EASTERN WHITE PINE DIEBACK: CURRENT COLONIZATION, BIOGEOGRAPHIC HISTORY, AND GENETIC DIVERSITY OF *MATSUCOCCUS MACROCICATRICES* AND ITS HOST

by

THOMAS EDWARD DANTAS WHITNEY

(Under the Direction of Kamal J.K. Gandhi)

ABSTRACT

Eastern white pine (*Pinus strobus* L.) is experiencing novel symptoms of dieback and mortality throughout its range in North America. Recent studies have found that two organisms, a fungal pathogen (*Caliciopsis pinea* Peck) and a scale insect (*Matsucoccus macrocicatrices* Richards), have emerged in association with a canker disease, which is driving the symptoms of this dieback phenomenon. Little is known about this insect-pathogen complex and its relationship to its host, including the colonization tendencies of the insect and pathogen, the nativity of the insect, and the genetic diversity and structure of the host in the Southern Appalachians region. I quantified the incidence of scale insects and pathogenic cankers within trees and among size classes and found that colonization of both the insect and the pathogen was consistent with the patterns of dieback and mortality of *P. strobus*, adding support to the hypothesis that these organisms are responsible for dieback symptoms and mortality. I also developed microsatellite markers to evaluate the present genetic diversity and structure of *M. macrocicatrices* populations within a biogeographical framework. I found that this insect

is native to all regions in North America, and I posit it has likely been co-occurring with *P. strobus*, its obligate host, since the last glacial maximum. Therefore, recent introductions and invasional processes likely did not contribute to the recent emergence of *M. macrocicatrices* and its novel association with pathogens and dieback symptoms. Lastly, I elucidated the genetic diversity and population structure of *P. strobus* in the understudied Southern Appalachians and found that populations in this region are highly variable but not more so than those in more northern latitudes. Although this important tree species faces new threats, its high genetic variation suggests it may also have a substantial amount of adaptive potential. My results and interpretations herein provide comprehensive and important foundational knowledge about the organisms involved in the eastern white pine dieback phenomenon. This research has built a framework for future investigations into (1) the mechanism(s) that allows *M. macrocicatrices* to facilitate *C. pinea* infection, (2) the reason(s) why *M. macrocicatrices* has recently emerged as a pest, and (3) the appropriate management and conservation strategies for *P. strobus* persistence in its native range.

INDEX WORDS: *Caliciopsis pinea*, Caliciopsis canker, eastern white pine bast scale, genetic diversity, microsatellites, *Pinus strobus*, population structure

EASTERN WHITE PINE DIEBACK: CURRENT COLONIZATION, BIOGEOGRAPHIC HISTORY, AND GENETIC DIVERSITY OF *MATSUCOCCUS MACROCICATRICES* AND ITS HOST

by

THOMAS EDWARD DANTAS WHITNEY

B.S., Gonzaga University, 2011

M.S., University of Kentucky, 2014

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2019

EASTERN WHITE PINE DIEBACK: CURRENT COLONIZATION, BIOGEOGRAPHIC HISTORY, AND GENETIC DIVERSITY OF *MATSUCOCCUS MACROCICATRICES* AND ITS HOST

by

THOMAS EDWARD DANTAS WHITNEY

Major Professor: Kamal J.K. Gandhi

Committee: Rima D. Lucardi
Joseph Nairn
Nathan Havill
Kenneth Ross

Electronic Version Approved:

Suzanne Barbour
Dean of the Graduate School
The University of Georgia
May 2019

ACKNOWLEDGEMENTS

I could not have reached this point without an incredible cast of people who supported me these last five years. First, let me give thanks to the countless individuals who sent me samples from near and far. This research depended heavily on their efforts, and I make sure to recognize each person by name later in the appropriate chapters. A few people deserve special recognition for their continuous help. I would like to especially thank Brittany Barnes for her support, assistance, and delightful presence throughout my time at UGA. She is an essential figure in the Gandhi Lab and was no less important to my individual success. I am also grateful to Michelle Cram for donating her time, resources, and energy to help me explore the curious nature of *M. macrocicatrices* for the first time. I would also like to thank my committee members for their continued support throughout. Ken Ross often encouraged me to consider all facets of the insect's biology and pushed back his retirement just for me (☺); Joe Nairn helped me extensively in the contemplation and planning of microsatellite development; and Nathan Havill shared DNA extraction protocols, gave me additional advice on molecular methods, and graciously donated his time to review a manuscript draft.

I would next like to thank my amazing co-advising dynamic duo. First, Kamal Gandhi has been so supportive and encouraging throughout my tenure. She has helped me navigate some difficult steps throughout my young career with aplomb, and for that I thank her. She has been a sterling example of the incredibly hard-working and high-achieving scientist who I aspire to become. Second, Rima Lucardi has been my main

mentor for the majority of my research focus. Her expansive knowledge of ecology has broadened the way I seek and attack research questions. She has sacrificed unfathomable amounts of her time and energy to ensure I was on the right track, and when the heaps of money and resources she sank into my work was not quite enough, she generously bankrolled and assisted me on a multi-day field trip to collect precious data. I cannot thank you two enough for all you have done.

