exposure to smoke had no direct effect on herbivores, but did alter predator response to the plant VOCs, we would not have detected these effects. Additionally, defense response varies within the species (Wu et al. 2008), and this accession may have a less sensitive genotype. While another possibility is that truly no plant defenses were induced or inhibited when *N. attenuata* was exposed to smoke, a broad survey of different plants may yet reveal plants that are sensitive to smoke exposure. Perhaps a longer-lived species, or a plant that re-sprouts after a fire rather than this fire-following (smoke-germinating) one, would show effects that were not seen in *N. attenuata.*

Changes in photosynthesis rates and plant defenses have complicated interactions. Plant defenses are metabolically expensive and generally impose a fitness cost, although this cost is often difficult to measure (Heil and Baldwin 2002). Because plant defenses are energetically costly, it is reasonable to expect that reduced photosynthesis might reduce available defenses. However, a decrease in photosynthesis is most often seen following attack by herbivores, and plants may even actively switch off local photosynthesis during defense initiation (Zangerl et al. 2002, Schwachtje and Baldwin 2008, Bolton 2009). In fact, decreased photosynthesis may play a direct role in defense signaling (Bolton 2009, Kerchev et al. 2012). Therefore, the temporary

reduction in photosynthesis seen in several studies of plant response to smoke exposure is not likely to have inhibited the plant defense response. Rather, any smoke-exposure induced decrease in photosynthesis may have actually enhanced subsequent defense initiation.

Despite not finding an effect of host plant exposure to smoke on herbivore growth in these experiments, we believe the potential for such novel interactions exists and should be examined in other systems. One reason this interaction should be pursued is that, given the sheer diversity of chemical and environmental stimuli to which plants do respond, and the frequency with which such responses impact insect performance, it almost seems implausible that plants would entirely lack a response to the miasma of chemicals in smoke. Whether through a change in growth-defense allocation brought on by a stress response (e.g. ozone stress; CO, CO_2, SO_2 or NO_X toxicity) or by growth stimulation (butenolide, NO fertilization), or a change in induced defense status by stimulation or suppression of defense signaling pathways, there are many potential mechanisms for an interaction between natural smoke and plant defenses.

If, on the other hand, plants are able to discriminate against smoke entirely, how do they maintain such specificity? Induced plant defenses stimulated by insect herbivory are thought to provide a plant fitness advantage in environments where herbivory is not always predictable. Specificity in response to herbivore attack is important if inducibility is to remain a benefit. Yet research has shown ozone and UVB exposure to induce components of plant defenses. Given the chemical complexity of smoke, what mechanisms prevent the misinterpretation of smoke borne compounds as biologically relevant information, or limit smoke interference with important signaling pathways? Current research on cross-talk between hormone signaling pathways regulated by abscisic acid, jasmonic acid, salicylic acid and ethylene are likely to be informative here (Fujita et al. 2006, Atkinson and Urwin 2012).

If smoke does have a yet undetected ecological effect on plant-insect interactions, what would be the implications? Plant defense priming or inhibition by smoke exposure could affect the success of insects escaping from a fire into nearby un-burned but smoke-exposed habitats. While timing of events in such an interaction would be important in the case of an adaptive response, the nature of all three components makes appropriate timing entirely feasible although not consistent across systems. Smoke exposure, for instance, may last for a single day to as much as several weeks or even months (Sandberg et al. 2002, Radojevic 2003, Bell and Adams 2009). Components of plant defense signaling pathways and defenses respond with different speeds, and may develop over seconds, hours or days, and may continue affecting herbivores as much as a year later (Underwood et al. 2005, Arimura et al. 2005, Peiffer et al. 2009). Herbivores may be impacted at any step from host-plant finding and oviposition (De Moraes et al. 2001, Kim et al. 2012, Reisenman et al. 2013) to juvenile development, across generations, and through third trophic level effects as well. Plants growing downwind of fire-prone habitats, either locally or at long range, depending on the dynamics of smoke transport, may be have altered susceptibility to herbivory due to exposure to natural smoke. The application of agricultural pesticides may even be influenced by the intensity of the fire season in distant mountains.

Because it can be difficult to detect small ecological effects, there is risk of accepting a statistical false negative. If we discount the possibility of smoke impacts on plant-insect interactions, we may fail to recognize important shifts in ecological interactions brought on by climatic or anthropogenic alterations to fire regimes.

1.5 Literature Cited

Agrawal, A. A. 1999. Induced responses to herbivory in wild radish: effects on several herbivores and plant fitness. Ecology 80:1713–1723.

Agrawal, S. B., and M. Agrawal (Eds.). 1999. Environmental Pollution and Plant Responses.

Alves, C. A., A. Vicente, C. Monteiro, C. Gonçalves, M. Evtyugina, and C. Pio. 2011. Emission of trace gases and organic components in smoke particles from a wildfire in a mixed-evergreen forest in Portugal. Science of the Total Environment 409:1466–75.

Andreae, M., and P. Merlet. 2001. Emission of trace gases and aerosols from biomass burning. Global Biogeochemical Cycles 15:955–966.

Arimura, G., C. Kost, and W. Boland. 2005. Herbivore-induced, indirect plant defences. Biochimica et Biophysica Acta 1734:91–111.

Arimura, G., R. Ozawa, T. Shimoda, T. Nishioka, W. Boland, and J. Takabayashi. 2000. Herbivory-induced volatiles elicit defence genes in lima bean leaves. Nature 406:512–5.

Atkinson, N. J., and P. E. Urwin. 2012. The interaction of plant biotic and abiotic stresses: from genes to the field. Journal of Experimental Botany 63:3523–43.

Bahulikar, R. A., D. Stanculescu, C. A. Preston, and I. T. Baldwin. 2004. ISSR and AFLP analysis of the temporal and spatial population structure of the post-fire annual, *Nicotiana attenuata*, in SW Utah. BMC Ecology 4:12.

Baldwin, I. T. 1988. The alkaloidal responses of wild tobacco to real and simulated herbivory. Oecologia 77:378–381.

Baldwin, I. T., L. Staszak-Kozinski, and R. Davidson. 1994. Up in smoke: I. Smoke-derived germination cues for postfire annual, *Nicotiana attenuata* Torr. Ex. Watson. Journal of Chemical Ecology 20:2345–71.

Barney, M., and N. Frischknecht. 1974. Vegetation changes following fire in the pinyon-juniper type of west-central Utah. Journal of Range Management 27:91–96.

Bell, T., and M. Adams. 2009. Smoke from wildfires and prescribed burning in Australia: effects on human health and ecosystems. Pages 289–316 *in* A. Bytnerowicz, M. Arbaugh, A. Riebau, and C. Andersen, editors. Developments in Environmental Science, Volume: 8. Elsevier Ltd.

Bell, T., S. Stephens, and M. A. Moritz. 2013. Short-term physiological effects of smoke on grapevine leaves. International Journal of Wildland Fire 22:933–946.

Beyaert, I., D. Köpke, J. Stiller, A. Hammerbacher, K. Yoneya, A. Schmidt, J. Gershenzon, and M. Hilker. 2012. Can insect egg deposition "warn" a plant of future feeding damage by herbivorous larvae? Proceedings of the Royal Society, Biological Sciences 279:101–8.

Blande, J. D., J. K. Holopainen, and T. Li. 2010. Air pollution impedes plant-to-plant communication by volatiles. Ecology Letters 13:1172–81.

Bolton, M. D. 2009. Primary metabolism and plant defense--fuel for the fire. Molecular Plant-Microbe Interactions: MPMI 22:487–97.

Bown, A. W., D. E. Hall, and K. B. MacGregor. 2002. Insect footsteps on leaves stimulate the accumulation of 4-aminobutyrate and can be visualized through increased chlorophyll fluorescence and superoxide. Plant Physiology 129:1430–1434.

Brown, N. A., and J. van Staden. 1997. Smoke as a germination cue: a review. Plant Growth Regulation 22:115–124.

Bruessow, F., C. Gouhier-Darimont, A. Buchala, J.-P. Metraux, and P. Reymond. 2010. Insect eggs suppress plant defence against chewing herbivores. The Plant Journal 62:876–85.

Bytnerowicz, A. 1996. Physiological aspects of air pollution stress in forests. Phyton 36:15–22.

Calder, W. J., G. Lifferth, M. A. Moritz, and S. B. St. Clair. 2010. Physiological effects of smoke exposure on deciduous and conifer tree species. International Journal of Forestry Research 2010:1–7.

Caldwell, M. M., C. L. Ballaré, J. F. Bornman, S. D. Flint, L. O. Bjorn, A. H. Teramura, G. Kulandaivelu, and M. Tevini. 2003. Terrestrial ecosystems, increased solar ultraviolet radiation and interactions with other climatic change factors. Photochemical & Photobiological Sciences 2:29–38.

Caldwell, M. M., J. F. Bornman, C. L. Ballaré, S. D. Flint, and G. Kulandaivelu. 2007. Terrestrial ecosystems, increased solar ultraviolet radiation, and interactions with other climate change factors. Photochemical & Photobiological Sciences 6:252–66.

Caputo, C., M. Rutitzky, and C. L. Ballaré. 2006. Solar ultraviolet-B radiation alters the attractiveness of Arabidopsis plants to diamondback moths (*Plutella xylostella* L.): impacts on oviposition and involvement of the jasmonic acid pathway. Oecologia 149:81–90.

Chen, M.-S. 2008. Inducible direct plant defense against insect herbivores: A review. Insect Science 15:101–114.

Cowan, P. D. 2010. Flammability, physiology and coexistence in fire prone environments. University of California, Berkeley.

Davies, S., and L. Unam. 1999. Smoke-haze from the 1997 Indonesian forest fires: effects on pollution levels, local climate, atmospheric CO 2 concentrations, and tree photosynthesis. Forest Ecology and Management 124:137–144.

Daws, M. I., J. Davies, H. W. Pritchard, N. A. Brown, and J. van Staden. 2006. Butenolide from plant-derived smoke enhances germination and seedling growth of arable weed species. Plant Growth Regulation 51:73–82.

De Moraes, C. M., W. Lewis, P. Pare, H. T. Alborn, and J. H. Tumlinson. 1998. Herbivore-infested plants selectively attract parasitoids. Nature 393:570–573.

De Moraes, C. M., M. C. Mescher, and J. H. Tumlinson. 2001. Caterpillar-induced nocturnal plant volatiles repel conspecific females. Nature 410:577–80.

De Moraes, C. M., J. C. Schultz, M. C. Mescher, and J. H. Tumlinson. 2004. Induced plant signaling and its implications for environmental sensing. Journal of Toxicology and Environmental Health, Part A 67:819–834.

Dicke, M., R. M. P. van Poecke, and J. G. de Boer. 2003. Inducible indirect defence of plants: from mechanisms to ecological functions. Basic and Applied Ecology 4:27–42.

Dixon, K. W., S. Roche, and J. Pate. 1995. The promotive effect of smoke derived from burnt native vegetation on seed germination of Western Australian plants. Oecologia 101:185–192.

Eckey-Kaltenbach, H., D. Ernst, W. Heller, and J. Sandermann, H. 1994. Biochemical plant responses to ozone: IV. Cross-induction of defensive pathways in parsley (*Petroselinum crispum* L.) plants. Plant Physiology 104:67–74.

Edwards, P. J., and S. D. Wratten. 1983. Wound induced defences in plants and their consequences for patterns of insect grazing. Oecologia 59:88–93.

Evtyugina, M., A. I. Calvo, T. Nunes, C. A. Alves, A. P. Fernandes, L. Tarelho, A. Vicente, and C. Pio. 2013. VOC emissions of smouldering combustion from Mediterranean wildfires in central Portugal. Atmospheric Environment 64:339–348.

Fares, S., J.-H. Park, E. Ormeño, D. R. Gentner, M. McKay, F. Loreto, J. Karlik, and A. H. Goldstein. 2010. Ozone uptake by citrus trees exposed to a range of ozone concentrations. Atmospheric Environment 44:3404–3412.

Frederickson, M. E., A. Ravenscraft, L. M. Arcila Hernández, G. Booth, V. Astudillo, and G. A. Miller. 2013. What happens when ants fail at plant defence? Cordia nodosa dynamically adjusts its investment in both direct and indirect resistance traits in response to herbivore damage. Journal of Ecology 101:400–409.

Frost, C. J., H. M. Appel, J. E. Carlson, C. M. De Moraes, M. C. Mescher, and J. C. Schultz. 2007. Within-plant signalling via volatiles overcomes vascular constraints on systemic signalling and primes responses against herbivores. Ecology Letters 10:490–8.

Frost, C. J., M. C. Mescher, J. E. Carlson, and C. M. De Moraes. 2008. Plant defense priming against herbivores: getting ready for a different battle. Plant Physiology 146:818–24.

Fujita, M., Y. Fujita, Y. Noutoshi, F. Takahashi, Y. Narusaka, K. Yamaguchi-Shinozaki, and K. Shinozaki. 2006. Crosstalk between abiotic and biotic stress responses: a current view from the points of convergence in the stress signaling networks. Current Opinion in Plant Biology 9:436–42.

Fürstenberg-Hägg, J., M. Zagrobelny, and S. Bak. 2013. Plant defense against insect herbivores. International Journal of Molecular Sciences 14:10242–97.

Geiselhardt, S., K. Yoneya, B. Blenn, N. Drechsler, J. Gershenzon, R. Kunze, and M. Hilker. 2013. Egg laying of cabbage white butterfly (Pieris brassicae) on Arabidopsis thaliana affects subsequent performance of the larvae. PloS One 8:e59661.

Ghebrehiwot, H., and M. G. Kulkarni. 2009. Smoke solutions and temperature influence the germination and seedling growth of South African mesic grassland species. Rangeland Ecology & Management 62:572–578.

Ghebrehiwot, H., M. G. Kulkarni, K. P. Kirkman, and J. van Staden. 2008. Smoke-water and a smoke-isolated butenolide improve germination and seedling vigour of *Eragrostis tef* (Zucc.) Trotter under high temperature and low osmotic potential. Journal of Agronomy and Crop Science 194:270–277.

Gilbert, M., and B. Ripley. 2002. The effect of smoke on the photosynthetic gas exchange of *Chrysanthemoides monilifera*. South African Journal of Botany 68:525–531.

Goodspeed, T. 1954. The Genus Nicotiana. Chronica Botanica Company, Waltham, MA.

Halitschke, R., U. Schittko, G. Pohnert, W. Boland, and I. T. Baldwin. 2001. Molecular interactions between the specialist herbivore *Manduca sexta* (Lepidoptera, Sphingidae) and its natural host *Nicotiana attenuata*. III. Fatty acid-amino acid conjugates in herbivore oral secretions are necessary and sufficient for herbivore-specifi. Plant Physiology 125:711–7.

Heiden, A., T. Hoffmann, and J. Kahl. 1999. Emission of volatile organic compounds from ozone-exposed plants. Ecological Applications 9:1160–1167.

Heil, M., and I. T. Baldwin. 2002. Fitness costs of induced resistance: emerging experimental support for a slippery concept. Trends in Plant Science 7:61–7.

Heil, M., and J. C. S. Bueno. 2007. Within-plant signaling by volatiles leads to induction and priming of an indirect plant defense in nature. Proceedings of the National Academy of Sciences of the United States of America 104:5467–72.

Hilker, M., and T. Meiners. 2011. Plants and insect eggs: how do they affect each other? Phytochemistry 72:1612–23.

Holopainen, J. K. 2004. Multiple functions of inducible plant volatiles. Trends in Plant Science 9:529–33.

Howe, G. A., and G. Jander. 2008. Plant immunity to insect herbivores. Annual Review of Plant Biology 59:41–66.

Izaguirre, M. M., C. A. Mazza, A. Svatos, I. T. Baldwin, and C. L. Ballaré. 2007. Solar ultraviolet-B radiation and insect herbivory trigger partially overlapping phenolic responses in *Nicotiana attenuata* and *Nicotiana longiflora*. Annals of Botany 99:103–9.

Izaguirre, M. M., A. Scopel, I. T. Baldwin, and C. L. Ballaré. 2003. Convergent responses to stress. Solar ultraviolet-B radiation and *Manduca sexta* herbivory elicit overlapping transcriptional responses in field-grown plants of *Nicotiana longiflora*. Plant Physiology 132:1755–1767.

Jain, N., and J. van Staden. 2006. A smoke-derived butenolide improves early growth of tomato seedlings. Plant Growth Regulation 50:139–148.

Kahl, J., D. H. Siemens, R. Aerts, and R. Gäbler. 2000. Herbivore-induced ethylene suppresses a direct defense but not a putative indirect defense against an adapted herbivore. Planta 210:336–342.

Karban, R. 2001. Communication between sagebrush and wild tobacco in the field. Biochemical Systematics and Ecology 29:995–1005.

Karban, R., I. T. Baldwin, K. J. Baxter, G. Laue, and G. W. Felton. 2000. Communication between plants: induced resistance in wild tobacco plants following clipping of neighboring sagebrush. Oecologia 125:66–71.

Karban, R., and J. Maron. 2002. The fitness consequences of interspecific eavesdropping between plants. Ecology 83:1209–1213.

Karban, R., J. Maron, G. W. Felton, G. Ervin, and H. Eichenseer. 2003. Herbivore damage to sagebrush induces resistance in wild tobacco: evidence for eavesdropping between plants. Oikos 100:325–332.

Keeley, J. E., and C. Fotheringham. 1998. Smoke-induced seed germination in California chaparral. Ecology 79:2320–2336.

Kerchev, P. I., B. Fenton, C. H. Foyer, and R. D. Hancock. 2012. Plant responses to insect herbivory: interactions between photosynthesis, reactive oxygen species and hormonal signalling pathways. Plant, Cell & Environment 35:441–53.

Kessler, A., and I. T. Baldwin. 2001. Defensive function of herbivore-induced plant volatile emissions in nature. Science 291:2141–2145.

Kessler, A., R. Halitschke, C. Diezel, and I. T. Baldwin. 2006. Priming of plant defense responses in nature by airborne signaling between *Artemisia tridentata* and *Nicotiana attenuata*. Oecologia 148:280–92.

Kessler, A., and M. Heil. 2011. The multiple faces of indirect defences and their agents of natural selection. Functional Ecology 25:348–357.

Khan, M. B., and J. B. Harborne. 1991. A comparison of the effect of mechanical and insect damage on alkaloid levels in *Atropa acuminata*. Biochemical Systematics and Ecology 19:529–534.

Kim, J., and G. W. Felton. 2013. Priming of antiherbivore defensive responses in plants. Insect Science 20:273–85.

Kim, J., J. F. Tooker, D. S. Luthe, C. M. De Moraes, and G. W. Felton. 2012. Insect eggs can enhance wound response in plants: a study system of tomato *Solanum lycopersicum* L. and *Helicoverpa zea* Boddie. PloS One 7:e37420.

Knight, T. M., and R. D. Holt. 2005. Fire generates spatial gradients in herbivory: an example from a Florida sandhill ecosystem. Ecology 86:587–593.

Koniak, S. 1985. Succession in pinyon-juniper woodlands following wildfire in the Great Basin. Western North American Naturalist 45:556–566.

Krock, B., S. Schmidt, C. Hertweck, and I. T. Baldwin. 2002. Vegetation-derived abscisic acid and four terpenes enforce dormancy in seeds of the post-fire annual, *Nicotiana attenuata*. Seed Science Research 12:239–252.