I'd also like to thank the incredible past and current members of the Gandhi forest entomology lab, the numerous friends I have made in Warnell, and all the wonderful folks at the USDA Forest Service building. Finally, thank you to the USDA Forest Service, the Daniel B. Warnell School of Forestry and Natural Resources, and the UGA Graduate School for the generous funding to make my research possible.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER ONE

INTRODUCTION

The versatile and important canopy tree species, eastern white pine (*Pinus strobus* L.), faces a new threat. "Eastern white pine dieback" was first noticed in the 1990s and is now considered a novel forest health phenomenon (Costanza et al. 2018). Across its range, *P. strobus* trees are experiencing similar symptoms, including branch flagging, branch dieback, resinosis (excessive resin outflow), crown thinning, and the presence of various sizes and types of cankers (dead sections of bark, often caused by pathogens). All size classes are affected, but saplings appear to be the most vulnerable to mortality, which presents a considerable concern for the regeneration of *P. strobus* (Asaro et al. 2018, Schulz et al. 2018a). The limited research thus far has failed to find any abiotic factors contributing to the phenomenon (Schulz et al. 2018a), however, certain biotic agents appear closely associated with symptom severity. A native pathogen, *Caliciopsis pinea* Peck (Lombard 2003, Munck et al. 2015, Schulz et al. 2018a), has been associated with a canker disease, "Caliciopsis canker," which appears to be driving most of the symptoms.

In Virginia during 2006, forest health experts discovered minute insect cysts embedded within branch crotches, under lichen, and along the edges of Caliciopsis cankers. These were later identified as nymphs of the species *Matsucoccus macrocicatrices* Richards (eastern white pine bast scale), which are specialist herbivores of *P. strobus*. Further investigation throughout the Southern Appalachians, particularly within the southernmost extent of the *P. strobus* range in Georgia, revealed the insects

were widely distributed, and their incidence was correlated with dieback symptoms (Mech et al. 2013). However, *M. macrocicatrices* has historically been described only as native to eastern Canada and New England, USA, and has never before been documented south of Massachusetts (Mech et al. 2013). Furthermore, this insect has long been assumed to be one of many benign herbivore affiliates of *P. strobus* and has never before been reported to be associated with significant tree damage. Today, *M. macrocicatrices* occupies the range-wide distribution its host and, along with the pathogen *C. pinea*, is associated with signature dieback symptoms and mortality. It is clear something has changed in recent years. Evidence from the last two decades suggests *M. macrocicatrices* is a new pest insect on *P. strobus*, perhaps facilitating infection of pathogens, like *C. pinea*. With such an overwhelming dearth of knowledge on the species, including foundational aspects of the insect species' biology, ecology, and evolution, additional research is needed to develop a better understanding of the present and historical relationship of this emergent pest and its host.

1.1 Eastern white pine

1.1.1 Recent evolutionary history

Eastern white pine, *P. strobus* L. (Figure 1.1a, b), belongs to subsection *Strobus* (the white pines) within subgenus *Strobus* (the soft pines/five-needled pine). The nine North American "white pine" species include: *P. albicaulis* Engelm., *P. aristita* Engelm., *P. balfouriana* Balf., *P. flexilis* James, *P. lambertiana* Dougl., *P. longaeva* D.K. Bailey, *P. monticola* Douglas ex D. Don, *P. strobiformis* Engelm., and *P. strobus*. Of them, *P. strobus* is the only known species occurring in the eastern USA and Canada (Gernandt et

al. 2005, Price et al. 1998). The divergence between *P. strobus* and the five-needled white pines of western North America likely occurred in the middle Eocene (48-38 million years ago) due to the unstable tectonic, climatic and biogeographic events that characterized this part of the Tertiary (Eckert et al. 2006, Richardson 1998). There are two recognized varieties of eastern white pine: *P. strobus* var. *strobus* and *P. strobus* var. *chiapensis* (also known as *P. chiapensis*). *Pinus chiapensis* is located in the humid mountain regions of southern Mexico and likely became a disjunct lineage during the late Pleistocene (Farjon 2010). The *strobus* variety, referred to as *P. strobus* herein, has a current range extending longitudinally from Minnesota to Maine and latitudinally from Newfoundland, through southern Canada and New England, to northern Georgia along high elevation sites of the Appalachian Mountains (Figure 1.1c). There are also non-native populations in central Europe, which are considered invasive (Hadincová et al. 2008, Mandák et al. 2013).

As ice receded after the last glacial maximum, refugial populations of *P. strobus* on the mid-Atlantic coast and northwest Georgia provided sources of propagules for rapid northward and westward recolonization (Davis 1983, Nadeau et al. 2015). Fossilized pollen evidence suggests *P. strobus* arrived in Virginia's Shenandoah Valley around 13,000 years ago, the northeastern United States approximately 10,000 years ago, southeastern Canada 8,000-10,000 years ago, the eastern Great Lakes region 8,000-9,000 years ago, the western Great Lakes region around 7,000 years ago, and its northern range-limit in northeastern Canada around 5,000 years ago (Craig 1969, Davis 1983, Jacobson 1992, Jacobson et al. 1987, MacDonald et al. 1998, Zinck and Rajora 2016). The climate then cooled, resulting in a southward range contraction (Davis 1983). During postglacial

recolonization north and west, *P. strobus* largely followed jack pine (*P. banksiana* Lamb.) and red pine (*P. resinosa* Aiton). In many regions, jack pine dominated until *P. strobus* arrived and established (MacDonald et al. 1998). Deciduous hardwoods then arrived and outcompeted *P. strobus* on mesic sites (Davis 1983). The interplay between climatic changes and species range shifts led to *P. strobus* reaching its highest density approximately 4,000 years ago in northeastern North America and then around 1,000 years ago in the western Great Lakes (Jacobson 1992).

1.1.2 Post-settlement history

Pinus strobus was so important as a building resource during European colonization and settlement, it was given the folkloric distinction as "the tree that built America" (Abrams 2001). This species is even referred to as the "tree that won the American Revolution" due to American colonists successfully thwarting the British Navy from accessing its lumber to replace their masts and planks in the 1700s (Harlow et al. 1979, Abrams 2001). At heights of up to 55 m and diameters up to 130 cm, individual *P. strobus* were some of the tallest and largest trees that settlers encountered in the forested American wilderness. Its straight grain and light weight solidified *P. strobus* as the ideal timber species, spawning the white pine logging industry (Abrams 2001). Focusing primarily on old-growth populations, *P. strobus* logging began in an organized effort around the year 1700 in northeastern North America, but it moved west into the Great Lakes region starting in the 1830s once over-harvesting began to limit the number of available large trees (Jacobson 1979, Abrams 2001). The extensive harvesting of the 17th

and 18th centuries has left only 0.2% of the three billion m^3 pre-Columbian stands of *P. strobus* remaining today (Mehes et al. 2009).

The clearing of forests for agriculture was also prevalent during the rise of the logging industry, especially in New England (Foster 1992, Hooker and Compton 2003). When industrialization during the 19th century shifted focus away from farming, many of these cleared agricultural lands became abandoned (Foster 1992, Hooker and Compton 2003, Barton et al. 2012). Due to its ability to colonize marginal soils (Dovciak et al. 2005), *P. strobus* often regenerated naturally on these abandoned farmlands and open pasture lands (Foster 1992). In the southern USA, *P. strobus* was a preferred species to reforest clear-cuts and abandoned agricultural lands. For example, the Civilian Conservation Corps planted more than 20 million *P. strobus* seedlings in North Carolina during the 1950s alone (Vimmerstedt 1962). The practice of fire exclusion in the 1900s resulted in denser, closed forest canopies, which limited regeneration of competitors, such as oaks (*Quercus* spp.) and southern pines (subg. *Pinus Pinus* spp.). Following European settlement, *P. strobus* populations were devastated in many areas, but due to fire exclusion and agricultural land abandonment in the 20th century, regeneration is increasing in some regions (Costanza et al. 2018).

The commercial use of *P. strobus* today is reduced but still important. There are more than 600 million m^3 of standing *P. strobus* timber ($\geq$12.7 cm diameter at breast height (DBH)) in the USA alone, and standing value of white pine saw logs is estimated to be at least $18.6 billion (Livingston 2016). This species is also widely grown for Christmas trees, for ornamental use in urban settings, for historical and cultural purposes, and for ecological benefits (Schroeder 1992, Ostry et al. 2010).

1.1.3 Biology and ecological importance

The vast distribution of *P. strobus* is a testament to its versatility to grow from sea level to 1220 m in elevation, on a wide variety of soils, and in 28 different Society of America Foresters (SAF) forest cover types (Wendel and Smith 1990). It grows especially well on low to moderate quality and well-drained sandy soils where hardwoods struggle (Wendel and Smith 1990). Densities of *P. strobus* are somewhat binomially distributed, tending to peak in riparian valleys and in dry, nutrient-poor uplands (Abrams 2001). Hardwoods typically outcompete *P. strobus* on mesic, rich sites, and eastern hemlock, *Tsuga canadensis* (L.) Carrière, usually dominates in cool, moist sites (Stiell et al. 1994, Abrams 2001).

Saplings and young *P. strobus* trees have thin, smooth bark that thickens and develops deep fissures as they mature. The root system consists of 3-5 large, wide-branching roots for support (Wendel and Smith 1990). Root grafting is common and thriving *P. strobus* individuals can exchange nutrients via interconnected root systems to suppressed individuals (Bormann 1966). When 5-10 years old, trees begin to bear cones and maintain reproductive rigor for more than 200 years (Wilson and McQuilkin 1963, Wendel and Smith 1990). Pulses of seed production typically occur every 3-5 years. The growth rate of *P. strobus* is high when compared to other pines and hardwoods, but early growth is slow due to its relatively low shade tolerance (Wendel and Smith 1990). Seedlings require at least 20% sunlight to survive, but once released from shading they grow rapidly (Wendel and Smith 1990). From ages 20-120, the growth rate of *P. strobus* often exceeds that of its competitor species (Barrett 1933). Additionally, *P. strobus* is the

tallest and one of the longest-lived trees in eastern North America (Wendel and Smith 1990). Behemoth individuals over 125 cm in DBH, over 50 m in height, and over 450 years in age have been reported from the northern part of its range (Abrams 2001). Even at more modest sizes and ages, *P. strobus* can easily remain a fixture in the canopy and supercanopy for hundreds of years across its distribution (Abrams 2001).