Krügel, T., M. Lim, K. Gase, R. Halitschke, and I. T. Baldwin. 2002. Agrobacterium -mediated transformation of *Nicotiana attenuata*, a model ecological expression system. Chemoecology 183:177–183.

Lamari, L. 2002. Assess - Image analysis software for plant disease quantification. American Phytopathological Society Press, St. Paul, Minnesota.

Lopes, C., and H. Vasconcelos. 2011. Fire increases insect herbivory in a neotropical savanna. Biotropica 43:612–618.

Lou, Y., and I. T. Baldwin. 2004. Nitrogen supply influences herbivore-induced direct and indirect defenses and transcriptional responses in *Nicotiana attenuata*. Plant Physiology 135:496–506.

Ma, G.-H., E. Bunn, K. W. Dixon, and G. Flematti. 2006. Comparative enhancement of germination and vigor in seed and somatic embryos by the smoke chemical 3-methyl-2H-furo[2,3-C]pyran-2-one in Baloskion tetraphyllum (Restionaceae). In Vitro Cellular & Developmental Biology - Plant 42:305–308.

McCullough, D. G., R. a Werner, and D. Neumann. 1998. Fire and insects in northern and boreal forest ecosystems of North America. Annual Review of Entomology 43:107–27.

McKenzie, D., D. L. Peterson, and J. J. Littell. 2009. Global warming and stress complexes in forests of western North America. Pages 319–337 *in* A. Bytnerowicz, M. Arbaugh, A. Riebau, and C. Anderson, editors. Developments in Environmental Science, Volume: 8. Elsevier Ltd.

Mithöfer, A., G. Wanner, and W. Boland. 2005. *Spodoptera littoralis* on lima bean leaves. II. Continuous mechanical wounding resembling insect feeding is sufficient to elicit herbivory-related volatile emission. Plant Physiology 137:1160–1168.

Musser, R. O., S. Hum-Musser, and H. Eichenseer. 2002. Herbivory: caterpillar saliva beats plant defences. Nature 416:599–600.

Ormeño, E., D. R. Gentner, S. Fares, J. Karlik, J.-H. Park, and A. H. Goldstein. 2010. Sesquiterpenoid emissions from agricultural crops: correlations to monoterpenoid emissions and leaf terpene content. Environmental Science & Technology 44:3758–64.

Parmenter, R. R., M. Kreutzian, D. I. Moore, and D. C. Lightfoot. 2011. Short-term effects of a summer wildfire on a desert grassland arthropod community in New Mexico. Environmental Entomology 40:1051–1066.

Peiffer, M., J. F. Tooker, D. S. Luthe, and G. W. Felton. 2009. Plants on early alert: glandular trichomes as sensors for insect herbivores. New Phytologist 184:644–656.

Preston, C. A., H. Betts, and I. T. Baldwin. 2002. Methyl jasmonate as an allelopathic agent: sagebrush inhibits germination of a neighboring tobacco, *Nicotiana attenuata*. Journal of Chemical Ecology 28:2343–69.

Radojevic, M. 2003. Chemistry of forest fires and regional haze with emphasis on Southeast Asia. Pure and Applied Geophysics 160:157–187.

Rao, M. V, H. Lee, R. A. Creelman, J. E. Mullet, and K. R. Davis. 2000. Jasmonic acid signaling modulates ozone-induced hypersensitive cell death. The Plant Cell 12:1633–46.

Reisenman, C. E., J. A. Riffell, K. Duffy, A. Pesque, D. Mikles, and B. Goodwin. 2013. Species-specific effects of herbivory on the oviposition behavior of the moth *Manduca sexta*. Journal of Chemical Ecology 39:76–89.

Sandberg, D., R. D. Ottmar, J. L. Peterson, and J. Core. 2002. Wildland fire in ecosystems: effects of fire on air. Page 77. Ogde, UT: U.S.

Schraudner, M., C. Langebartels, J. Negrel, and J. Sandermann, H. 1993. Plant defense reactions induced in tobacco by the air pollutant ozone. Pages 286–290 *in* B. Fritig and M. Legrand, editors. Mechanisms of Plant Defense Responses. Kluwer Academic Publishers, Netherlands.

Schultz, M. G., A. Heil, J. J. Hoelzemann, A. Spessa, K. Thonicke, J. G. Goldammer, A. C. Held, J. M. C. Pereira, and M. van het Bolscher. 2008. Global wildland fire emissions from 1960 to 2000. Global Biogeochemical Cycles 22:GB2002.

Schwachtje, J., and I. T. Baldwin. 2004. Smoke exposure alters endogenous gibberellin and abscisic acid pools and gibberellin sensitivity while eliciting germination in the post-fire annual, *Nicotiana attenuata*. Seed Science Research 14:51–60.

Schwachtje, J., and I. T. Baldwin. 2008. Why does herbivore attack reconfigure primary metabolism? Plant physiology 146:845–51.

Schwartz, M. W., S. M. Herman, and C. S. Vogel. 1995. The catastrophic loss of *Torreya taxifolia*: assesing environmental induction of disease hypotheses. Ecological Applications 5:501–516.

Shallwani, P. 2009. Interview with Kent Kirshenbaum. http://web.archive.org/web/20130202014550/http://www.slashfood.com/2009/06/23/liquid-smoke-what-is-it.

Sofiev, M. 2013. Wildland fires: Monitoring, plume modelling, impact on atmospheric composition and climate. Pages 451–472 *in* R. Matyssek, N. Clarke, P. Cudlin, T. Mikkelsen, J. Tuovinen, G. Wieser, and E. Paoletti, editors. Climate Change, Air Pollution and Global Challenges. First edition. Elsevier Ltd.

Sparg, S. G., M. G. Kulkarni, and J. van Staden. 2006. Aerosol smoke and smoke-water stimulation of seedling vigor of a commercial maize cultivar. Crop Science 46:1336.

Swengel, A. 2001. A literature review of insect responses to fire, compared to other conservation managements of open habitat. Biodiversity & Conservation:1141–1169.

Takabayashi, J., M. Dicke, S. Takahashi, M. A. Posthumus, and T. A. Van Beek. 1994. Leaf age affects composition of herbivore-induced synomones and attraction of predatory mites. Journal of Chemical Ecology 20:373–86.

Takabayashi, J., S. Takahashi, M. Dicke, and M. A. Posthumus. 1995. Developmental stage of herbivore Pseudaletia separata affects production of herbivore-induced synomone by corn plants. Journal of Chemical Ecology 21:273–287.

Thaler, J. S. 1999. Jasmonate-inducible plant defences cause increased parasitism of herbivores. Nature 399:17–19.

Tian, D., J. F. Tooker, M. Peiffer, S. H. Chung, and G. W. Felton. 2012. Role of trichomes in defense against herbivores: comparison of herbivore response to woolly and hairless trichome mutants in tomato (Solanum lycopersicum). Planta 236:1053–66.

Underwood, N., K. Anderson, and B. Inouye. 2005. Induced vs. constitutive resistance and the spatial distribution of insect herbivores among plants. Ecology 86:594–602.

Urbanski, S., W. M. Hao, and S. Baker. 2009. Chemical composition of wildland fire emissions. Pages 79–107 *in* A. Bytnerowicz, M. Arbaugh, A. Riebau, and C. Andersen, editors. Developments in Environmental Science, Volume: 8. Elsevier Ltd.

Van Staden, J., A. Jager, M. E. Light, and B. Burger. 2004. Isolation of the major germination cue from plant-derived smoke. South African Journal of Botany 70:654–659.

Van Staden, J., S. G. Sparg, M. G. Kulkarni, and M. E. Light. 2006. Post-germination effects of the smoke-derived compound 3-methyl-2H-furo[2,3-c]pyran-2-one, and its potential as a preconditioning agent. Field Crops Research 98:98–105.

Vuorinen, T., A.-M. Nerg, and J. K. Holopainen. 2004. Ozone exposure triggers the emission of herbivore-induced plant volatiles, but does not disturb tritrophic signalling. Environmental Pollution 131:305–11.

Ward, D. E., and C. C. Hardy. 1991. Smoke emissions from wildland fires. Environment International 17:117–134.

Wells, P. 1959. An ecological investigation of two desert tobaccos. Ecology 40:626–644.

Winz, R. A., and I. T. Baldwin. 2001. Molecular interactions between the specialist herbivore *Manduca sexta* (Lepidoptera, Sphingidae) and its natural host *Nicotiana attenuata*. IV. Insect-Induced ethylene reduces jasmonate-induced nicotine accumulation by regulating putrescine N-methyltransfer. Plant Physiology 125:2189–202.

Wu, J., C. Hettenhausen, M. C. Schuman, and I. T. Baldwin. 2008. A comparison of two *Nicotiana attenuata* accessions reveals large differences in signaling induced by oral secretions of the specialist herbivore *Manduca sexta*. Plant Physiology 146:927–39.

Young, J., and R. Evans. 1978. Population dynamics after wildfires in sagebrush grasslands. Journal of Range Management 31:283–289.

Zangerl, A. R., J. G. Hamilton, T. J. Miller, A. R. Crofts, K. Oxborough, M. R. Berenbaum, and E. H. de Lucia. 2002. Impact of folivory on photosynthesis is greater than the sum of its holes. Proceedings of the National Academy of Sciences of the United States of America 99:1088–91.

Chapter Two

Geographic genetic structure and diversity of native range and introduced Colorado potato beetle, Leptinotarsa decemlineata *(Coleoptera: Chrysomelidae)*

2.1 Introduction

Invasions are dramatic events in the history of species that can have significant genetic consequences. The invasion process is classically thought to reduce genetic diversity in invasive populations through founder effects and population bottlenecks, both of which can cause the loss of rare alleles (Nei et al. 1975, Sakai et al. 2001). Strong selection to novel environmental conditions in an introduced range can reduce genetic diversity. Introductions that stem from a single introduction tend not to display geographic genetic structure, as with the Argentine ant (Tsutsui and Case 2001). Thus for several reasons, both genetic diversity and population genetic structure are expected to be reduced in invasive populations relative to native populations. However, an increasing number of studies have shown that invasive populations may be founded by introductions from multiple populations, allowing these admixed populations to be more invasive. Recent work including studies of anole lizards (Kolbe et al. 2004), reed canarygrass (Lavergne and Molofsky 2007), walnut husk fly (Chen et al. 2010), wasp spiders (Krehenwinkel and Tautz 2013) and a marine mollusk (Riquet et al. 2013) suggest expected reductions in diversity and structure are not always found in studies of actual invasive populations. Processes including admixture and hybridization between multiple sources can help maintain or enhance genetic diversity in introduced populations, and may play more important roles in successful invasions than previously realized.

Colorado potato beetle, *Leptinotarsa decemlineata* (Say), is one of the most serious insect pests of potato, and invasive across the northern hemisphere. While Colorado potato beetle (CPB) does not feed on potato in its native range in southwestern North America, but rather principally on the wild buffalo bur *Solanum rostratum* Dunal (Hsiao 1978) and *S. angustifolium* Mill., invasive populations cause significant economic damage to potato crops as well as other cultivated solanaceous crops. CPB was first collected from its native host in 1811 along the Iowa-Nebraska boarder, and was not recorded from potato until 1859, despite the presence of potato in the area as early as 1820 (Casagrande 1985, Jaques 1988). While many early histories of CPB include the eastern slopes of the Rocky Mountains as part of its native range, currently most researchers (Lu and Logan 1994, Azeredo-Espin et al. 1996, Piiroinen et al. 2013) think the original range of CPB was limited to central and Southern Mexico. Subsequent to the conquest of Mexico by Spain, the distribution of CPB extended northward into the Central Plains following the spread of *S. rostratum* by Spanish travelers and Mexican vaqueros herding cattle (Hsiao 1978, Lu and Lazell 1996, Piiroinen et al. 2013). Although the invasive spread of CPB,

first across North America and then Eurasia, has been well documented, less is known about the very early stages of range expansion and invasion. There are still unresolved questions regarding the origins of the invasive population, the extent of gene flow, and population structure of CPB across the native and invaded range of North America.

Most studies of genetic diversity in CPB have focused on introduced populations of CPB collected from potato, and results have varied in the level of genetic diversity and population structure found across introduced populations of CPB (Zehnder et al. 1992, Azeredo-Espin et al. 1996, Grapputo et al. 2005, Zhang et al. 2012). Differentiation is typically found between native and introduced populations of CPB in those studies that include samples from the center of origin (Jacobson and Hsiao 1983, Zehnder et al. 1992, Piiroinen et al. 2013), but limited sampling in the native range has constrained the conclusions that can be drawn about diversity and population structure there.

Population genomics methods offer powerful resources for understanding the demographic and evolutionary forces that have shaped the genomes of species. Genome-by-sequencing techniques utilizing restriction-site associated DNA (RAD) markers (Miller et al. 2007) makes genomic analysis at the population level feasible by targeting sequencing at a reduced portion of the genome and thereby allowing for the pooling of multiple individuals in a single sequencing lane (Baird et al. 2008, Davey and Blaxter 2010, Narum et al. 2013). These techniques have been employed to examine population differentiation, genetic diversity, population structure and phylogeography in a diverse range of model and non-model organisms, from mosquitoes (Emerson and Merz 2010), to fish (Catchen et al. 2013a, Corander et al. 2013, Keller et al. 2013), voles (White et al. 2013), and yeast (Hyma and Fay 2013).

In this study we used restriction-site associated DNA sequencing (RADseq) (Baird et al. 2008, Hohenlohe et al. 2010, Davey and Blaxter 2010, Baxter et al. 2011) to investigate the population structure, genetic diversity, and extent of differentiation between populations of *L. decemlineata* across its native range in Mexico, as well as among several locations in the introduced range along the East Coast of the USA. Our goal was to genetically characterize native and introduced populations, shed light on gene flow between populations, and increase understanding of the origins and early invasion history of Colorado potato beetle in the USA. The SNPs that we identified can be used to look for signatures of selection in CPB, to identify candidate genes for its success as an invasive crop pest, and to explore the potential for rapid evolution in Colorado potato beetle.

2.2 Methods

Sampling and DNA preparation

Insects were collected by Izzo (Izzo et al. in prep) from Mexico during July and August of 2009 and 2010, and from United States locations during June of the same years. Colorado potato beetle samples used in this analysis were collected from 11 locations in Mexico: Magdalena JA (n=5), Santa Rosalía JA (n=4), Caimanero JA (n=4), Iguala GR (n=6), San Marcos Tlapazola OA (n=4), San Pablo Villa de Mitla OA (n=2), Xochitepec MO (n=1), Palmar de Bravo PU (n=1), Texcoco de Mora ME(n=1), Calpulalpan TL (n=2), and Hueyotlipan TL (n=1); and three locations in the USA; Exmore VA (n=2), Hampstead MD (n=2) and South Hero VT (n=3) (Table and Figure 2.1). Although including many individuals in an analysis is always more informative than including few, revealing analyses can yet be conducted with smaller sample

sizes because the depth of sequencing possible within individuals compensates for the lower number of individuals. Mexico and United States samples were collected across a similar geographic distance of approximately 1000 km between the most distant collection sites within each country. Genomic DNA extractions were completed at UVM using the DNeasy Blood & Tissue Kit (Qiagen) following the manufacturer's instructions. 100 µl aliquots of purified genomic DNA were shipped to UC Berkeley on dry ice.

Extractions were treated with RNase A to remove any RNA contamination using a protocol from the Cornell University Institute for Genomic Diversity's Interactive Resource Center (1 µl heat-treated RNase A in 60 µl TE buffer DNA suspension, incubated 30 minutes at 37°C), precipitated with sodium acetate (1:10 volumes of NaOH (3M) and 2:1 volumes 95% EtOH in 1 volume of DNA solution, incubated on ice 30 minutes), centrifuged, decanted, washed, dried overnight and re-suspended in TE buffer. DNA concentration was checked in all samples using Qubit DNA quantification. Low-concentration samples were re-precipitated and suspended in ½ volume TE, and quantified again.

RADseq protocol

RAD-tag library preparation

Multiplexed DNA libraries were prepared for paired-end RAD sequencing using protocols that generally followed Etter and Johnson (2012) and Baxter et al. (2011)**.**

Barcodes and adapter prep

We used custom *R* scripts to design and select 48 barcodes for multiplexing samples, each 6 bp long and differing by at least 2 nucleotides. P1 adapters were designed following Etter and Johnson (2012) with 6 bp barcode ends and *PstI* overhangs (Table 2.2). We re-suspended HPLC purified single stranded oligonucleotides (Integrated DNA Technologies) in 1x TE buffer to prepare 200 nM stocks of each, then annealed the adapters using 10 µM each of complementary top and bottom strands in 1x annealing buffer (AB) (100mM Tris and 500mM NaCl for 10x AB). Annealing reactions were performed in a thermocycler for 2 minutes at 95°C, followed by a ramp to 25°C by 0.1°C/s, a 30-minute hold at 25°C and final hold at 4°C. Annealed adapters were first diluted to approximately 2 µM in 1x AB as a stock solution, then annealing was checked with fluorometric quantification (Qubit 2.0 HS assay; Invitrogen). Stock solutions of annealed adapters were then diluted to final 0.1 µM working concentration based on Qubit analysis of actual stock concentrations. One adapter was lost due to a cracked PCR well, leaving 47 barcoded primers for library preparation.

Table 2.1: Sampling locations for Colorado potato beetle.

Site ID	Location	Longitude	Latitude	Host plant	Region assignment	N
JA-1	Magdalena , Jalisco, MEX	103°59'41.70"W	20°54' 0.90"N	S. rostratum	Western Mexico	5
JA-2	Santa Rosalía, Jalisco, MEX	104° 3'34.26"W	20°49' 4.80"N	S. rostratum	Western Mexico	4
JA-3	Caimanero, Jalisco, MEX	103°58'57.42"W	20°31'58.20"N	S. rostratum	Western Mexico	4
GR	Iguala, Guerrero, MEX	99°29'57.96"W	18°19'37.50"N	S. angustifolium	Southern lowland Mexico	6
OX-1	San Marcos Tlapazola, Oaxaca, MEX	96°30'44.10"W	16°55'35.70"N	S. angustifolium	Southern lowland Mexico	4
OX-2	San Pablo Villa de Mitla, Oaxaca, MEX	96°22' 7.80"W	16°55'35.70"N	S. rostratum	Southern lowland Mexico	2
XM	Xochitepec, Morelos, MEX	99°13'57.48"W	18°47'52.38"N	S. angustifolium	Southern lowland Mexico	1
PU	Palmar de Bravo, Puebla, MEX	97°32'15.75"W	18°51'14.18"N	S. rostratum	Central highland Mexico	1
TE	Texcoco de Mora, Mexico State, MEX	98°53' 8.34"W	19°31' 8.34"N	S. rostratum	Central highland Mexico	1
TL-1	Calpulalpan, Tlaxcala, MEX	98°36'51.30"W	19°35'40.86"N	S. rostratum	Central highland Mexico	2
TL-2	Hueyotlipan, Tlaxcala, MEX	98°22'13.80"W	19°28'54.12"N	S. rostratum	Central highland Mexico	1
VA	Exmore, Virginia, USA	75°47' 5.75"W	37°35' 7.48"N	S. tuberosum	East Coast United States	2
MD	Hampstead, Maryland, USA	76°50'59.92"W	39°36'17.37"N	S. tuberosum	East Coast United States	2
VT	South Hero, Vermont, USA	73°19'20.29"W	44°38'27.48"N	S. tuberosum	East Coast United States	3

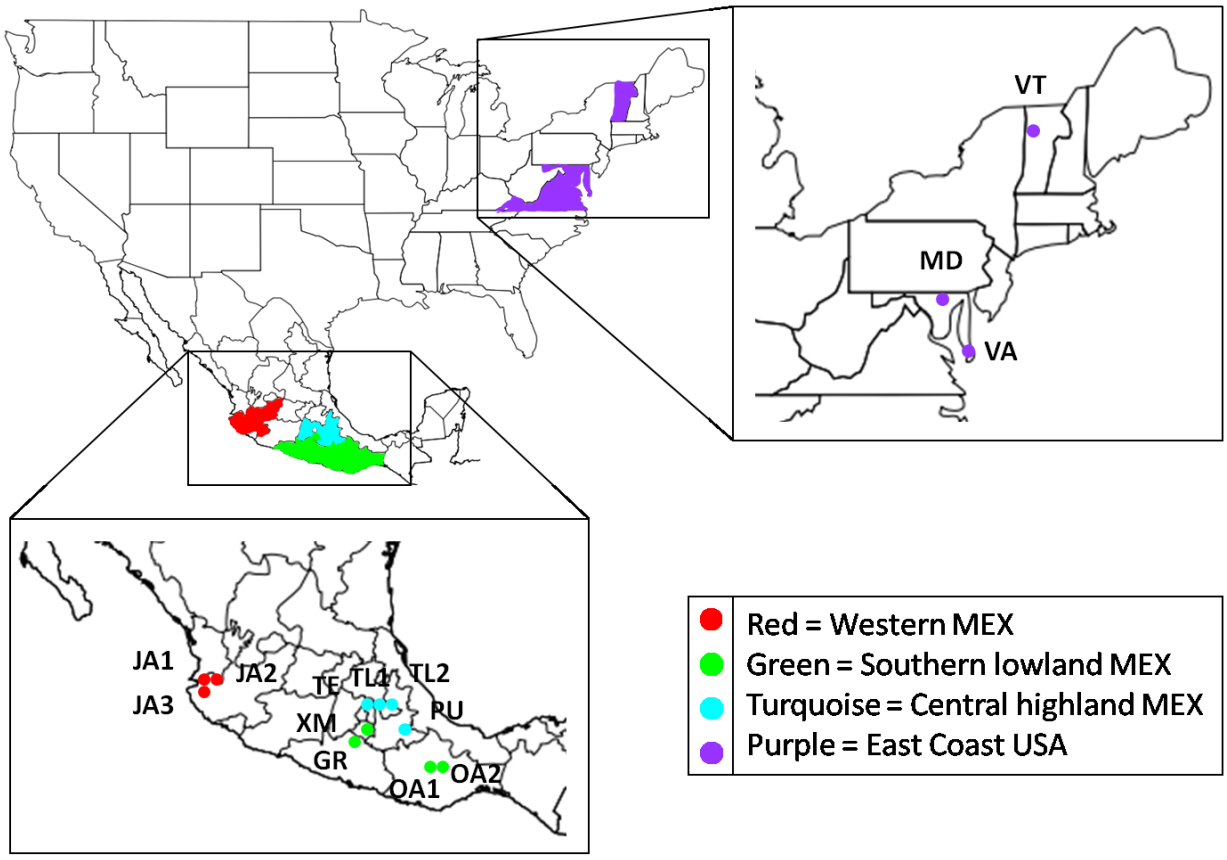

Figure 2.1: Map showing sampling locations of *L. decemlineata* in the United States and Mexico. Colors correspond to region assignment for inferred genetic clusters from discriminant analysis.

Restriction enzyme digest

We divided CPB DNA extracts into two batches for library preparation. Based on fluorometric quantification of DNA concentrations, 47 samples identified as having high DNA concentration were prepared in parallel for one multiplexed library, and 47 samples with low DNA concentrations were prepared for a second library. Extracted CPB genomic DNA was aliquoted into individual wells of a 96-well PCR plate to give 1 µg DNA per high-concentration sample, and 0.5 µg DNA per low-concentration sample, and dried overnight. Restriction endonuclease digestions were done using 0.5 µl (0.3 µl for low conc. samples) high fidelity *PstI* (CTGCAG) restriction enzyme (New England BioLabs, Ipswitch, MA) and 2 µl buffer 4 (NEB) in 17.5 µl H_2O (total reaction volume 20 µl). Reactions were incubated at 37°C for 90 minutes in a thermal cycler (iCycler; Bio-Rad), followed by heat inactivation according to manufacturer's instructions (20 minutes at 80°C) and a slow 45 minute cool-down to 4°C.

Adapter ligation and size selection

Prepared P1 adapters were ligated onto restriction-enzyme digested DNA as follows. To each restriction-digested sample we added 1 µl of P1 adapter, followed by 19 µl of prepared 1x ligation buffer master mix (5 µl Buffer 2 (NEB), 0.5 µl rATP (Promega) in 13.5 µl H_2O) and 10 µl of ligase master mix (0.4 µl (0.2 µl for low conc. samples) concentrated T4 DNA Ligase 2000 U/µl (NEB) in H_2O) for a total reaction volume of 50 µl. Ligation reactions proceeded for 30 minutes at room temperature, and then we held them at 65°C for 20 minutes before allowing them to cool back to room temperature. We then multiplexed libraries by combining 2 µl (4 µl for low concentration samples) of each adapter-ligated DNA sample in H_2O to produce 2 different libraries with 1.88 µg total DNA (0.04 µg DNA per sample) in either 100 µl volume (high concentration library) or 200 µl volume (low concentration library).

Multiplexed libraries were randomly sheared using sonication (Bioruptor; diagenode) for 10 cycles of 30 seconds each, then cleaned with MinElute PCR purification columns (Qiagen) and eluted to 20 µl volumes. Size selection was performed by running the products on a 2% agarose gel made with TALE and SYBR Gold Nucleic Acid Gel Stain (Invitrogen). The gel was run at 120 v for 45 minutes, and a 300-700bp fragment was excised on a UV light table. We used MinElute gel extraction kits to extract the size-selected library from the gel fragment, and eluted the DNA in 20 µl elution buffer.

We prepared P1-ligated, multiplexed libraries for P2-adapter ligation by repairing the ends and adenylating the 3' overhang. Ends were repaired using the Quick Blunting Kit (NEB), with 1 µl blunting enzyme, 2.5 µl dNTP, and 2.5 µl 10x quick blunting buffer (total volume 25 µl) incubated at room temperature for 30 minutes. Blunt-end libraries were cleaned with MinElute PCR columns, and eluted to 24 µl volume. We adenelated the libraries by adding 3 µl Klenow Fragment (5,000U/ml, NEB), 1 µl dATP (100mM, Fermentas), and 3 µl Buffer 2 (NEB) to each sample for a 30 µl reaction volume, and incubated them at 37°C for 30 minutes followed by a 15 minute slow cool back to room temperature. Adenelation was followed with another MinElute cleanup, and elution in 25 µl elution buffer.

The P2 paired-end adapter (Table 2.2), designed according to Etter and Johnson (2012), was HPLC purified (10 µM, IDT) and the top and bottom strands were annealed as for the P1 adapters. Adapters were ligated to DNA samples by combining 1 µl of 10 µM P2 adapter, 1 µl rATP (100mM, Promega), 3 µl of Buffer 2 (10x, NEB) and 0.5 µl T4 Ligase with prepared samples for a total reaction volume of 30 µl. Ligation reactions were held at room temperature

for 30 minutes, then cleaned using a MinElute PCR column and eluted to 50 µl of completed multiplexed library template.

PCR amplification, final size selection and sequencing

The library template was PCR amplified and size selected before being submitted for sequencing. For PCR amplification, we combined a 5 µl aliquot of library template with 50 µl Phusion High-Fidelity PCR Master Mix with HF Buffer (NEB), 5 µl each of the P1 forward PCR primer and P2 reverse PCR primer (Table 2.2), and water to a final volume of 100 µl. The PCR mix was split into 4 equal volumes and amplified on a thermal cycler (iCycler; Bio-Rad) for 18 cycles (30 seconds at 98°C, then 18x [98°C for 10 seconds, 65°C for 30 seconds and 72°C for 30 seconds] followed by 5 minutes at 72°C and final hold at 4°C. Reactions were combined and cleaned with MinElute PCR columns, and eluted in 20 µl elution buffer.

Final library size selection was done using a 2% agarose gel with TALE and SYBR Gold. We loaded the full 20 µl library volumes with 4 µl orange loading dye, with two empty wells between samples, and ran the gel at 120 v for 45 minutes. We excised 300-600 bp gel fragments, extracted the final libraries using MinElute Gel extraction kits, and eluted each into 25 µl elution buffer. Libraries were quantified on a Bioanalyzer, and submitted to the Delaware Biotechnology Institute's DNA Sequencing & Genotyping Center for sequencing on Illumina Hi-Seq 2000.

RADseq processing and population analysis

Processing

We processed raw sequence reads with the *process_radtags* program in *Stacks* 1.02 software pipeline (Catchen et al. 2011, 2013b). This program recovers individual sample data from a multiplexed library based on sample barcodes and presence of the restriction site overhang. We used 1 bp error correcting in the barcode and restriction site, filtered out adapter sequences, and trimmed all reads to 85 base pairs. We used the *Stacks: denovo_map* wrapper program to build stacks, identify RAD loci and call SNPs in the samples. We set processing parameters based on examples and default settings; we required a minimum coverage of three identical reads within an individual to form a stack (m=3) and allowed stacks with up to three nucleotide mismatches between them to be merged into a single RAD locus within an individual (M=3). We allowed up to two mismatches between RAD loci of different individuals when building the catalog (n=2), and enabled the lumberjack-stack removal algorithm to remove or break up highly repetitive RAD tags (t). Samples with fewer than 250,000 sequences recovered by *process_radtags* were removed from subsequent analyses, leaving 77 samples in the early processing. Read quality was significantly lower for low-DNA concentration samples than for high-DNA concentration samples, so analyses proceeded for the 38 high-concentration samples only.

Table 2.2: Adapter and primer oligonucleotide sequences used for library preparation and PCR in amplification prior to Illumina Hi-Seq paired-end sequencing. X and Y indicate forward and reverse locations of unique 6 bp barcodes.

		Sequence
P1 Adapter	top	5′- ACACTCTTTCCCTACACGACGCTCTTCCGATCTxxxxxxTGC*A -3′
	bottom	5′- P-yyyyyAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT -3′
P2 Adapter	top	5′- P-CTCAGGCATCACTCGATTCCTCCGAGAACAA -3′
	bottom	5′- CAAGCAGAAGACGGCATACGACGGAGGAATCGAGTGATGCCTGAG*T -3′
PCR Primers	P1-forward	5′- AATGATACGGCGACCACCG*A -3′
	P2-reverse	5′- CAAGCAGAAGACGGCATACG*A -3′

SNP calling

As a starting point to data analysis, we arbitrarily grouped samples into 10 preliminary populations by considering collection locations within the same state as a single state-wide population. Samples collected within a state were always geographically closer to each other than to collection sites from neighboring states. After removing samples with very low representation in the original catalog of loci, we used *Stacks*: *populations* (Catchen et al. 2013b) to call SNPs in these state-wide populations.

Regional population structure and diversity indexes

We restricted the SNPs output for further processing such that any loci processed were present in nine of the ten state-wide populations. Only the first SNP from each locus was included in the output used for analysis of population genetic structure, making all loci independent. In order to test for significant population structure, we used the Discriminant Analysis of Principal Components (*DAPC*) module of the *R: adegenet* package (Jombart 2008) to analyze these 30,426 SNPs. We used the *adegenet* function *find.clusters* to determine the optimal number of groups. We then used *Stacks*: *populations* to calculate population statistics using these genetic clusters as regional populations, and required that included loci be present in at least 50% of the individuals within each of these new regional populations. All diversity indexes, including F_{ST} and heterozygosity values, were calculated within the *Stacks: populations* program. *Populations* uses an implementation of Wright's F_{ST} that is applicable for bi-allelic SNP data (Catchen et al. 2013b). We analyzed the regional population structure based on these SNPs with *DAPC*, which maximizes differences between groups while minimizing differences within groups, first transforming data with Principal Component Analysis and then conducting a Discriminant Analysis of the retained PCs. We retained only the first three PCAs since that was sufficient for describing the genetic clusters without causing over-fitting of the data. Discriminant analysis used all three discriminant eigenvectors.

2.3 Results

RAD sequencing and loci identification

After filtering for a correct barcode and *PstI* recognition site, and applying a lower cut-off for the number of valid sequences returned for each individual, we generated an average of 951,949 reads across 77 samples (Table 2.3). For the 38 individual CPB from 14 locations included in this analysis (Table 2.1), the mean number of reads was 1,300,995. We identified a total of 1,048,365 RAD loci across these samples, and 4,542 of these passed our filter of being present in at least 50% of individuals assigned to each regional group. The first SNP from each of these RAD loci was included in the discriminant analysis.

Table 2.3: Sequencing reads processed and RADtags utilized by *Stacks*.

	Mean depth	Valid RADtags	RADtags used to build stacks	Stacks matched to RAD loci catalog
Cut-off minimum for processing	---	250,000	100,000	40,000
Mean for 77 samples retained in early processing	6.7	951,949	733,683	---
Mean for 38 samples retained in final analysis	6.6	1,300,995	1,139,135	134,180

Geographic genetic structure

We identified four genetic clusters, with each cluster representing one geographic region; Southern lowland Mexico (Oaxaca, Morelos, and Guerrero), Western Mexico (Jalisco), Central highlands of Mexico (Tlaxcala, Texcoco, Puebla), and East Coast USA (Vermont, Maryland, Virginia) (Figure 2.2.c). As the BIC values for two, three, and four clusters were not appreciably different from each other (kstat ranged from 220.8146 to 221.2846) (Figure 2.3), we chose to complete our analysis using the maximum number of distinct clusters supported by our data. Under a two-cluster K-means model, Western Mexico and Southern Mexico samples were placed in one cluster (Figure 2.2.a), with Central Mexico and USA samples in the other cluster. Using three clusters further separated Western and Southern Mexico individuals from each other (Figure 2.2.b). Of over 1,000 iterations requiring that *find.clusters* identify four clusters, the groupings used in the following analyses were created 98.4% of the time. Of the 16 conflicting runs, East Coast USA samples were lumped with Central highland Mexico 6 times, with six of the Oaxaca individuals split out of the Southern lowland Mexico group into a separate cluster. Western Mexico and Southern lowland Mexico samples were lumped together the remaining 10 times. Of these runs, East Coast USA was split into two clusters 7 times, twice a Central highland Mexico sample formed a cluster with two USA samples while leaving the balance of these groups intact, and once the Central highland Mexico cluster was split in two (Table 2.4, Figure 2.4)

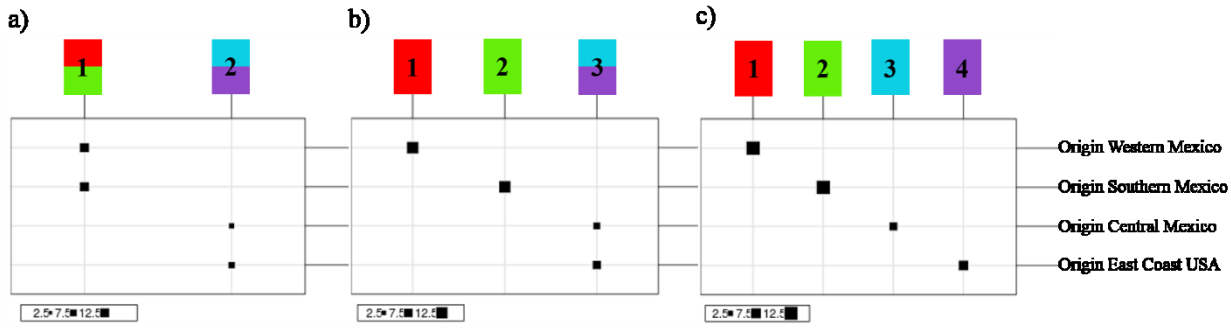

Figure 2.2: Genetic clusters identified by *find.clusters* using two, three, and four clusters. Collection region is on the horizontal axis, and inferred clusters on the vertical axis. The size of the cell is scaled by the number of individuals in the group.

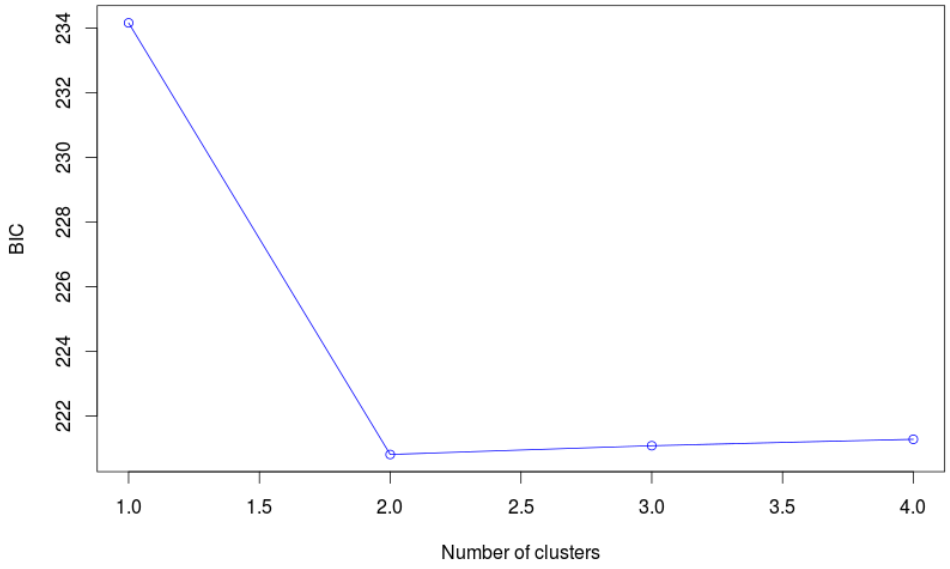

Figure 2.3: BIC values for different numbers of clusters in *find.clusters*.

Table 2.4: Number and type of alternative genetic clusters identified by 1.6% of iterations of *find.clusters* from DAPC analysis.