Dendrochronological records show that *P. strobus* was seldom the first-rank species at the regional level in pre-settlement forests, which is similarly the case today (Abrams 2001). However, it does represent a major component of five SAF forest cover types: (a) eastern white pine, (b) red pine, (c) white pine-chestnut oak, (d) white pine-hemlock, and (e) white pine-northern red oak-red maple (Wendel and Smith 1990). This species is a particularly fundamental associate of eastern hemlock (*Tsuga canadensis*) and Carolina hemlock (*T. caroliniana* Engelm.) forests along the entirety of the Appalachian Mountains (Abrams 2001). The role of *P. strobus* as a canopy tree species has become even more important during the current widespread decline of hemlock trees due to the non-native insect, the hemlock woolly adelgid (*Adelges tsugae*) (Lovett et al. 2006). Changes in forest structure and function are inevitable with the loss of one or both conifer tree species, as these areas may become dominated by mid-story (e.g., *Rhododendron* spp.) and upper-story hardwood (e.g., black birch, *Betula lenta*) trees (Small et al. 2005).

Pinus strobus responds well to disturbance, excelling as a pioneer species. With its relatively low shade tolerance but high growth rate compared to other pines and hardwoods, *P. strobus* easily establishes in early-mid successional communities, quickly colonizing abandoned fields and filling canopy gaps created by fire, windfall, and

insect/disease outbreaks (Wendel & Smith 1990, Abrams 2001, Black and Abrams 2005). Despite the lack of serotinous cones or the ability to sprout vegetatively, *P. strobus* grows most favorably under a disturbance regime consisting of a 150 to 300-year fire cycle with intermittent surface-level fires occurring every 20 to 40 years (Frelich 1992). Once established and mature, its bark is thick enough to withstand surface-level fires, and its height and longevity allow it to dominate the canopy for centuries (Frelich 1992, Abrams 2001). Given its valuable ecopysiological traits, such as high growth rate, longevity, and ability to establish on marginal soils, *P. strobus* is commonly planted to control erosion, reclaim surface mined sites, and increase productivity on previously cultivated slopes (Hepp et al. 2015).

Countless wildlife species rely on *P. strobus* in various ways. Pure white pine stands provide storm shelter for wildlife when young and dense, and they foster a well-developed herbaceous layer when mature, increasing habitat richness for wildlife (Carey 1993). As a scattered supercanopy tree in mixed stands, *P. strobus* provides a vertical heterogeneity in the canopy that other pines fail to replicate (Rogers and Lindquist 1992). This multilayered quality yields more foraging and nesting opportunities for birds and mammals than in communities where *P. strobus* is absent (Rogers and Lindquist 1992). For instance, black bears (*Ursus americanus* Pallus) prefer these trees as refuge and bedding sites when raising their cubs (Elowe and Dodge 1989), and both bald eagles [*Haliaeetus leucocephalus* (Linnaeus)] and osprey [*Pandion haliaetus* (Linnaeus)] prefer these trees to more commonly occurring species for nest building (Kingsley and Ramquist 1993). The seeds, bark, and foliage of *P. strobus* provide important resources for birds and wildlife too. The seeds represent a major component of seed caches

belonging to white-footed mice [*Peromyscus leucopus* (Rafinesque)] and red-backed voles [*Clethrionomys gapperi* (Vigors)] (Abbott and Quink 1970), its inner bark is a favorite winter food of porcupines [*Erethizon dorsatum* (Linnaeus)] (Hazard 1982), and its foliage supports insect food for pine warblers [*Setophaga pinus* (Linnaeus)] (Green 1992). More examples of bird and wildlife species that utilize *P. strobus* are listed in Rogers and Lindquist (1992) and Green (1992).

1.1.4 Historical biotic threats

Long-lived organisms, such as trees, are usually highly resilient and must endure a variety of disturbances over their lifespans. A particularly common disturbance trees experience is attack from arthropods and pathogens. An estimated 277 insects and 110 disease agents have been reported to consume or infect *P. strobus*, but the majority of these associates are considered to be relatively innocuous (Wendel and Smith 1990). Only a handful of insects and pathogens that use *P. strobus* as hosts cause significant injury or mortality. For instance, the epidemic of white pine blister rust (caused by the non-native fungal pathogen, *Cronartium ribicola* J.C. Fisch) lasted decades, almost eliminating the seed source for regeneration at the northernmost range limit of *P. strobus* (Kinloch 2003). This non-native pathogen, introduced in 1906, has two obligate hosts: five-needle pines and currants (*Ribes* spp.). Because the pathogen only transmits from pine to currant and vice-versa, the predominant management response in eastern North America became to eradicate all currants. For example, in Maine starting in 1917, it became illegal to plant or grow currants, and as a result, there was greater than 50% reduction in blister rust incidence after 70 years (Ostrofsky et al. 1988). Similar programs

achieved similar results in other parts of the range. Blister rust is still present in eastern North America but is currently causing much more damage and mortality in the five-needle pines of western North America (Brar et al. 2015, Munck et al. 2015b). White pine weevil (*Pissodes strobi* Peck) has historically been considered the other major biotic threat to *P. strobus*, especially in commercial stands. This native species attacks and kills the terminal leader of mature trees, causing multiple, crooked, and/or bushy stems (Wilson 1978, Wendel and Smith 1990, Hamid et al. 1995). Mortality from weevil damage is rare and with appropriate silvicultural management, such as shelterwood cuts, damage to trees can be greatly reduced (Ostry et al. 2010).