Number of iterations	Individuals lumped into one cluster	Individuals split from other clusters
7	Western Mexico + Southern Mexico	E.USA spilt into: MD x2 + VT VA x2 + VT x2
2	Western Mexico + Southern Mexico	TE + MD + VT split from otherwise coherent C.MEX and E.USA clusters
1	Western Mexico + Southern Mexico	C.MEX split into: TE + TXL PU + TXL x2
6	Central Mexico + East Coast USA	OX x6 split from S.MEX clusters

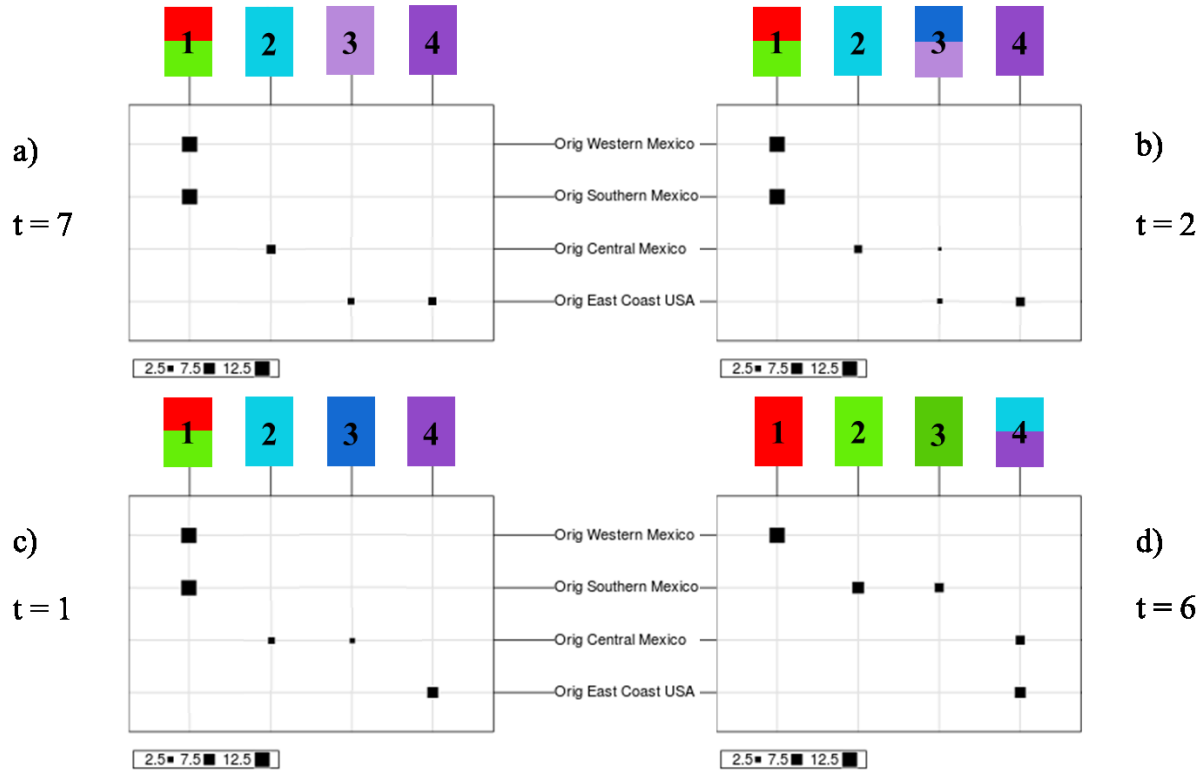

Figure 2.4: Alternative clusters identified in 1,000 iterations of *find.clusters*. Collection region is on the horizontal axis, and inferred clusters on the vertical axis. The size of the cell is scaled by the number of individuals in the group, and t = the number of times a clustering formation was identified. Colors refer to assignment of samples to clusters as identified in 98.4% of iterations.

Using only the first three principal components in the DAPC, individuals were correctly assigned with 100% probability to their collection region, and each of these four regions formed non-overlapping genetic clusters. Results using 4,542 SNPs present in at least 50% of individuals from each cluster (Figure 2.5) were not substantially different from analysis using only 488 SNPs present in 75% of individuals from each cluster (Figure 2.6).

Each regional genetic cluster was distinct, with the Western and Southern Mexico regions as nearest neighbors, sharing more similarity with each other than with other groups. The Central highland Mexico group fell between Southern Mexico and East Coast USA group, which were distant nearest neighbors compared to the difference between Western and Southern Mexico clusters. Host plant itself does not seem to be associated with assignment to a regional cluster in the native range. Beetles assigned to the Western and Central Mexico groups were collected on *Solanum rostratum*, and beetles from the Southern Mexico region were collected from both *S. rostratum* and *S. angustifolium*. Samples identified as belonging to the East Coast USA cluster were all collected from *S. tuberosum* in the introduced range.

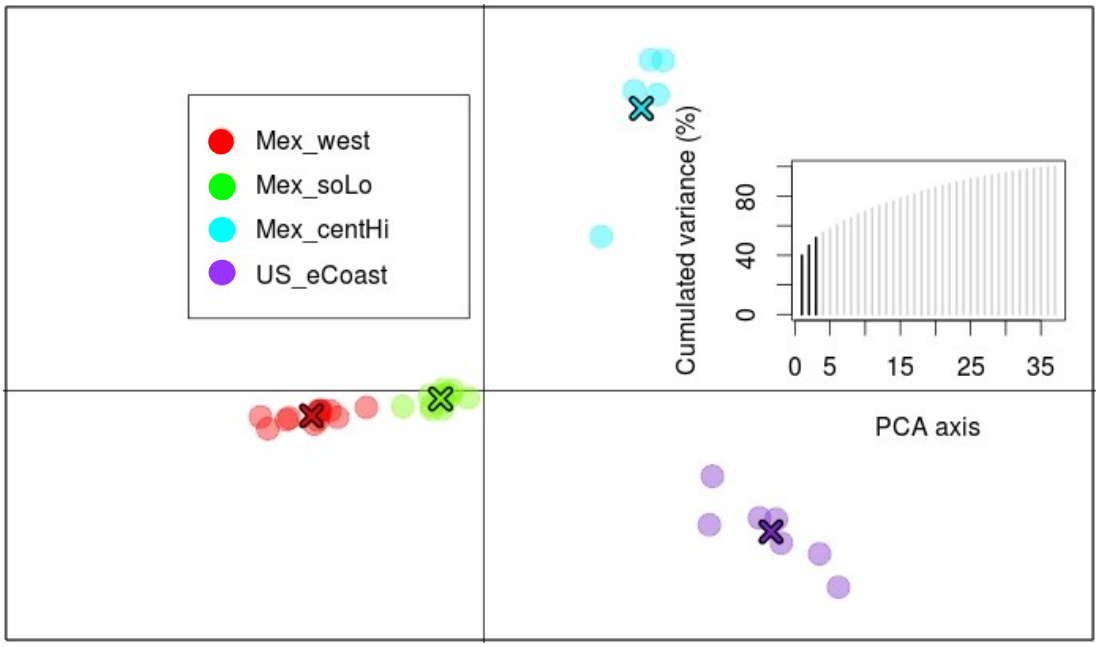

Figure 2.5: Genetic clustering of 38 Colorado potato beetles by discriminant analysis of the first three principal components using 4,542 SNPs. SNP loci included in this analysis were present in at least 50% of individuals from each of the four regions. The centroids of ellipses are marked by an X. Inset figure shows the percent of variance within and between groups that is explained by the three principal components.

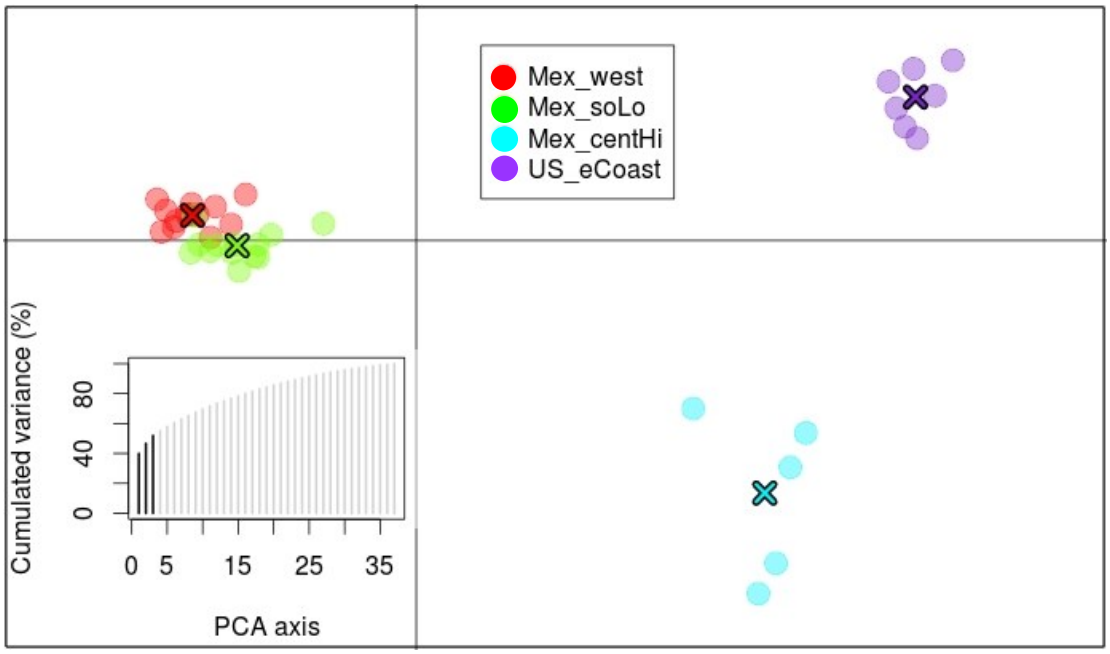

Figure 2.6: Genetic clustering of CPB using 488 SNPs present in at least 75% of individuals from each of the four regions. Xs mark ellipse centroids. Inset figure shows the percent of variance within and between groups that is explained by the three principal components.

Genetic differentiation

Overall, average pairwise F_{ST} values indicate three levels of differentiation between regions, ranging from moderate to very high differentiation (Table 2.5). Genetic differentiation was greatest between Western Mexico and both the East Coast USA and Central Mexico. The most similar region pairs were USA with Central Mexico, followed by Western Mexico with Southern Mexico, with moderate differentiation between pairs. Other pairwise comparisons indicate intermediate levels of differentiation between regions.

Genetic diversity

Genetic diversity was highest in the East Coast USA population and lowest in the Western Mexico population, as measured by the percent of polymorphic loci and the observed and expected heterozygosities (Table 2.6). The proportion of polymorphic loci was higher in the Southern lowland Mexico population than in the Central highland Mexico population, while the Central highland population had a higher expected heterozygosity than the Southern population as well as slightly higher observed heterozygosity. The number of private alleles was lowest in the Western Mexico population, with twice as many found in the Southern and Central Mexico populations and three times the number of private alleles in the East Coast USA population.

Table 2.5: Average pairwise FST values between regions for 38 individuals across 4,542 RADloci.

	Western Mexico	Southern highland Mexico	Central lowland Mexico	East Coast USA
Western Mexico	—			
Southern lowland Mexico	0.0745	—		
Central highland Mexico	0.2662	0.1984	—	
East Coast USA	0.2672	0.2113	0.0660	—

Table 2.6: Proportion of polymorphic sites, and observed and expected heterozygosities for variant positions. The number of samples at each locus (n) varied within each region due to random sampling introduced during sequencing. These values were calculated across all SNPs within the included RADloci, rather than only the first SNP per loci as used in DAPC.

Regions	Average n across alleles	Private alleles	% Polymorphic loci	Heterozygosity	
				Observed	Expected
Western Mexico	9.74	1881	1.48893	0.0527	0.0629
Southern lowland Mexico	8.57	3773	1.97180	0.0559	0.0877
Central highland Mexico	3.46	3743	1.63554	0.0564	0.0999
East Coast USA	5.05	5652	2.05794	0.0721	0.1073

2.4 Discussion

Our results show that Colorado potato beetles collected from *S. tuberosum* in the East Coast of the USA are genetically distinct from beetles collected from *Solanum* host plants in Mexico (Figure 2.5). Beetles from the USA introduced range are more genetically similar to native-range beetles from the Central highlands of Mexico than from other regions, therefore supporting this region as a more likely source population than the Western or Southern lowland regions of Mexico. Furthermore, our analysis indicates the presence of geographic patterns of genetic variation at the regional level in the native range. Genetic differentiation between individuals was higher among regions than within regions, and high discrimination between genetic clusters shows that these genetic markers can be used to identify a clear region of origin for these beetles. This pattern of distinct regional genomic diversity suggests that gene flow between these populations is limited enough to allow divergence between populations. Beetles from Western and Southern Mexico share more genetic similarities with each other than with beetles from other regions, and there is less discrimination between these two groups than there is between other groups.

United States samples collected across a large geographic distance (~1000 km) along the East Coast cluster together. This may be due to two different alternative, though not mutually exclusive, causes. On one hand, gene flow may be higher within agricultural settings than within the native range or between native and introduced ranges. On the other hand, genetic clustering across a large geographic distance may be the result of low initial diversity followed by recent spread. The historical evidence of very rapid spread of CPB across the USA following its host shift suggests that invasion history is likely an important contributor to this pattern. One way to distinguish between these alternatives is by looking for linkage disequilibrium. Low initial diversity followed by recent spread and subsequent mutation would result in individual unlinked SNPs being scattered throughout the genomes. In contrast, high gene flow among different populations would cause groups of SNPs with common ancestry to co-occur within introgressed linkage groups. However, high gene flow among genetically similar populations would be

difficult to distinguish from patterns produced by recent spread. Any assessment in this study of gene flow in the USA is limited, as sampling efforts focused on collecting across the native range in Mexico. Future analyses that include samples collected by Izzo et al. (unpublished) from Massachusetts, Missouri, Kansas, Texas and Washington may help resolve the extent of gene flow in the invaded range. Additionally, such sampling would answer whether genomic markers support genetic differentiation between populations in the invaded range as found in previous studies using mtDNA markers and AFLP analysis (Azeredo-Espin et al. 1996, Grapputo et al. 2005), or the limited discrimination found between introduced CPB in other studies (Izzo et al. n.d., Zehnder et al. 1992).

The pairwise F_{ST} values ranged from 0.066 to 0.267 (Table 2.5). These values are within the range of genetic differentiation seen in other studies of Coleptera. A 1987 survey of genetic differentiation in Coleptera based on allozymes found F_{ST} to range from 0.030 to 0.154 (McCauley and Eanes 1987). Included in this survey (McCauley and Eanes 1987) is an F_{ST} for Colorado potato beetle collected from 9 locations across the USA and one population from each of Mexico, Spain, and the Netherlands ($F_{ST} = 0.068$) calculated from Jacobson and Hsiao (1983). In a recent study of worldwide diversity in the red flour beetle, *Tribolim castaneum*, pairwise F_{ST} values for individual microsatellite loci ranged from 0.0289 to 0.353 with an overall F_{ST} of 0.180 across all loci and all populations (Drury et al. 2009). A comparison of two species of bean weevils in their native range using allozymes found differences in their overall F_{ST} values (*Acanthoscelides obtectus*: $F_{ST} = 0.082$; *A. obvelatus*, $F_{ST} = 0.036$; González-Rodríguez et al. 2000). The salt-mash beetle *Pogonus chalceus* showed regional differences in F_{ST} based on allozymes (Mediterranean: $F_{ST} = 0.045$, Atlantic: $F_{ST} = 0.147$) but not microsatellites (Mediterranean: $F_{ST} = 0.049$, Atlantic: $F_{ST} = 0.04$), and overall F_{ST} was higher for allozymes than for microsatellites (allozymes: $F_{ST} = 0.213$, microsatellites: $F_{ST} = 0.079$) (Dhuyvetter et al. 2004).

Our F_{ST} values indicate a greater overall similarity between East Coast USA and Central Mexico than evidenced by DAPC. Although F_{ST} values are useful in describing partitioning of variation between populations, because the invasion is relatively recent and F_{ST} values therefore have likely not reached equilibrium, they may not be accurate estimates of migration rates (Fitzpatrick et al. 2012). Small effective population size, for example due to sampled individuals being close relatives, small sample sizes, and variable sample size between populations can bias estimation of F_{ST} values. We sequenced few samples from some sites, and variable numbers of samples across sites, so these values may be over-estimates of population differentiation. Additionally, simulations of RADseq data indicate that F_{ST} values can be inflated by missing data in RADseq analyses due to incomplete sequencing and null allele dropout from mutations in the restriction enzyme recognition site (Arnold et al. 2013, Gautier et al. 2013).

Contrary to the expectation of reduced genetic diversity in introduced populations, and in contrast to Azeredo-Espin et al. (1996) who found lower mitochondrial diversity in East Coast samples than in other regions, we found the highest genetic diversity in the introduced population from the East Coast of the United States. We found nearly twice as many private alleles in samples from the East Coast USA region as the average number of private alleles in Mexican regions (average of MEX regions: private alleles = 3,132; East Coast USA: private alleles = 5,652; Table 2.6). Heterozygosity and percent polymorphic alleles were also highest in the introduced beetles (Table 2.6). The relatively low genetic diversity in the Western Mexico population suggests that this was a small, isolated population, or that it recently went through a reduction in population. Studies of recent invasions by CPB (Grapputo et al. 2005, Zhang et al.

2012) show classic patterns of reduced variation relative to source populations. In contrast, the high diversity found in the introduced population in this study, relative to the native range diversity, is characteristic of invasions produced by admixture and hybridization. This suggests that multiple genetically different introductions, hybridization, heterozygote advantage, or some other factor is more than compensating for historic small-population demographic processes. Phylogenomic analysis of our data, and ongoing whole genome sequencing in combination with more native range sampling, may allow us to distinguish between these forces.

There are several possible causes for the high genomic diversity and large number of private alleles observed in the US. First, this may be a response to different ecological conditions in the introduced range compared to those experienced by beetles across their native Mexican range. Alternatively, crop-rotation in potato fields may increase divergence of introduced CPB from native populations through genetic drift, by causing oscillations in population size as populations shift through phases of near-extinction and re-colonization (Azeredo-Espin et al. 1996). Third, the native range may not have been sampled sufficiently to include samples from the dominant source population or all of the primary source populations. Finally, hybridization with another species of *Leptinotarsus* might have played a role in the initial host shift, producing a uniquely invasive sub-population and facilitating the dramatic range expansion of the invasive Colorado potato beetle. As further analyses are pursued based on our data, and as more genomic resources become available for CPB and other insects, it may become possible to determine which of these factors contributed to the high genomic diversity in introduced populations.

Examining both native and introduced insect populations can provide insight into the history and the genomic impact of the introduction process. Our analysis of RADseq derived SNPs in CPB demonstrated significant genetic structure in the native range, with variable levels of genomic diversity. Based on levels of genetic differentiation, the Central highland Mexico region was more likely than other regions of Mexico to be the origin of the introduced beetles. Despite limited sampling of CPB in the introduced USA range, our data confirmed that these samples were genetically distinct from native Mexican range individuals, and additionally showed high genomic diversity. CPB in the USA thus provides another example of an introduced insect population possessing higher genetic diversity, rather than lower, in the introduced range. Further examination of the genomic diversity identified by RADseq analysis may provide insight into particular alleles that have been important in the success of CPB as a major crop pest, and could prove useful in its control.

Our study makes major contributions to the understanding of the invasion history of CPB, the geographic and genomic diversity of CPB across the native range, and the comparison of native range diversity to that of CPB from the East Coast of the USA. Further analysis of this data and future comparisons of additional samples across the native and introduced ranges may confirm a specific region of origin for the introduced beetles. If so, then management of this important crop pest may be aided by identification of potential biological control agents from that region. In the following chapter, we present a phylogenomic analysis of CPB as a next step in this process. If future and ongoing genomic studies of CPB investigate selection in CPB, this analysis can guide sample choice to maximize genomic diversity across ecological gradients. Our analysis suggests that, to encompass the diversity of native range CPB, future genomic analyses should include samples collected from all three regions of Mexico identified here. The SNPs identified here will be a valuable resource for guiding targeted re-sequencing efforts aimed at identifying genomic regions that have diverged between native and introduced populations. Our findings thus provide a foundation for future research in understanding the genetic basis for

the success of CPB on potato, and future efforts to transform that understanding into successful agricultural management of this critical crop pest. Finally, our findings add to the growing list of examples of invasions that do not conform to the classic expectations of reduced genetic diversity in introduced ranges. The specific mechanisms responsible for the high levels of genetic diversity in these introduced populations remains an open question.

2.5 Literature Cited

Arnold, B., R. B. Corbett-Detig, D. Hartl, and K. Bomblies. 2013. RADseq underestimates diversity and introduces genealogical biases due to nonrandom haplotype sampling. Molecular Ecology 22:3179–90.

Azeredo-Espin, A. M. L., R. Schroder, and G. K. Roderick. 1996. Intraspecific mitochondrial DNA variation in the Colorado potato beetle, *Leptinotarsa decemlineata* (Coleoptera: Chrysomelidae). Biochemical Genetics 34:253–268.

Baird, N. A., P. D. Etter, T. S. Atwood, M. C. Currey, A. L. Shiver, Z. A. Lewis, E. U. Selker, W. A. Cresko, and E. A. Johnson. 2008. Rapid SNP discovery and genetic mapping using sequenced RAD markers. PloS One 3:e3376.

Baxter, S. W., J. W. Davey, J. S. Johnston, A. M. Shelton, D. G. Heckel, C. D. Jiggins, and M. L. Blaxter. 2011. Linkage mapping and comparative genomics using next-generation RAD sequencing of a non-model organism. PloS One 6:e19315.

Casagrande, R. 1985. The Iowa Potato Beetle, its Discovery and Spread to Potatoes. Bulletin of the Entomological Society of America 31:27–29.

Catchen, J. M., A. Amores, P. A. Hohenlohe, W. A. Cresko, and J. H. Postlethwait. 2011. Stacks: building and genotyping Loci de novo from short-read sequences. G3 1:171–82.

Catchen, J. M., S. Bassham, T. Wilson, M. C. Currey, C. O'Brien, Q. Yeates, and W. A. Cresko. 2013a. The population structure and recent colonization history of Oregon threespine stickleback determined using restriction-site associated DNA-sequencing. Molecular Ecology 22:2864–83.

Catchen, J. M., P. A. Hohenlohe, S. Bassham, A. Amores, and W. A. Cresko. 2013b. Stacks: an analysis tool set for population genomics. Molecular Ecology 22:3124–40.

Chen, Y. H., S. H. Berlocher, S. B. Opp, and G. K. Roderick. 2010. Post-colonization temporal genetic variation of an introduced fly, *Rhagoletis completa*. Genetica 138:1059–75.

Corander, J., K. K. Majander, L. Cheng, and J. Merilä. 2013. High degree of cryptic population differentiation in the Baltic Sea herring *Clupea harengus*. Molecular Ecology 22:2931–40.

Davey, J. W., and M. L. Blaxter. 2010. RADSeq: next-generation population genetics. Briefings in Functional Genomics 9:416–23.

Dhuyvetter, H., E. Gaublomme, and K. Desender. 2004. Genetic differentiation and local adaptation in the salt-marsh beetle *Pogonus chalceus*: a comparison between allozyme and microsatellite loci. Molecular ecology 13:1065–74.

Drury, D., A. Siniard, and M. Wade. 2009. Genetic differentiation among wild populations of *Tribolium castaneum* estimated using microsatellite markers. Journal of Heredity 100:732–741.

Emerson, K., and C. Merz. 2010. Resolving postglacial phylogeography using high-throughput sequencing. Proceedings of the National Academy of Sciences of the United States of America 107:16196–16200.

Gautier, M., K. Gharbi, T. Cezard, J. Foucaud, C. Kerdelhué, P. Pudlo, J.-M. Cornuet, and A. Estoup. 2013. The effect of RAD allele dropout on the estimation of genetic variation within and between populations. Molecular Ecology 22:3165–78.

González-Rodríguez, A., B. Benrey, A. Castaneda, and K. Oyama. 2000. Population genetic structure of *Acanthoscelides obtectus* and *A. obvelatus* (Coleoptera: Bruchidae) from wild and cultivated *Phaseolus* spp.(Leguminosae). Annals of the Entomological Society of America 93:1100–1107.

Grapputo, A., S. Boman, L. Lindström, A. Lyytinen, and J. Mappes. 2005. The voyage of an invasive species across continents: genetic diversity of North American and European Colorado potato beetle populations. Molecular Ecology 14:4207–19.

Hohenlohe, P. A., S. Bassham, P. D. Etter, N. Stiffler, E. A. Johnson, and W. A. Cresko. 2010. Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. PLoS Genetics 6:e1000862.

Hsiao, T. H. 1978. Host plant adaptations among geographic populations of the Colorado potato beetle. Entomologia Experimentalis et Applicata 24:237–247.

Hyma, K. E., and J. C. Fay. 2013. Mixing of vineyard and oak-tree ecotypes of Saccharomyces cerevisiae in North American vineyards. Molecular Ecology 22:2917–30.

Izzo, V. M., D. J. Hawthorne, S. D. Schoville, and Y. H. Chen. (n.d.). No Title.

Jacobson, J. W., and T. H. Hsiao. 1983. Isozyme variation between geographic populations of the Colorado potato beetle, *Leptinotarsa decemlineata* (Coleoptera: Chrysomelidae). Annals of the Entomological Society of America 76:162–166.

Jaques, R. L. 1988. Early history of the potato beetle. Pages 29–35 *in* E. J. Brill, editor. The Potato Beetles: the Genus *Leptinotarsa* in North America (Coleoptera: Chrysomelidae). Flora & Fauna Handbook No. 3. New York.

Jombart, T. 2008. adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics (Oxford, England) 24:1403–5.

Keller, I., C. E. Wagner, L. Greuter, S. Mwaiko, O. M. Selz, A. Sivasundar, S. Wittwer, and O. Seehausen. 2013. Population genomic signatures of divergent adaptation, gene flow and

hybrid speciation in the rapid radiation of Lake Victoria cichlid fishes. Molecular Ecology 22:2848–63.

Kolbe, J. J., R. E. Glor, L. Rodríguez Schettino, A. C. Lara, A. Larson, and J. B. Losos. 2004. Genetic variation increases during biological invasion by a Cuban lizard. Nature 431:177–81.

Krehenwinkel, H., and D. Tautz. 2013. Northern range expansion of European populations of the wasp spider Argiope bruennichi is associated with global warming-correlated genetic admixture and population-specific temperature adaptations. Molecular Ecology 22:2232–48.

Lavergne, S., and J. Molofsky. 2007. Increased genetic variation and evolutionary potential drive the success of an invasive grass. Proceedings of the National Academy of Sciences of the United States of America 104:3883–8.

Lu, W., and J. Lazell. 1996. The voyage of the beetle. Natural History:36–39.

Lu, W., and P. A. Logan. 1994. Genetic variation in oviposition between and within populations of *Leptinotarsa decemlineata* (Coleoptera: Chrysomelidae). Annals of the Entomological Society of America:634–640.

McCauley, D., and W. Eanes. 1987. Hierarchical population structure analysis of the milkweed beetle, *Tetraopes tetraophthalmus* (Forster). Heredity 58:193–201.

Miller, M. R., J. P. Dunham, A. Amores, W. A. Cresko, and E. A. Johnson. 2007. Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. Genome Research 17:240–8.

Narum, S. R., C. A. Buerkle, J. W. Davey, M. R. Miller, and P. A. Hohenlohe. 2013. Genotyping-by-sequencing in ecological and conservation genomics. Molecular Ecology 22:2841–7.

Nei, M., T. Maruyama, and R. Chakraborty. 1975. The bottleneck effect and genetic variability in populations. Evolution 29:1–10.

Piiroinen, S., L. Lindström, A. Lyytinen, J. Mappes, Y. H. Chen, and V. Izzo. 2013. Pre-invasion history and demography shape the genetic variation in the insecticide resistance-related acetylcholinesterase 2 gene in the invasive Colorado potato beetle. BMC Evolutionary Biology 13:13.

Riquet, F., C. Daguin-Thiébaut, M. Ballenghien, N. Bierne, and F. Viard. 2013. Contrasting patterns of genome-wide polymorphism in the native and invasive range of the marine mollusc *Crepidula fornicata*. Molecular Ecology 22:1003–18.

Sakai, A., F. Allendorf, J. Holt, D. Lodge, J. Molofsky, K. With, S. Baughman, R. Cabin, J. Cohen, N. Ellstrand, D. McCauley, P. O'Neil, I. Parker, T. JN, and S. Weller. 2001. The population biology of invasive species. Annual Review of Ecology and Systematics 32:305–332.

Tsutsui, N. D., and T. J. Case. 2001. Population genetics and colony structure of the Argentine ant (*Linepithema humile*) in its native and introduced ranges. Evolution 55:976–985.

White, T. A., S. E. Perkins, G. Heckel, and J. B. Searle. 2013. Adaptive evolution during an ongoing range expansion: the invasive bank vole (*Myodes glareolus*) in Ireland. Molecular Ecology 22:2971–85.

Zehnder, G., L. Sandall, A. Tisler, and T. Powers. 1992. Mitochondrial DNA diversity among 17 geographic populations of *Leptinotarsa decemlineata* (Coleoptera: Chrysomelidae). Annals of the Entomological Society of America 85:234–240.

Zhang, J., J. Yang, Y. Li, N. Liu, and R. Zhang. 2012. Genetic relationships of introduced Colorado potato beetle *Leptinotarsa decemlineata* populations in Xinjiang, China. Insect Science 20:643–54.

Chapter Three

Phylogenomics of the Colorado potato beetle in North America

3.1 Introduction

Insect pests of crops typically display complex evolutionary histories because of their association with agriculture (Miller et al. 2005, Nardi et al. 2010). Characterizing ancestral and descendent relationships within a geographic context can provide insight into the evolutionary genetics associated with pest evolution. By reconstructing the history of a species, phylogenomic analysis can describe the relationship of pest insects with wild-type ancestors, and a phylogeographic perspective can help place this evolutionary history in a geographic context (e.g. Nardi et al. 2010). When coupled with population genetics, these analyses provide more complete information on the number of introductions and their origins.

The Colorado potato beetle, *Leptinotarsa decemlineata* (Say), is one of the most serious insect pests of potato (Hare 1990). Both adults and larvae feed on leaves, and can severely defoliate plants. Left uncontrolled, they can destroy entire potato fields. The beetle is notorious for its ability to develop resistance to insecticides; the modern pesticide industry effectively began with early attempts to control Colorado potato beetle (CPB) (Alyokhin et al. 2008). However, the beetle has repeatedly developed resistance to insecticides. Resistance levels can be very high and widespread, cross-resistance and multiple resistance are common, and populations of CPB have variously become resistant to every major class of insecticide (Alyokhin et al. 2008). It is an invasive pest, with a history of rapid spread. CPB first appeared as a pest of potato in 1859, in eastern Nebraska, and quickly spread to the East Coast (Tower 1906, Casagrande 1985). Today, CPB damages potato crops across the Northern Hemisphere, and continues to expand its range (Alyokhin et al. 2013).

While the spread of CPB as a pest of potato is very well documented, the origins of the beetle are somewhat less clear. There is considerable confusion over its origins and native range that can still be found in the literature—both popular and academic—to this day (for example, see Wikipedia entry). The most widely accepted and generally well supported history is that the beetle's native range was originally confined to Mexico (Tower 1906, Casagrande 1985), where it fed on wild solanaceous plants. Even now, CPB is not a pest in Mexico and beetles there typically resist ovipositing or even feeding on potatoes (Izzo et al. n.d., 2013, Lu and Logan 1994a, 1994b). Burrs of the primary native host, *Solanum rostratum*, likely clung to the fur of cattle and were transported northward as the grazing industry developing in northern Mexico in the mid 1500's quickly expanded northward (Tower 1906, Brand 1961). It has been assumed that the beetles subsequently followed their host north into the Central Plains and the eastern slopes of the Rocky Mountains. CPB first encountered the potato when European settlers migrated westward with the colonization of the American West. Northern populations of the beetle made the shift to this new plant sometime between 1820, when the earliest potato fields within that

early expanded range of wild CPB were documented, and 1859, when the first outbreaks on potato were reported (Tower 1906, Casagrande 1985). Following its first reported occurrence on potato, the range of CPB rapidly expanded across North America. The beetle reached the US East Coast by 1874 and overran the Northern States within three years. CPB had begun making incursions into Europe as early as 1875, and by 1900 it had spread across the entire eastern part of the USA and lower Canada (Tower 1906). Quarantine and eradication efforts in Europe were successful until 1922, when CPB finally became established in France. Its current range worldwide is likely more than the 16 million square kilometers in the Northern Hemisphere often cited from Weber (2003), as the expansion front in China, at least, is active (Zhang et al. 2012).

Despite considerable effort by the scientific community, control of this highly adaptable beetle continues to be very challenging. Understanding the origins of this invasive pest may provide insight into the genetic basis for its success on potato and across a diverse range of ecological conditions, and thus contribute to the control of CPB. Although it has been a well-studied insect, comparatively little is known about CPB at the genomic level. Additionally, while considerable effort has been made to examine the population genetics of beetles from the introduced range in the USA, beetle populations in the proposed native range have been much less studied.

We conducted a phylogenomic analysis of CPB in the native range of Mexico and the introduced USA range using a RAD-seq derived SNP data set. Using a genomic analysis may answer questions that are not fully resolved by traditional genetic methods (mSat and mtDNA) alone. While the first draft of a non-annotated CPB genome has been produced by the i5K Initiative (The 5,000 Insect Genome Project) (BCM-HGSC 2014a), and an annotated transcriptome by the Grapputo group (Kumar et al. 2014), our study was initiated before any reference genome was available. For an organism previously without a reference genome, a reduced-representation sequencing technique such as RAD-seq can be an effective way to produce genomic data from many individuals. We asked whether the data supported a single introduction event, or multiple introductions from the native range. We also sought to identify the population in the native range with which the introduced populations are most closely related, and thereby determine from which geographic region the introduced populations most likely originated.

3.2 Methods

Data collection

Colorado potato beetle genomic SNP data generated for Chapter Two were used as the starting data for a phylogenomic analysis. Briefly, 38 CPB were collected from 11 locations in the native range of Mexico and three locations in the introduced USA range (Table 2.1, Figure 2.1). Whole genomic DNA was extracted from the samples using the DNeasy Blood & Tissue Kit (Qiagen). Insect collection and DNA extractions were conducted by V. Izzo et al. A multiplexed *PstI*-digested RAD-tag library was prepared for paired-end sequencing from these samples using a protocol based on Etter and Johnson (2012) and Baxter et al. (2011). The library was sequenced on an Illumina Hi-Seq 2000 by the Delaware Biotechnology Institute's DNA Sequencing & Genotyping Center. We used *Stacks* 1.02 (Catchen et al. 2011, 2013) to process the raw first-read sequence reads and call SNPs in the identified sample RAD loci. Our previous discriminant analysis of principle components (DAPC) in *R* based on this SNP data set identified four clusters

that corresponded to broad geographic regions from which CPB were collected (Figure 2.5). Samples collected across the native Mexican range were identified as belonging to the Western Mexico, Southern lowland Mexico, or Central highland Mexico groups, and samples collected from the introduced range along the East Coast of the USA formed a single group.

SNP matrix preparation

SNP selection

For SNP selection, we exported a subset of RAD loci from the catalog of all loci produced the using *denovo_stacks* program in the *Stacks* pipeline. We began SNP selection with the set of RAD loci previously used for population genetic analysis. This comprised 4,542 RAD loci that were present in all four of the regional clusters previously identified by DAPC analysis, and present in at least 50% of the samples from each region. From these RAD loci we then identified 1,545 loci that contained 1-4 SNPs per locus, for a total of 4,540 SNPs. As initial SNP calling in *Stacks* was conducted using a more diverse set of samples, not all SNP sites were variable within this set of 38 samples. Therefore, a total of 1,418 SNP sites were removed as being invariant or invariant across all individuals except those that were heterozygous at the SNP. This selection process produced a SNP matrix of 3,122 variable sites (Figure 3.1). All SNP loci were concatenated together into a single alignment in Phylip format, with missing data (-) entered as needed for loci with incomplete representation across samples. All remaining heterozygous positions were converted to standard IUPAC ambiguity codes. A more conservative data set containing only SNPs present in 70% or more of the samples was also produced for additional maximum likelihood analysis (details below).

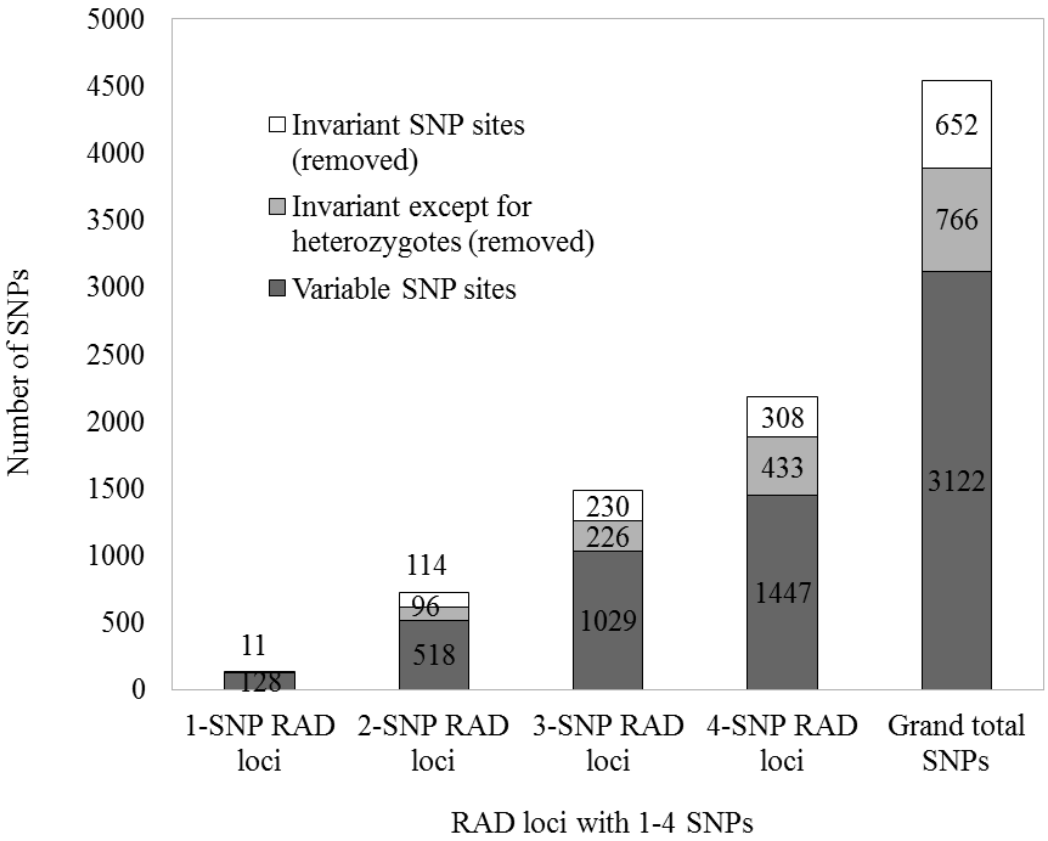

Figure 3.1: The number of SNPs at RAD loci originally containing 1, 2, 3 or 4 SNPs, before and after SNPs were filtered to remove invariant sites.

Heterozygous SNP haplotype resampling

Typically, phylogenetic analysis is conducted to understand deep evolutionary relationships. In such analyses, fixed differences between groups are the primary data. In analyses of closely related species or of closely related populations of diploid individuals within one species, ancestral polymorphisms and heterozygous loci can make up a significant proportion of the data. However, heterozygous loci are typically discarded in favor of fixed differences, or coded as ambiguous sites. In order to incorporate potentially informative phylogenetic data that is contained in heterozygous sites, we used the program *RRHS* (Repeated Random Haplotype Sampler) (Lischer et al. 2013). Using this program, we produce 1000 different SNP matrices where all ambiguous SNP sites in our input matrix were each replaced with a randomized selection of the two bases present in that individual at that SNP.