1.1.5 Eastern white pine dieback

Forest health specialists began observing novel symptomology in *P. strobus* starting in the 1990s within the northeastern USA, which could not be attributed to the usual suspects (e.g., white pine weevil and white pine blister rust). Reductions in timber quality resulted in downgrades and lost revenue for timber companies. Subsequent reports revealed that the symptoms were present range-wide, but with slight regional variation. In the Northeast, branch and stem cankers are prevalent on trees >10 cm DBH, causing severe resinosis (Figure 1.2a), crown thinning, and large bark cracks/fissures (Vermont Department of Forests, Parks, and Recreation 2001, Lombard 2003, Maine Forest Service 2008). In the Southern Appalachians, symptoms were first reported between 2006 and 2007 in Virginia and West Virginia, and in 2010 in Georgia (Asaro 2011, Rose 2011, Mech et al. 2013, Schulz 2015). Multiple branch and/or stem cankers are also commonly observed, and this necrotic tissue, enlarging over time, can eventually

girdle the branches or the main stem on saplings and mature trees (Mech et al. 2013, Schulz et al. 2018b). Flagged needles, which appear as brown, dried fascicles, appear first in the lower canopy (Figure 1.2b), followed by branches dying back in a bottom-up pattern (Figure 1.2c). This dieback progression results in highly reduced crowns (Figure 1.2d) and can eventually cause mortality in mature trees (Figure 1.2e). Stem girdling, and hence mortality, is most common in young seedlings and saplings (Figure 1.2f, g) (Asaro et al. 2018). In the Great Lakes region of the USA, seedling mortality and canker formation have been observed since 2006 (Chhin 2013, Michigan Department of Natural Resources 2015). Needle discoloration and canker occurrence were documented as early as 2013 on mature trees in southern Canada as well (Llewellyn 2013). The overlap in symptomology and the similarity of reports indicates this is a novel, range-wide phenomenon. Two understudied organisms, previously assumed to be rather innocuous, have emerged as key associates to the symptoms: a native, fungal pathogen, *Caliciopsis pinea* Peck (Coryneliales: Coryneliaceae), and a scale insect, the eastern white pine bast scale, *Matsucoccus macrocicatrices* Richards (Hemiptera: Matsucoccidae).

1.1.6 Caliciopsis canker

"Caliciopsis canker" is the disease most attributed to eastern white pine dieback symptoms and is caused by the ascomycetous fungal pathogen, *Caliciopsis pinea*. Of all the fungi isolated from cankers of symptomatic *P. strobus* in a recent study, *C. pinea* was by far the most virulent and, once inoculated on seedlings, was the only species to form girdling cankers (Schulz et al. 2018b). Unlike white pine blister rust, this fungal species is (assumed to be) native to North America. First reported near Ithaca, New York, in

1880, it has long been known to infect *P. strobus* (Fitzpatrick 1920, Peck 1880, Ray 1936). *Caliciopsis pinea* creates cankers on stems and branches, characterized by reddish-brown depressions with globose, asexual fruiting bodies and stalked, spore-carrying fruiting bodies (Figure 1.3a, b) (Ray 1936, Funk 1963). Once mature, the black ascospores (resembling clusters of eyelashes) are disseminated by wind, rain splash, or stemflow (Funk 1963, Delatour 1969). Previous studies report damage and cite occasional mortality due to Caliciopsis canker, but they were infrequent (Fitzpatrick 1920, Overholts 1930, Ray 1936, Fitzpatrick 1942). Since the 1990s, however, reports of Caliciopsis canker incidence, *P. strobus* damage, and associated mortality have increased substantially (Lombard 2003, Asaro 2011, Rose 2011, Chhin 2013, Mech et al. 2013, Munck et al. 2015a).

Caliciopsis cankers vary in appearance, but there are three general types: (1) small, sunken, reddish-brownish depressions in the bark (Figure 1.4a), (2) areas of roughened bark adjacent to branch nodes (Figure 1.4b), and (3) bark cracks and fissures colonized by ascospores (Figure 1.4c) (Overholts 1930, Ray 1936, Lombard 2003, Munck et al. 2015a). All are associated with internal damage to the host tree, such as necrosis of the xylem tissue (Figure 1.4d). Surface cankers, if not calloused over by the tree, can expand over time - multiple cankers can even coalesce. If this necrotic growth fully envelopes the circumference of a branch or stem, girdling of vascular flow can occur, effectively killing the branch or tree (Figure 1.3b). To first establish and subsequently form cankers, fungal hyphae must first aggregate under the bark and grow into the vascular cambium. Hence, *C. pinea* spores first need an entry point ("infection court") such as a mechanical wound, natural bark fissure, or an insect feeding site to

successfully colonize its host (Fitzpatrick 1942, Funk 1963, Lombard 2003, Munck et al. 2015a). There is a growing body of correlative evidence to suggest that the eastern white pine bast scale (*Matsucococcus macrocicatrices*), which has also emerged in association with *P. strobus* dieback symptoms, may be facilitating the initial infection of *C. pinea* and development of the Caliciopsis canker disease.

1.2 The eastern white pine bast scale (*Matsucoccus macrocicatrices*)

There is an overall paucity of research concerning *M. macrocicatrices*. Before Mech et al. (2013) first documented its presence in the Southern Appalachians, only two refereed studies had addressed *M. macrocicatrices* as the focal organism and were published over 50 years prior (Richards 1960, Watson et al. 1960). The following sections cite all current knowledge about this species' taxonomy, biology, and ecology to date, which have been gleaned from papers on congener species, the two *M. macrocicatrices* papers published in 1960, papers from a 2018 special issue in *Forest Ecology and Management* volume 423*,* and my own research presented, herein.