Phylogenomic analysis

Distance based trees

The *seqboot* program in *Phylip* 3.695 (Felsenstein 2013) was used to produce 1000 bootstrapped SNP datasets to allow calculation of bootstrap support values for distance. A pairwise genetic distance matrix was calculated for each bootstrapped dataset using *Phylip: dnadist*. We used a gamma coefficient of 1.096 under the F84 distance model, which uses two different change rates

for transitions and transversions, and allows equilibrium frequencies of bases to differ. Distance based neighbor joining and Fitch-Margoliash square root optimized trees were inferred from these matrices using *Phylip*. Replicate trees were combined to form neighbor joining and Fitch-Margoliash consensus trees using the *SumTrees* 3.3.1 program (Sukumaran and Holder 2010) in the *DendroPy* 3.12.0 package of *Python*.

Parsimony tree

The most parsimonious tree was selected by heuristic algorithm in *TNT* 1.1 (Goloboff et al. 2003, 2008), software sponsored by the Willi Hennig Society. First, the SNP data matrix was opened in *Mesquite* 2.75 (Maddison and Maddison 2011) and exported to *TNT* with gaps treated as missing. Then, default settings for "new technology searches" were used in *TNT*, with sectorial search, ratchet, and tree fusing for 100 random-addition searches. The resulting trees from this initial search were then subject to "traditional search" using TBR. This search found five tees of 4725 steps, from which a strict consensus tree was calculated in *TNT*. Standard Bootstrap was done in *TNT* for 1000 replicates using TBR.

Maximum likelihood trees

Two maximum likelihood trees were inferred from the data using RAxML 8.0.14 (Stamatakis 2014). First, an analysis was conducted using the full set of 3,122 SNPs. We used the GTR+ Γ nucleotide substitution model, with ascertainment bias to correct for the lack of invariant positions as recommended in the RAxML Manual 8. While invariant positions are typically used to help estimate model parameters in phylogenetic analysis, the ascertainment bias parameter in the new 8.0 version of RAxML makes the use of a SNP-only dataset less problematic. Bootstrap support was estimated from 1,000 replicate searches from random starting trees.

We also conducted a second more conservative analysis using the same parameters. For this analysis, we reduced the amount of missing data by excluding from the analysis all SNPs that were present in fewer than 70% of the samples, or 27 individuals. This left a matrix of 1,875 SNPs present across 38 individuals. While raising the coverage cutoff by 20%, this caused a 40% reduction in the number of SNPs from the full data set, and only resulted in a 4.9% decrease in the amount of missing data.

ML resampled haplotype consensus tree

Maximum likelihood trees were also inferred for each of our 1000 haplotype-resampled data sets using *RAxML* 8.0.14, with bootstrap support estimated from 100 replicate searches with random starting trees using the GTR+ Γ nucleotide substitution model with ascertainment bias. *RAxML* trees were combined using the *SumTrees* 3.3.1 to form a consensus tree.

Graphical presentation of trees

All trees were printed as midpoint rooted trees using *Figtree*. For all trees, branches with less than 25% support were collapsed, either by using *TreeGraph* (Stöver and Müller 2010) for parsimony, ML full, and ML conservative analyses, or by setting a minimum 25% consensus rule for inclusion of branches in trees produced by *SumTrees* for other analyses. Branch bootstrap support values greater than 50% are shown on figures.

3.3 Results

Sequence polymorphism

We identified a set of 3,122 polymorphic SNPs that were present in at least 50% of individuals from all four geographic regions. Sample representation across SNPs ranged from 21 to 36 samples represented at a given SNP loci, with the greatest number of SNPs present in 27 of the 38 samples (Figure 3.2). Approximately 60% of the SNPs were found in 70% or more of the individuals (Figure 3.3). In total, 27.9% of the data in our SNP matrix was missing. Of the SNPs included in the analysis, 1,610 (51.6%) were homozygous within all individuals and 1,512 (48.4%) were heterozygous in at least one individual (Figure 3.4). SNPs that were heterozygous in 1-3 individuals accounted for 89% of all the SNPs with heterozygosities (43.2% of all SNPs), with the remainder comprising SNPs that were heterozygous in 4-12 individuals. With the exception of the ML analysis based on resampled data, these heterozygous sites were included as ambiguous positions in analyses.

At the individual level, samples contained on average 2,252 SNPs, ranging from 980 to 3,047 SNPs (Figure 3.5). Both low and high coverage individuals were distributed across all regions. All individuals contained some heterozygous sites. The average number of heterozygous SNPs in an individual was 77.9, and ranged from 17 to 222 sites per individual (Figure 3.5).

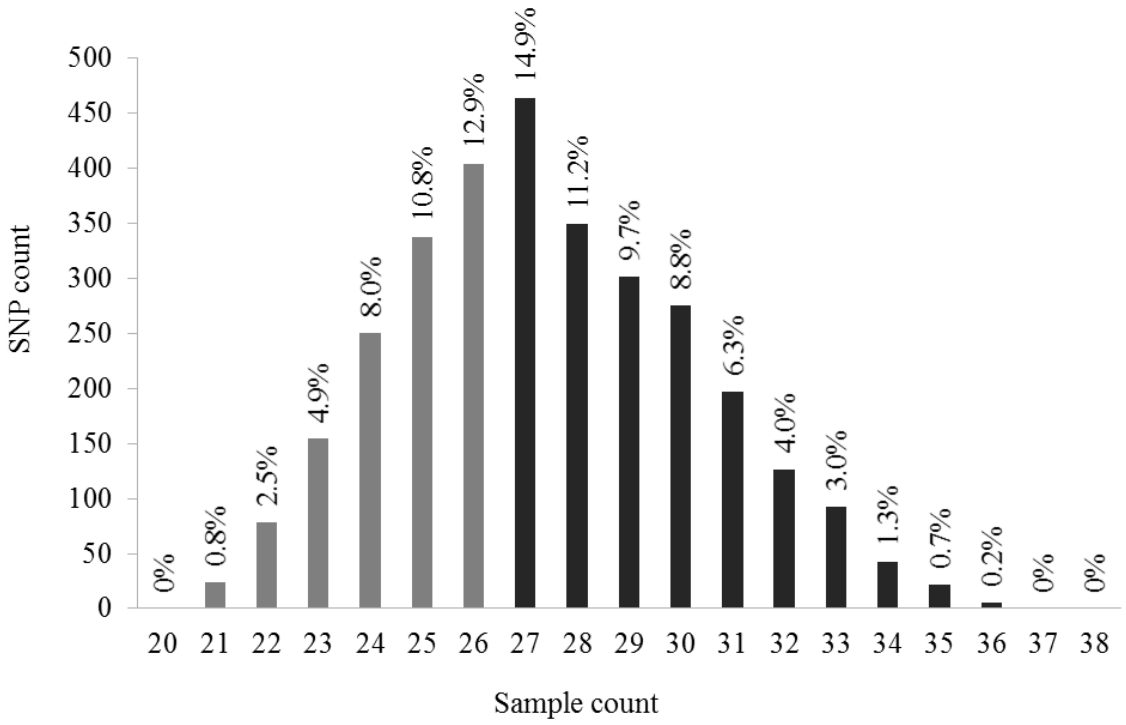

Figure 3.2: Sample coverage varied by SNP. Data labels indicate the percentage of the total SNPs made up by SNPs from each category of sample coverage. Grey bars indicate sites that were removed from SNP matrix for conservative ML phylogenomic analysis.

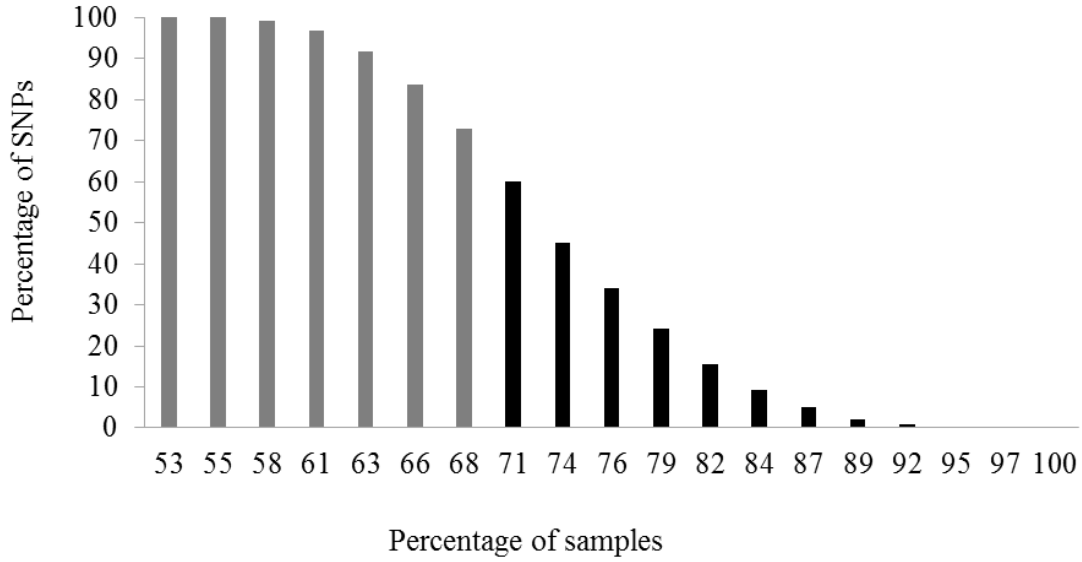

Figure 3.3: Tail distribution of minimum sample coverage for SNPs. Grey bars indicate SNP sites that were removed from SNP matrix for conservative ML phylogenomic analysis.

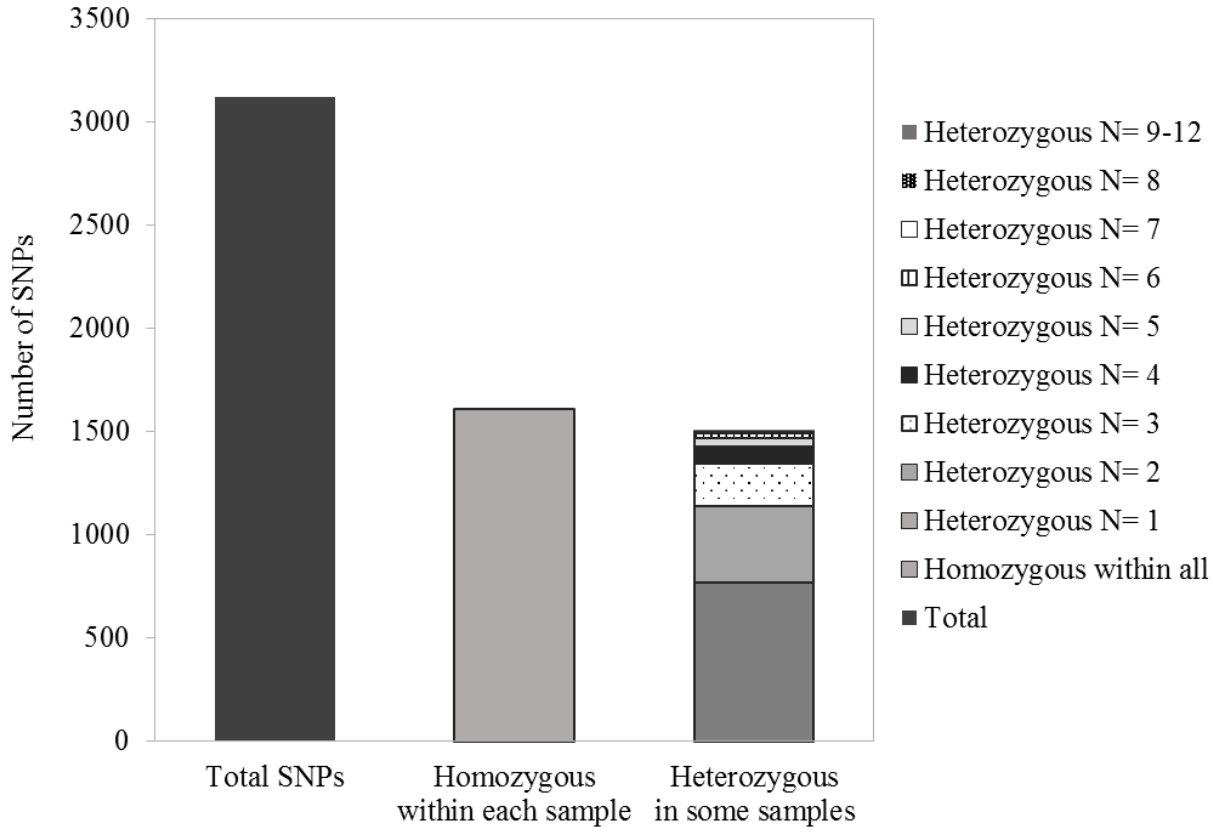

Figure 3.4: Heterozygosity on a per-SNP basis. N = the number of individuals in which a given SNP was heterozygous.

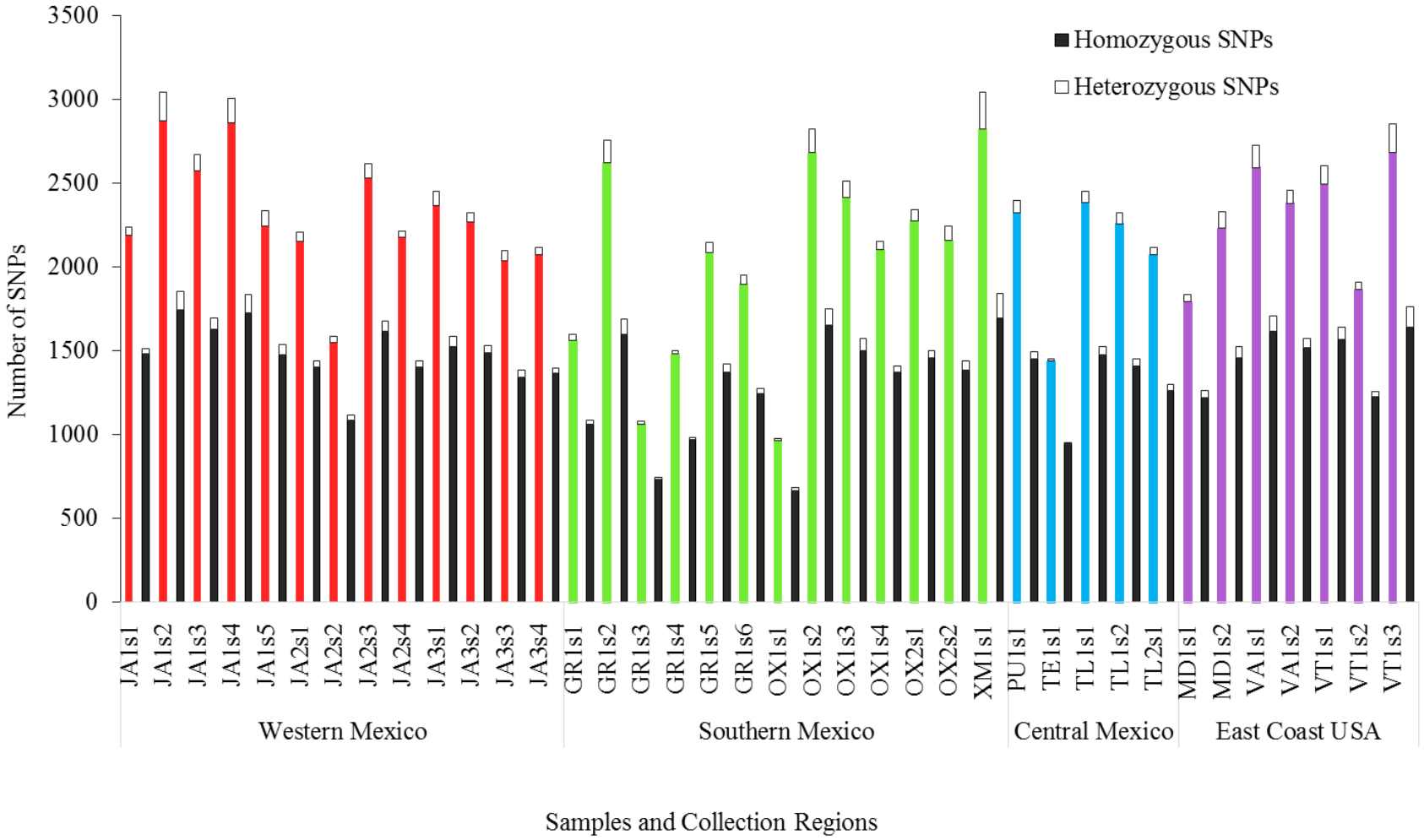

Figure 3.5: Total number of SNPs per individual, and the abundance of heterozygous SNPs in each sample. Colored bars (left) indicate homozygous SNPs used in full analyses, black bars (right) indicate homozygous SNPs used in conservative ML analysis, and open bars (top) indicate the respective number of heterozygous SNPs in individual samples, which were re-sampled in RRHS analysis.

Phylogenomic analysis

All phylogenomic analyses (Figures 3.6 to 3.11) of the Colorado potato beetle data produced consistent inferences about the major topological features of the tree. In all analyses, the East Coast USA samples (MD, VA, VT) formed a single cluster with 100% support for the branch separating this group from the rest of the tree. Internal branches had low support. Individuals collected from the central highland region of Mexico (PU, TE, TL) formed a single group with the following exceptions. In both the neighbor joining (Figure 3.6) and the Fitch-Margoliash distance trees(Figure 3.7), the sample from Puebla was placed immediately outside the Central Mexico cluster. In the Fitch-Margoliash tree, one Tlaxcala sample also was placed outside the Central Mexico cluster. Branches within this group had low to moderate levels of support. However, the branch separating this group from others had high support, ranging from 65% to 100% across different methods. An exceptionally long branch separated all East Coast USA and Central Mexico samples from the native range samples in the rest of the tree. This branch had 100% support in all methods.

Samples from Western Mexico (JA) formed a single cluster under all methods, separated from other samples by a branch with 100% support. Specific placement of individuals within the Western Mexico cluster was inconsistent, with low branch support in all trees. The topology of the Southern Mexico group (GR, OX, XM) was least consistent between algorithms, generally forming a paraphyletic group of undifferentiated polytomies. Samples from Oaxaca and Guerrero tended to differentiate from each other, but with inconsistent topology and low support. Combining random haplotype resampling with maximum likelihood, however, recovered some internal nodes in the Southern Mexico group with moderate (56-79%) support (Figure 3.11).

Reducing the amount of missing data by approximately 5% somewhat reduced the length of the branch between the USA/Central Mexico and Southern/Western Mexico groups (Figure 3.10) relative to ML analysis based on the full SNP matrix (Figure 3.9). Support for most branches was also reduced by the loss of data required to achieve the 5% reduction in missing data. Haplotype resampling did not change tree topology relative to the ML tree using the full data set, but did increase support values for internal branches in the East Coast USA and Central highland Mexico clusters and the Southern Mexico cluster (Figure 3.11).