1.2.1 Taxonomy

All members of family Matsucoccidae, formerly part of the family Margarodidae *s.l.*, inhabit the Holarctic and feed exclusively on *Pinus* species. A main distinguishing characteristic of the family is the presence of large simple pores (or cicatrices) arranged dorsally in rows along the abdomen of the female (Foldi 2004). There are two genera: one extinct, *Eomatsucoccus* Koteja (4 fossil species), and one extant, *Matsucoccus* Cockrell (33 extant species and 6 fossil species) (Foldi 2004). Seventeen species of

Matsucoccus are found in North America, five of them located east of the Mississippi River (Gill 1993). There is an incredible homogeny of characters among *Matsucoccus* species, making them difficult to distinguish, especially in the nymphal stages, but certain characters in the adult female are considered reliable (Foldi 2004, Ray 1982). Universal COI primers have proven problematic for many species of the Coccoidea (Deng et al. 2012, Wang et al. 2015), so members of *Matsucoccus* have been routinely distinguished using 18S and 28S nuclear rDNA genes as barcodes (Booth and Gullan 2006; e.g. Mech et al. 2013).

Matsucoccus spp. are referred to as the "pine bast scales," presumably because their stylets penetrate and ingest sap from the "bast" region of the tree (an uncommon term for the outermost layer of the phloem, containing large cells that move and store sugars). Although most *Matsucoccus* species cause negligible feeding damage to their hosts, some are notorious pests. Notable examples include *M. matsumurae* (Kuwana) killing red pine (*P. resinosa* Sol. ex Aiton) in New England states (Booth and Gullan 2006), *M. feytaudi* Ducasse killing maritime pine (*P. pinaster* Aiton) in Mediterranean countries (Kerdelhué et al. 2014), and *M. josephi* Bodenheimer and Harpaz killing Aleppo pine (*P. halepensis* Miller) in Israel (Mendel 1998).

The eastern white pine bast scale, *Matsucoccus macrocicatrices* Richards (Figure 1.5a, b), is the only matsuccocid found on *P. strobus*. The only other two North American matsucoccids known to colonize white pines (i.e. *Pinus* subg. *Strobus* subsect. *Strobus*) are *M. acalyptus* Herbert (found on *P. flexilis* E. James and *P. lambertiana* Douglas) and *M. paucicicatrices* Morrison (found on *P. flexilis*, *P. lambertiana, P. albicaulis* Engelm. and *P. monticola* Douglas ex D. Don) (Ray 1982). According to Richards (1960), the

distinguishing morphological character between *M. macrocicatrices* and *M. paucicicatrices* is merely the diameter of the dorsal cicatrices in adult females. *Matsucoccus macrocicatrices* lacks an official common name but has been referred to in the literature as the Canadian pine scale (Kosztarab 1996), the cicatrix pine scale (Ray 1982), and the white pine fungus scale (Encyclopedia of Life, eol.org). With consultation from other *Matsucoccus* researchers (Ray Gill pers. comm., Penny Gullan pers. comm., Zvi Mendel pers. comm.), we advocate the common name to be the "eastern white pine bast scale." This is the most descriptive name and differentiates this species from its congeners that exist on western white pine species.

1.2.2 Life cycle

All described species within *Matsucoccus* are extremely host specific and none are known to feed on more than one subgenus of *Pinus* (Ray 1982). *Matsucoccus macrocicatrices* is a specialist of *P. strobus*, but its phenology is not fully understood. Before 2006, the only studies on *M. macrocicatrices* were conducted in Canada during the 1950s, where they report a two-year life cycle (Richards 1960, Watson et al. 1960). This appears to be true of populations at high (northern) latitudes, but we have observed that development completes in a single year in the warmer southeastern USA states (Mech et al. 2013, Thomas Whitney pers. obs.). The latitudinal threshold where the lifespan transitions from two years to one is unknown. Relative to several congeners with multi-voltine life cycles (Foldi 2004), this species has a lengthy generation time. A detailed account of the *M. macrocicatrices* life cycle, observed under laboratory conditions, can be found in Appendix 1. All *M. macrocicatrices* spend the majority of

their life cycle as 2nd instar cysts, the stage most noticeable to observers (Figure 1.5a, b; Figure 1.6a). They resemble black, oblong pearls often found embedded in tight spaces on *P. strobus* bark, such as in branch crotches, under lichen, or along the edge of cankers (Figure 1.5c-f). This is a sessile feeding stage; cysts extract sap starting in summer from the tree through their long stylets (Figure 1.6b), which are perpetually inserted into the vascular tissue ("bast"). This is also the overwintering stage, unlike in *M. gallicolis* (egg), or *M. pini* and *M. matsumurae* (1st instar crawler) (Watson et al. 1960). The cold tolerance of *M. macrocicatrices* is quite high during the winter, even at the warmer, low latitudes of its range (see Appendix 2). Once the cysts have fully grown to ~1.5 mm in spring, the adults emerge. Pre-adult males and adult females are indistinguishable to the naked eye (Figure 1.6c, f), but males tend to emerge two weeks earlier to undergo further development (Appendix 1). Although *M. macrocicatrices* are hemimetabolous (incomplete metamorphosis), males must go through a developmental pathway that is pseudo-holometabolous (complete metamorphosis) (Gullan and Kosztarab 1997). This is because once pre-adult males (3rd instar) emerge from their cyst stage, they envelop themselves in a "cocoon" of silk. In this "pupal" stage (Figure 1.5d), males will completely transform morphologically, eventually developing into a final winged adult resembling a gnat (Figure 1.5e). They have a pair mesothoracic wings with little veination, a pair of metathoracic wings reduced to haltere-like structures, a pair of long, hairy antennae, and highly reduced, probably non-functional, mouthparts (Gullan and Kosztarab 1997). The timing between final adult male development and the emergence of the adult females from their cyst stage is synchronized (Appendix 1). Adult females are neotenous, meaning they are nymph-like and retain juvenile traits (Gullan and Kosztarab