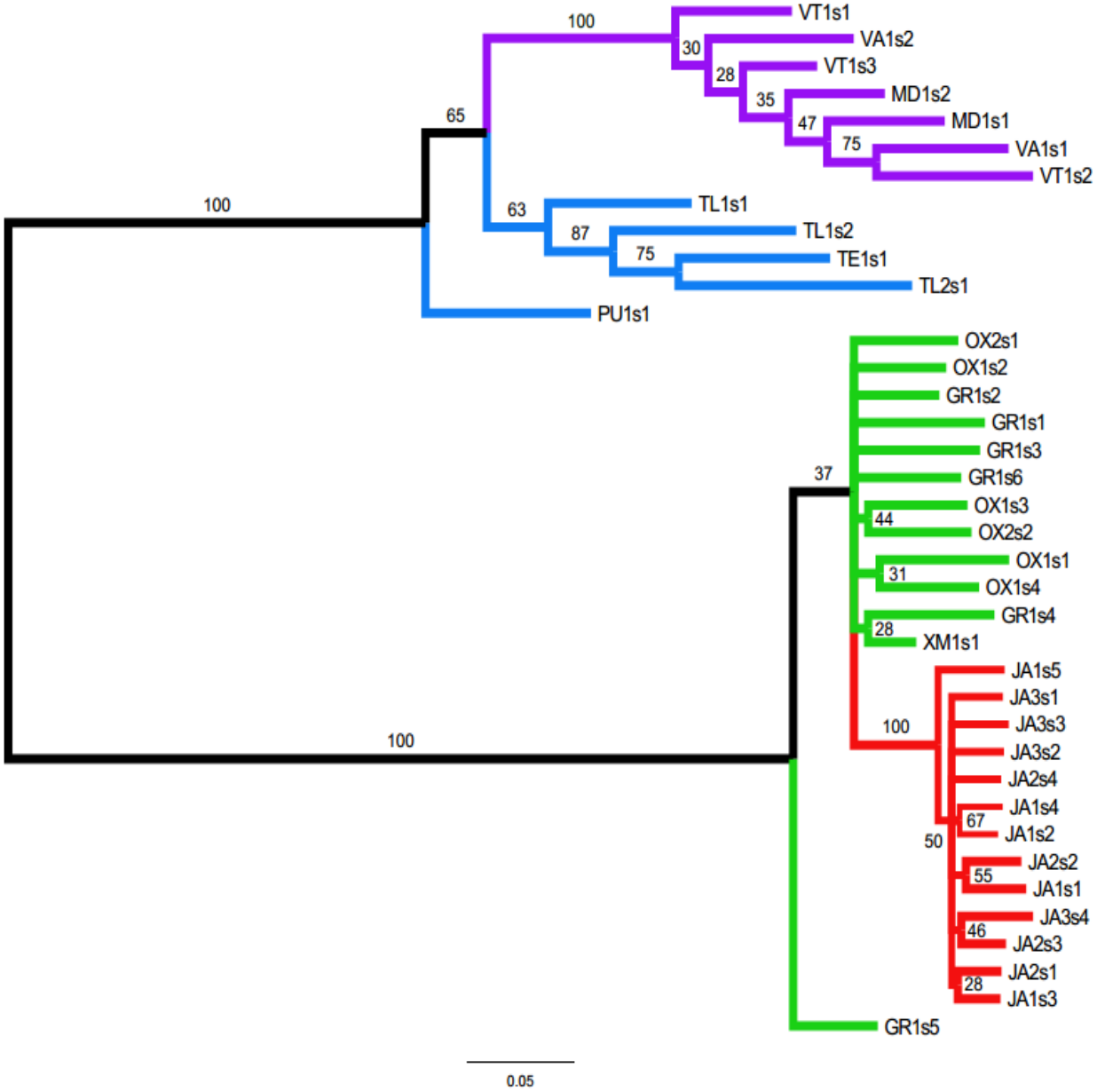

Figure 3.6: Midpoint rooted neighbor joining tree. Branches are colored by collection region. Branch coloration corresponds to geographic regional populations identified by discriminant analysis (Chapter 2) as follows: purple, East Coast USA; turquoise, Central highland Mexico; green, Southern lowland Mexico; red, Western Mexico.

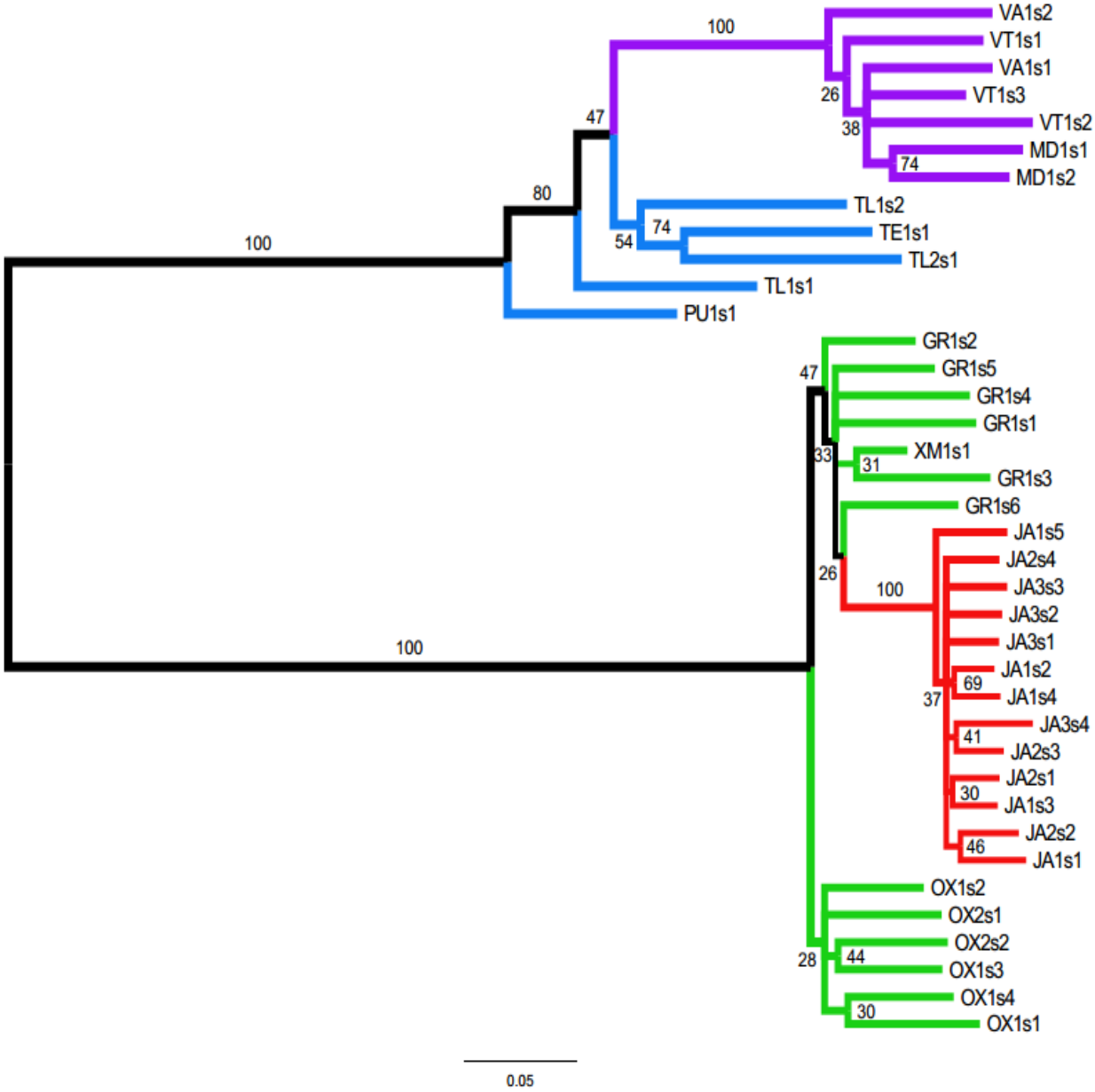

Figure 3.7: Midpoint rooted Fitch-Margoliash distance tree. Branches are colored as in Figure 3.6.

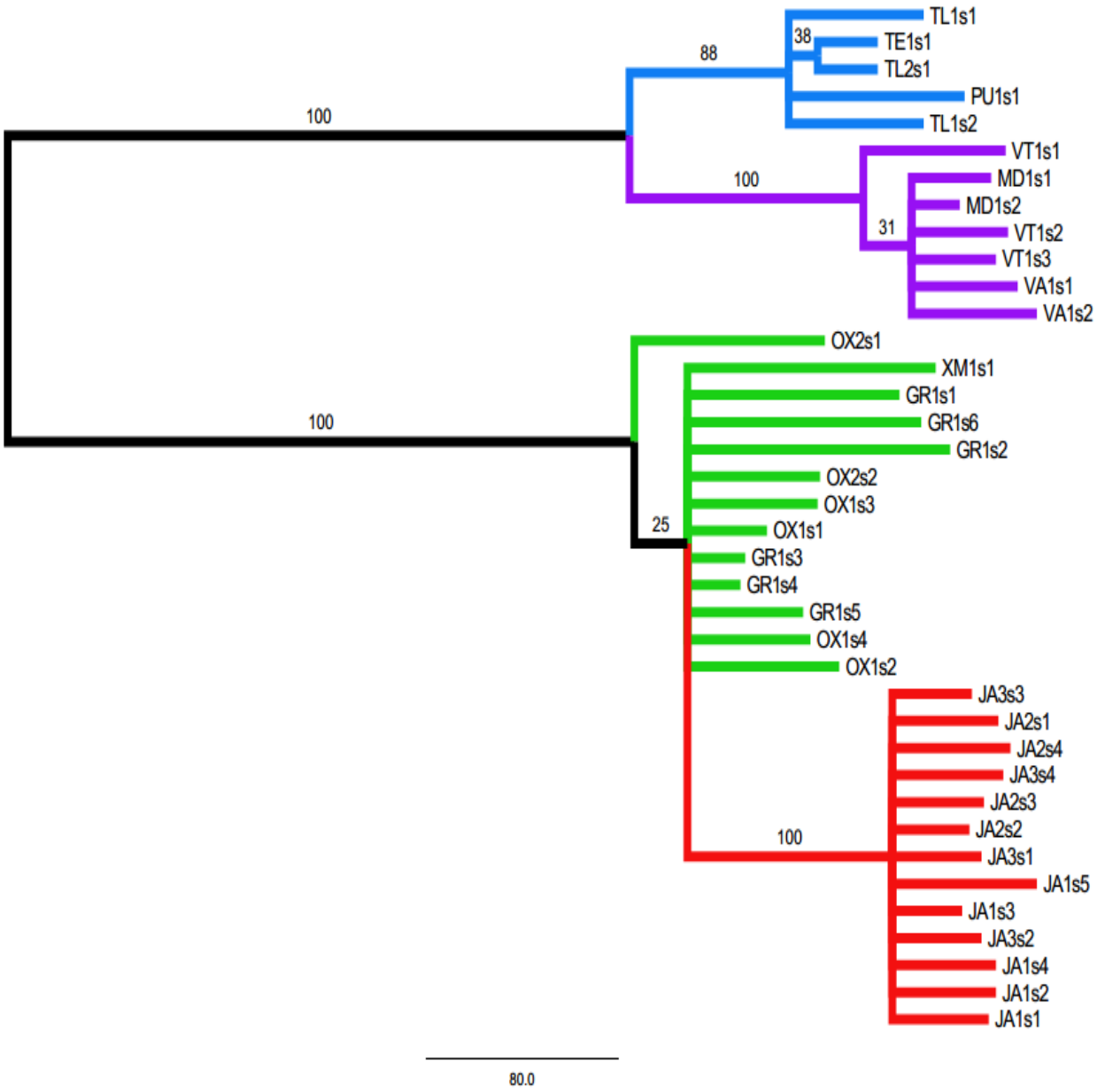

Figure 3.8: Midpoint rooted maximum parsimony tree. Branches are colored as in Figure 3.6.

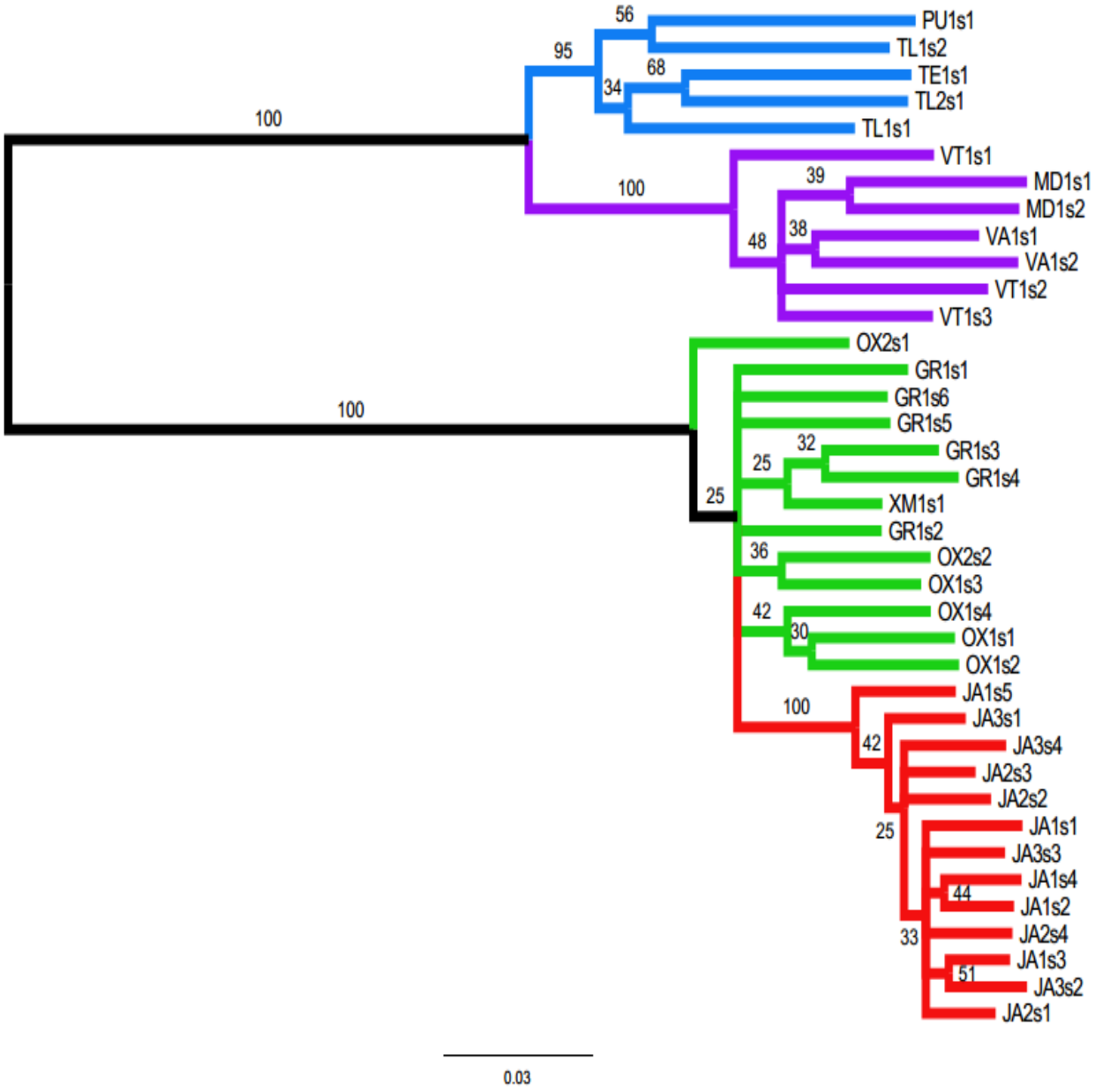

Figure 3.9: Midpoint rooted maximum likelihood (full data set) tree. Branches are colored as in Figure 3.6.

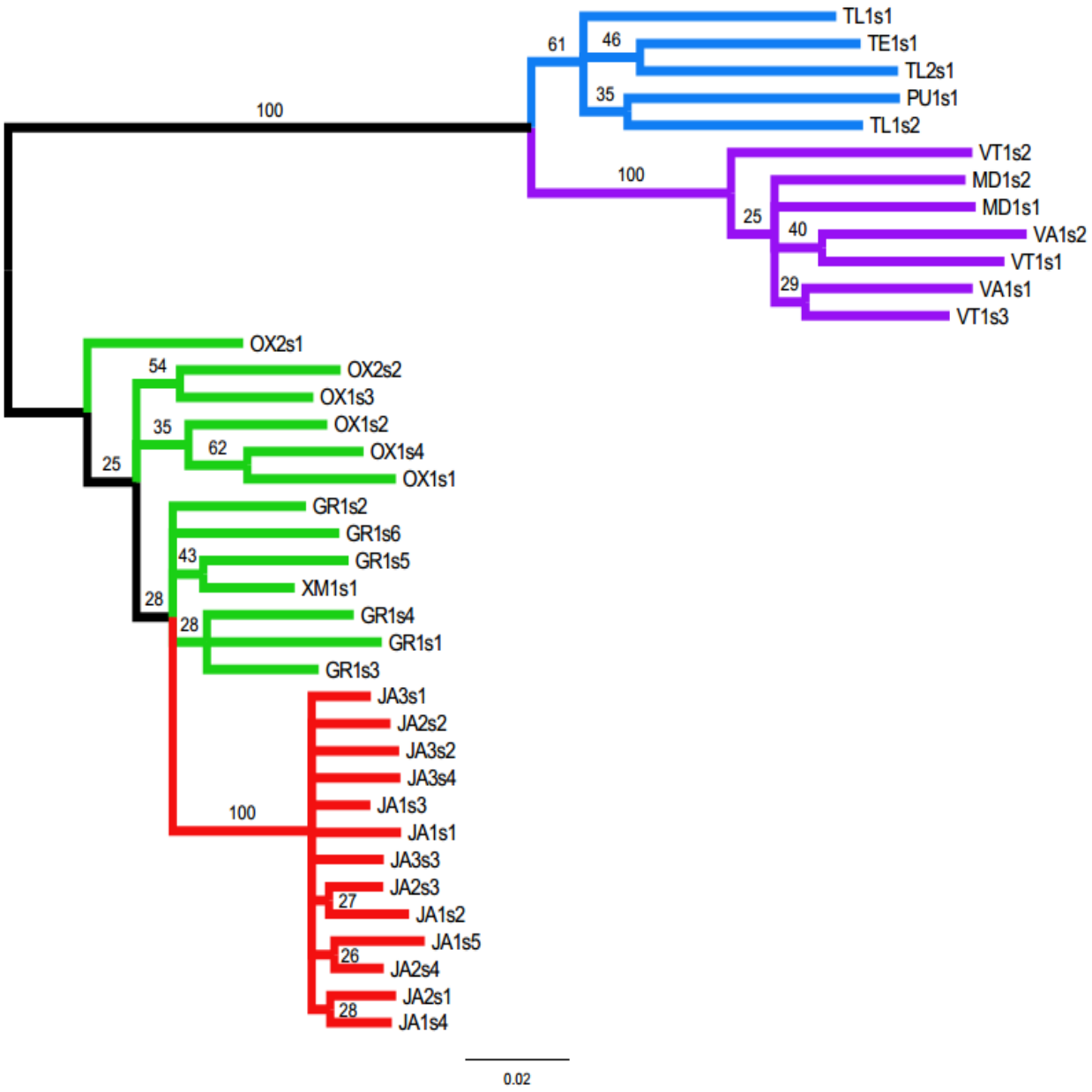

Figure 3.10: Midpoint rooted ML (conservative) tree. Branches are colored as in Figure 3.6.

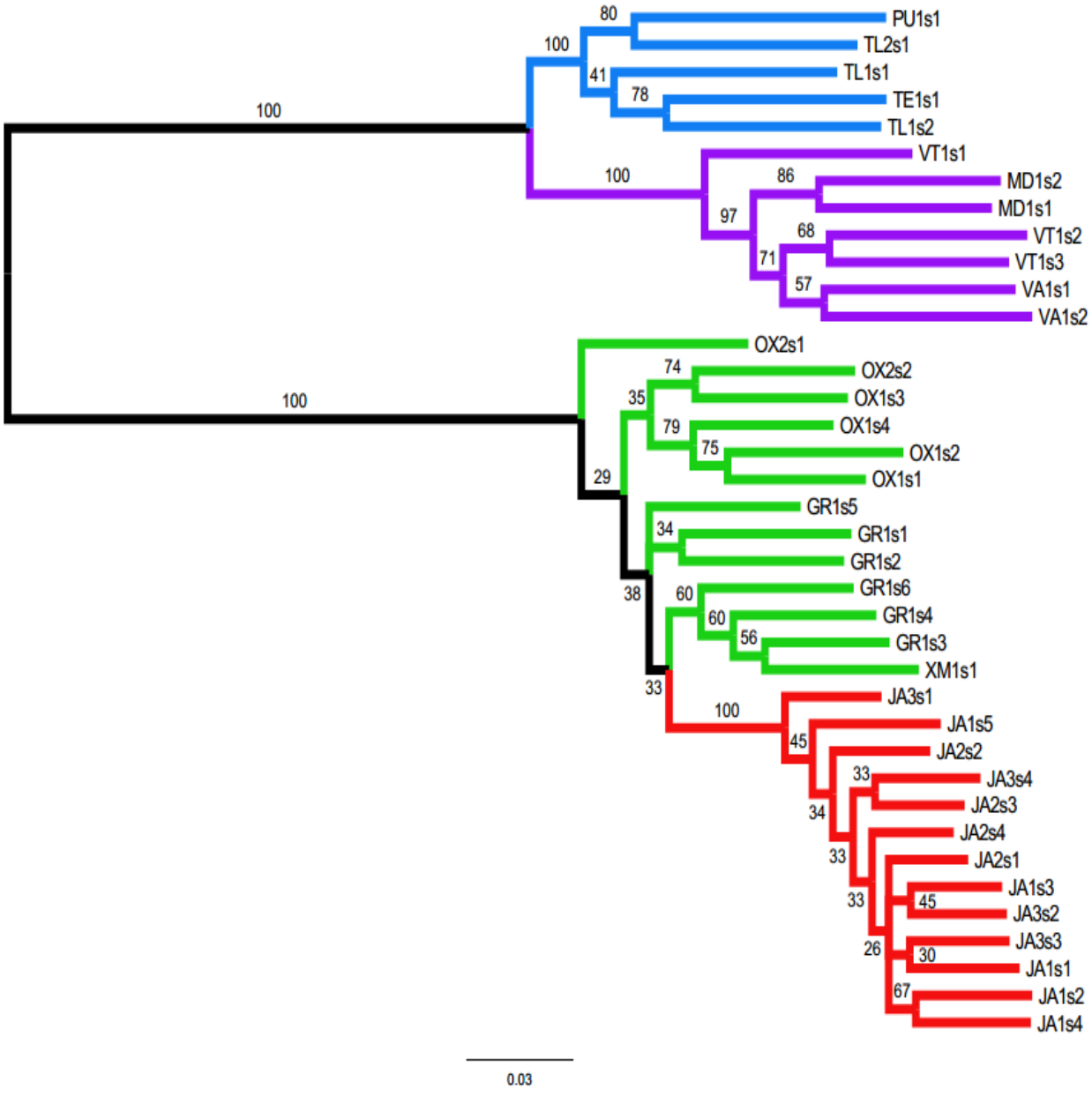

Figure 3.11: Midpoint rooted ML haplotype resampling tree. Branches are colored as in Figure 3.6.

3.4 Discussion

Our phylogeographic analysis confirmed the geographic clustering found in our population genetics analysis (Chapter 2). Discriminant analysis identified four genetic clusters (Figure 2.5), three of which were supported as separate clusters by all of the phylogenetic algorithms we tested, with the fourth forming a separate but paraphyletic cluster (Figures 3.6-3.11). Tree topology was also consistent with genetic differentiation based on F_{ST} values and DAPC in terms of both the close pairwise relationships between the USA and Central Mexico branches and the Southern and Western Mexico branches, and the separation between these two groups.

One of the principle reasons to conduct a phylogeographic analysis of CPB was to identify the region or regions of origin of the introduction. Based on the short branch lengths between the Central Mexico and East Coast USA groups, and the long branch separating these groups from the rest of the tree, the most likely native source of the introduced US beetles is the Central highlands of Mexico. Unfortunately, we cannot be certain of this interpretation without the use of an outgroup with which to polarize the tree. The Asian longhorn beetle, *Anoplophora glabripennis*, has just been sequenced by the i5K Initiative, and may be useful for this purpose (BCM-HGSC 2014b). However, a Central Mexico origin for the introduced beetles is consistent with research conducted by Izzo et al. (2013), who found that beetles from this region of Mexico exhibit ecological and behavioral traits more similar to US beetles than do beetles from other parts of the native range.

The single cluster recovered for the samples collected from potato in the introduced range suggests that the invasive beetles originated from a single introduction event. At the very least, any repeated introductions are most likely to have come from within the same geographic region rather than from across the breadth of the native range. The alternative hypothesis, of multiple introductions from different regions, would instead have produced a phylogeny wherein introduced beetles did not form a monophyletic group, but rather were scattered across the tree.

Haplotype resampling did not significantly change tree topology relative to the basic full maximum likelihood analysis. This is consistent with the work of Lischer et al. (2013), who found differences in convergence time estimates but not differences in tree topology when using the method for vole phylogenomics. Our data also had only a fraction of the heterozygosity observed in the vole data, so limited changes due to haplotype resampling are not unexpected. However, branch support was significantly higher for a number of internal nodes reconstructed using the resampled data sets. This suggests that some useful phylogenetic signal was in fact retrieved by this method. The lack of change seen in the Western Mexico group suggests that, while the frequency of heterozygous SNPs was fairly consistent across regions, the pattern of heterozygosity may have been different in these samples. For example, Western Mexico samples may have shared with each other the same alternative alleles at the same SNP sites more often than were shared within other regions.

Low resolving power of internal branches within the four geographic clusters may have several causes, none of which are mutually exclusive. First, there may be little phylogenetic signal at the local scale, for example due to sampled individuals being close relatives or simply a lack of phylogenetically informative mutations. Alternatively, there may be phylogenetic signal but it is relatively weak. The recovery of weak phylogenetic signal depends greatly on the amount of background noise in the data set. A third possible factor in low branch support, then, is the presence of high levels of noise in the data, due to errors in sequencing, genotyping, and SNP calling. While there is no lower bound for the number of samples included in a

phylogenetic analysis in order to recover strong phylogenetic signal from background noise, recent studies indicate that the use of many samples, even at low coverage, can improve the accuracy of genotyping and SNP calling in the face of the higher error rates produced by next generation sequencing relative to Sanger sequencing (Pool et al. 2010, Fumagalli 2013). Therefore, the best chance of resolving more nodes is by minimizing the amount of error in the data set. This would most likely be best accomplished using a larger set of samples than were available for this analysis.

One interesting feature of our inferred tree topology is the long branch separating the USA and Central Mexico clusters from the rest of the tree. With a parsimony branch length of more than 500 changes, this branch is nearly 10 times the average branch length for the rest of the tree. This is a significant finding that should be further investigated. While missing data can produce long branches in phylogenies (Lemmon et al. 2009, but see also Wiens 2006, Roure et al. 2013), and systematically missing data is often a feature of RAD-seq data due to null-allele dropout caused by mutations at restriction enzyme recognition sites (Arnold et al. 2013, Davey et al. 2013, Gautier et al. 2013), this does not mean that these data are necessarily less informative regarding potentially important differences between CPB populations. RAD loci that are present in USA and Central Mexico individuals but missing from other individuals, or the reverse, may still point out genomic regions that could be important in differentiating invasive and non-invasive genotypes. Furthermore, restricting the analysis to those SNPs present in a higher percentage of samples (70% rather than 50%), did not significantly alter either the topology or the length of the branch between the two main nodes (Figure 3.10). This suggests that the long branch is not an artifact of missing data. In fact, early work by Hsiao and Hsiao (Hsiao and Hsiao 1982, 1983, Hsiao 1985) identified a stable pericentric inversion on the No. 2 chromosomes of some populations of CPB. An inversion could explain the long branch identified in our analysis. The metacentric version of the chromosome, without the inversion, was identified in the sampled populations from Mexico, as well as in some southwestern US populations. The inversion was found in USA samples both as homozygous and heterozygous forms, but not in samples from Mexico (Hsiao 1985). If this inversion proves to be the cause of the long branch, it would suggest that the mutation occurred earlier, and further south, than was hypothesized by Hsiao.

Of practical significance is the identification of genes that may aid scientists in the development of control programs for this important crop pest. In addition to those alleles that produce the long branch and thus may be worth investigating as potential pre-adaptations in the historic CPB genotype, alleles that separate USA samples from the native-range clusters may be particularly important in the current success of introduced CPB in the agricultural setting. By comparing the highlighted RAD loci with the annotated genome that will soon be completed, it may be possible to identify the function of some of the genes marked by these SNPs. Such genes could then be explored for roles in a variety of potentially important differences between wild type and pest beetles, including host recognition and acceptance, cold tolerance and diapause behavior, and pesticide resistance.

The high genetic diversity of introduced beetles (Chapter 2) seems to be a feature unique to the invasive group. If high diversity was due to multiple origins, as described above, we would not expect to see the clearly monophyletic grouping of introduced beetles. Rather than this high within-group diversity being a sum of the diversity found across the native range, it is likely the result of processes acting during and after invasion. There are several possible hypotheses that may contribute to this pattern. For example, hybridization or introgression with an unidentified species may have contributed to the genomic diversity of these beetles. However, this scenario

seems unlikely based on the close relationship inferred between the introduced East Coast USA beetles and the native Central highland Mexico beetles. Bayesian analysis of demographic parameters could also provide an indication of strictly demographic processes that may have influenced diversity in introduced CPB. However, as most demographic forces operating during an invasion tend to reduce genetic diversity, this alone is unlikely to resolve the question. Another possibility is that exposure of large populations to varying selection in the agricultural setting has facilitated the generation and maintenance of genetic diversity in introduced beetles. Examination of more individuals from different geographic populations in the introduced range that have historically experienced different climatic and agricultural management conditions may provide insight into this possibility. It is also possible that one of the genetic changes in the introduced range is in some way explicitly linked with regulation of the rate of mutation in pest CPB. A thorough genome annotation would be required to even begin exploration of such a possibility. Regardless of the source of high genomic diversity in introduced CPB, it is likely that this diversity contributes to CPB's success as an invasive crop pest.

In conclusion, our data places the three genetic groups previously identified in the native range in two distinct clusters and a third paraphyletic cluster, and indicates that the invasive pest beetles are derived from a single introduction that may have originating in the Central highlands of Mexico. Based on the presence of a long branch separating Central Mexico individuals from the rest of the native range individuals, some genetic differentiation may have occurred in Mexico prior to the host shift of CPB. Future studies focusing on genetic differences between the two main branches of the phylogeny may identify genes that contribute to invasiveness in CPB, while studies that focus on the genetic characteristics that unite the East Coast USA samples may identify genes that have evolved in direct response to selection in the agricultural setting. Our phylogenomic analysis of Colorado potato beetle, and the associated genomic data, thus make a significant contribution to our understanding of the history of CPB, and provide a powerful new resource for understanding the genetic underpinnings of the success of this critical agricultural pest.

3.5 Literature Cited

Alyokhin, A., M. Baker, D. Mota-Sanchez, G. Dively, and E. Grafius. 2008. Colorado potato beetle resistance to insecticides. American Journal of Potato Research 85:395–413.

Alyokhin, M. Udalov, and G. Benkovskaya. 2013. Chapter 2-The Colorado Potato Beetle. Pages 11–17 *in* A. Alyokhin, C. Vincent, and P. Giordanengo, editors. Insect Pests of Potato: Global Perspective on Biology and Management.

Arnold, B., R. B. Corbett-Detig, D. Hartl, and K. Bomblies. 2013. RADseq underestimates diversity and introduces genealogical biases due to nonrandom haplotype sampling. Molecular Ecology 22:3179–90.

Baxter, S. W., J. W. Davey, J. S. Johnston, A. M. Shelton, D. G. Heckel, C. D. Jiggins, and M. L. Blaxter. 2011. Linkage mapping and comparative genomics using next-generation RAD sequencing of a non-model organism. PloS One 6:e19315.

BCM-HGSC. 2014a. Colorado Potato Beetle Genome Project. https://www.hgsc.bcm.edu/arthropods/colorado-potato-beetle-genome-project.

BCM-HGSC. 2014b. Asian Long-horned Beetle Genome Project. https://www.hgsc.bcm.edu/arthropods/asian-long-horned-beetle-genome-project.

Brand, D. 1961. The early history of the range cattle industry in northern Mexico. Agricultural History 35:132–139.

Casagrande, R. 1985. The Iowa Potato Beetle, its Discovery and Spread to Potatoes. Bulletin of the Entomological Society of America 31:27–29.

Catchen, J. M., A. Amores, P. A. Hohenlohe, W. A. Cresko, and J. H. Postlethwait. 2011. Stacks: building and genotyping Loci de novo from short-read sequences. G3 1:171–82.

Catchen, J. M., P. A. Hohenlohe, S. Bassham, A. Amores, and W. A. Cresko. 2013. Stacks: an analysis tool set for population genomics. Molecular Ecology 22:3124–40.

Davey, J. W., T. Cezard, P. Fuentes-Utrilla, C. Eland, K. Gharbi, and M. L. Blaxter. 2013. Special features of RAD Sequencing data: implications for genotyping. Molecular Ecology 22:3151–64.

Etter, P. D., and E. A. Johnson. 2012. RAD Paired-end sequencing for local de novo assembly and SNP discovery in non-model organisms. Pages 135–151 *in* F. Pompanon and A. Bonin, editors. Data Production and Analysis in Population Genomics. Humana Press, Totowa, NJ.

Felsenstein, J. 2013. Phylip (Phylogenetic Inference Package). Distributed by the author. Department of Genome Sciences, University of Washington, Seattle.

Fumagalli, M. 2013. Assessing the effect of sequencing depth and sample size in population genetics inferences. PloS One 8:e79667.

Gautier, M., K. Gharbi, T. Cezard, J. Foucaud, C. Kerdelhué, P. Pudlo, J.-M. Cornuet, and A. Estoup. 2013. The effect of RAD allele dropout on the estimation of genetic variation within and between populations. Molecular Ecology 22:3165–78.

Goloboff, P. A., J. S. Farris, and K. C. Nixon. 2008. TNT, a free program for phylogenetic analysis. Cladistics 24:774–786.

Goloboff, P. A., S. Farris, and K. Nixon. 2003. TNT: Tree analysis using New Technology. Program and documentation available from the authors, and at www.zmuc.dk/public/phylogeny, Tucuman, Argentina.

Hare, J. D. 1990. Ecology and management of the Colorado potato beetle. Annual Review of Entomology:81–100.

Hsiao, C., and T. H. Hsiao. 1982. Inheritance of three autosomal mutations in the Colorado potato beetle Leptinotarsa decemlineata (Coleoptera: Chrysomelidiae). Canadian Journal of Genetics and Cytology 24:681–686.

Hsiao, T. H. 1985. Ecophysiological and genetic aspects of geographic variations of the Colorado potato beetle. Research Bulletin-Massachusetts Agricultural Experiment Station:63–77.

Hsiao, T. H., and C. Hsiao. 1983. Chromosomal analysis of Leptinotarsa and Labidomera species (Coleoptera: Chrysomelidae). Genetica 60:139–150.

Izzo, V. M., D. J. Hawthorne, and Y. H. Chen. 2013. Geographic variation in winter hardiness of a common agricultural pest, Leptinotarsa decemlineata, the Colorado potato beetle. Evolutionary Ecology 28:505–520.

Izzo, V. M., D. J. Hawthorne, S. D. Schoville, and Y. H. Chen. (n.d.). No Title.

Kumar, A., L. Congiu, L. Lindström, S. Piiroinen, M. Vidotto, and A. Grapputo. 2014. Sequencing, de novo assembly and annotation of the Colorado potato beetle, Leptinotarsa decemlineata, transcriptome. PLoS One 9:e86012.

Lemmon, A. R., J. M. Brown, K. Stanger-Hall, and E. M. Lemmon. 2009. The effect of ambiguous data on phylogenetic estimates obtained by maximum likelihood and Bayesian inference. Systematic Biology 58:130–45.

Lischer, H., L. Excoffier, and G. Heckel. 2013. Ignoring heterozygous sites biases phylogenomic estimates of divergence times: implications for the evolutionary history of Microtus voles. Molecular Biology and Evolution:1–42.

Lu, W., and P. A. Logan. 1994a. Geographic variation in larval feeding acceptance and performance of Leptinotarsa decemlineata (Coleoptera: Chrysomelidae). Annals of the Entomological Society of America:460–469.

Lu, W., and P. A. Logan. 1994b. Genetic variation in oviposition between and within populations of Leptinotarsa decemlineata (Coleoptera: Chrysomelidae). Annals of the Entomological Society of America:634–640.

Maddison, W. P., and D. R. Maddison. 2011. Mesquite: a modular system for evolutionary analysis. http://mesquiteproject.org.

Miller, N., A. Estoup, S. Toepfer, D. Bourguet, L. Lapchin, S. Derridj, K. S. Kim, P. Reynaud, L. Furlan, and T. Guillemaud. 2005. Multiple transatlantic introductions of the western corn rootworm. Science (New York, N.Y.) 310:992.

Nardi, F., A. Carapelli, J. L. Boore, G. K. Roderick, R. Dallai, and F. Frati. 2010. Domestication of olive fly through a multi-regional host shift to cultivated olives: comparative dating using complete mitochondrial genomes. Molecular Phylogenetics and Evolution 57:678–86.

Pool, J. E., I. Hellmann, J. D. Jensen, and R. Nielsen. 2010. Population genetic inference from genomic sequence variation. Genome Research 20:291–300.

Roure, B., D. Baurain, and H. Philippe. 2013. Impact of missing data on phylogenies inferred from empirical phylogenomic data sets. Molecular Biology and Evolution 30:197–214.

Stamatakis, A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics (Oxford, England):2010–2011.

Stöver, B. C., and K. F. Müller. 2010. TreeGraph 2: combining and visualizing evidence from different phylogenetic analyses. BMC Bioinformatics 11:7.

Sukumaran, J., and M. T. Holder. 2010. DendroPy: a Python library for phylogenetic computing. Bioinformatics (Oxford, England) 26:1569–71.

Tower, W. 1906. An investigation of evolution in chrysomelid beetles of the genus Leptinotarsa. Carnegie Institution, Washington, D.C.

Weber, D. 2003. Colorado beetle: pest on the move. Pesticide Outlook 14:256.

Wiens, J. J. 2006. Missing data and the design of phylogenetic analyses. Journal of Biomedical Informatics 39:34–42.

Zhang, J., J. Yang, Y. Li, N. Liu, and R. Zhang. 2012. Genetic relationships of introduced Colorado potato beetle Leptinotarsa decemlineata populations in Xinjiang, China. Insect Science 20:643–54.