1997). Their mobility is highly limited, and they are thought to never disperse. Instead, immediately upon emerging, females emit pheromones from their large dorsal cicatrices to attract males. After copulation (Figure 1.5g), they lay hundreds of eggs in bark crevices within a silken mass (Figure 1.5h). Mobile 1st instar "crawlers" (Figure 1.5i) hatch from eggs in late spring, disperse by wind (most likely), and once landing successfully on another *P. strobus* host, they choose new feeding sites to settle. Afterward, individuals undergo a molt, thus completing the cycle into legless, 2nd instar cysts once again.

1.2.3 Dispersal

In general, *M. macrocicatrices* may be capable of long-distance migration and colonization, but different life stages vary greatly in this ability. Adult females probably do not disperse. They are wingless, and based on what we know about congeneric species, they likely lack the ability to stray from the original tree they first settled as a 1st instar crawler (Kerdelhue et al. 2014). Like pollen grains, the winged adult males can contribute to gene flow, but not colonization. However, they are very weak fliers and probably cannot travel far (Thomas Whitney, pers. obs.). Female *Matsucoccus* species emit a pheromone attracting the males, but the maximum distance these chemicals can reach potential mates is unclear. The main dispersal stage is likely the 1st instar crawler. Crawlers can contribute to gene flow via passive, windborne dispersal to existing populations, and they can also colonize new, vacant habitats, similar to plant seeds. Soft scales (Hemiptera: Coccidae), with similar life histories to matsucoccids, assume a "dispersal posture," with their anterior end and antennae outstretched, to be deliberately

and more effectively carried off by wind currents (Yardeni and Rosen 1990). In *M. matsumurae*, 1st instar crawlers have been observed to stay airborne for almost half a kilometer (Bean and Godwin 1971); other scale insect species can be carried over several kilometers (Greathead 1972, 1990, Gullan and Kosztarab 1997). Alternatively, some scale insect crawlers can exhibit phoresy: the act of hitchhiking on a better-dispersing animal for transport. For instance, Magsig-Castillo et al. (2010) found that crawlers of an armored scale species (Hemiptera: Diaspididae) use specialized leg structures to attach themselves to flies, ladybeetles, and ants. However, there is no evidence yet of any *Matsucoccus* spp. exhibiting phoresy.

1.2.4 Ecological interactions

Endosymbiotic relationships between scale insects and microorganisms are remarkably prevalent, especially within the Margarodidae *s.l.* (Gullan and Kosztarab 1997). *Matsucoccus macrocicatrices* has not been investigated for any endosymbioses or screened for microorganism infection, except for the bacteria *Wolbachia* (see Appendix 3). This is a very common group of bacterial endosymbionts in arthropods, which alter reproduction of their hosts in various ways (Werren et al. 2008). Almost 30% of *M. macrocicatrices* cysts out of 45 samples tested positive for *Wolbachia* (Appendix 3), suggesting infection is common and widespread in this species. The particular role it plays in the lives of *M. macrocicatrices*, if any, is unknown.

No natural enemies (predators, parasitoids, or parasites) of *M. macrocicatrices* are known, likely due to the dearth of research on the species. A member from the genus of *Matsucoccus*-specialist predators, *Elatophilus brimleyi* Kelton (Hemiptera:

Anthocoridae), has been collected on Canadian *P. strobus* trees in the past, however there is no direct evidence of a predator-prey relationship with *M. macrocicatrices* (Mendel et al. 1991). No museum holdings for *Elatophilus* spp. exist in the University of Georgia collections. Other *Matsucoccus* spp. are common prey for generalist predators, such as lady beetles (Coleoptera: Coccinellidae) and lacewing larvae (Neuroptera: Crysopidae) (e.g. Bean and Godwin 1971). A diverse generalist predator community, including coleopterans, dipterans, and neuropterans, has been surveyed on *P. strobus* in the Southern Appalachians (Wantuch et al. 2019), so it is conceivable that they may also prey upon *M. macrocicatrices*.

Quite unique to the matsucoccids, *M. macrocicatrices* has been described to have a symbiotic relationship with the epiphytic fungus, *Septobasidium pinicola* Snell (Figure 1.7a-d). Couch (1938) describes *S. pinicola* as a common fungus that forms small mats superficially on the bark of *P. strobus* but requires parasitizing a scale insect-host from which to derive nourishment. Specifically, Couch (1938) cites a *Chermes sp.* as the only insect host (an archaic genus with various species absorbed by both the Adelgidae and Physillidae). This proposed fungal symbiont compromises the bodies of a few individuals in exchange for the protection of many individuals, perhaps from adverse weather, predation, or parasitism. Couch (1938) writes:

> "In trying to find an appropriate term to apply to the relationship…, it is clear that the term parasitism is inadequate… Parasitism implies the living together of two organisms to the benefit of one and the detriment or even death of the other… The fungus and insects live together interdependently; the fungus furnishes a home and protection for the insects, while, in return, the insects furnish food and a means of distribution for the fungus. Some of the insects are sacrificed, but to the advantage of the insect colony as a whole; for, from the juices taken from the parasitized insects, the fungus grows and forms more houses for the non-parasitized as well as

> the parasitized insects… The relationship here is, therefore, obviously one of symbiosis."

Watson et al. (1960) was the first to identify the relationship between *S. pinicola* and *M. macrocicatrices*. The authors observed insect crawlers populating the edges of the *S. pinicola* epiphytic mats, indicating they were colonizing there with intent, and found haustoria-like fungal structures within body cavities of 2nd instar cysts. The symbiosis is probably not obligatory, but the authors hypothesized it is a beneficial relationship: fungal nourishment for insect protection (Watson et al. 1960). More recently, while surveying sites in Georgia, Mech et al. (2013) found that of the 120 symptomatic *P. strobus* branches examined (91.7% of which were infested with *M. macrocicatrices*), only 2.5% had *S. pinicola* fungal mats present. My personal observations confirm this – only on a rare occasion will *S. pinicola* be present on *P. strobus* where *M. macrocicatrices* is also present. Therefore, the scale insect can clearly survive without the fungal association, and crawlers routinely find other means of protection. Cysts are more often found settled under lichen, deep in branch nodes, and embedded within cankers of fungal pathogens (Figure 1.5c-f) (Mech et al. 2013, Thomas Whitney pers. obs.).

1.2.5 Association with C. pinea and eastern white pine dieback

All indications from the limited research available on this insect suggest its historical distribution is limited to the Canadian Maritime Provinces and the New England region of the USA. Richards (1960) used type specimens from Ontario and New Brunswick to describe the species. Watson et al. (1960) studied the relationship of *M. macrocicatrices* and *S. pinicola* in Ontario, New Brunswick, and Nova Scotia, Canada. Mech et al. (2013) conducted a search for accessed specimens in 27 museum collections

across eastern North America, and found individuals only from Ontario, New Brunswick, and Vermont. According to Drooz (1985) and Kosztarab (1996), New Hampshire and Massachusetts also have recorded the insect. There is currently no good explanation for its presence outside of its historically described distribution. *Matsucoccus macrocicatrices* has now been recorded in 14 states within the USA, ten of which occur outside its purported native range (Figure 1.8).

Even within its historically described distribution in New England, USA, and southeastern Canada, *M. macrocicatrices* has generally been considered a benign herbivore of *P. strobus*, never before reported as a cause of significant damage or of raising any forest health concerns prior to 2006. However, recent work suggests this insect is a main associate, and possibly a contributing cause, of current dieback symptoms and mortality among North American *P. strobus*. Within the Southern Appalachians, Mech et al. (2013) found that a strong association between *P. strobus* showing hallmark symptoms and the presence of *M. macrocicatrices*. Schulz et al. (2018b) also found that the frequency and severity of seedling dieback across the Northern and Southern Appalachian Mountains was significantly correlated with *M. macrocicatrices* incidence. *Caliciopsis pinea* and *M. macrocicatrices* incidence, together, was also highly correlated with dieback symptoms (Schulz et al. 2018a, b). Although correlative, there is a compelling body of evidence forming to suggest that these organisms likely form a facultative insect-pathogen complex contributing to the recent phenomenon of eastern white pine dieback and mortality.

1.3 Research Objectives

When a new potential pest emerges, it is important to gather basic, foundational knowledge of its biology, ecology, and evolution. This is necessary to develop appropriate management solutions. With the eastern white pine bast scale, *Matsucoccus macrocicatrices*, there are many unknowns. Paramount to all questions is: why has this insect – previously believed to be a negligible herbivore and limited to a narrow region of North America – now become a destructive and range-wide pest of its host, eastern white pine, *P. strobus*? The following dissertation research comprises of studies aimed at unpacking this overarching question. The principal objectives were as follows:

1. Examine the distribution of *M. macrocicatrices* cysts and Caliciopsis cankers within individual trees and among size classes, to determine if colonization patterns are consistent with dieback patterns.
2. Explore the most likely historical biogeography of *M. macrocicatrices* to determine if it is non-native to regions outside of its putative historical range in northeastern North America.
3. Assess the genetic diversity, variation, and population structure of the host tree, *P. strobus*, within the understudied Southern Appalachian region, to determine its relative adaptive potential for future conservation implications.

1.4 References

Abbott, H.G., Quink, T.F., 1970. Ecology of eastern white pine seed caches made by small forest mammals. Ecology 51, 271-278.

Abrams, M.D., 2001. Eastern white pine versatility in the presettlement forest. Bioscience 51, 967-979.

Asaro, C., Chamberlin, L.A., Rose, J.A., Mooneyham, K., Rose, A.K., 2018. Mortality of eastern white pine (*Pinus strobus* L.) in association with a novel scale insect-pathogen complex in Virginia and West Virginia. Forest Ecology and Management 423, 37-48.

Barrett, L.I., 1933. Growth rate of northern white pine in the Southern Appalachians. Journal of Forestry 31, 570-572.

Bean, J.L., Godwin, P.A., 1971. Red Pine Scale. Forest Pest Leaflet 10. USDA Forest Service, p. 6.

Black, B.A., Abrams, M.D., 2005. Disturbance history and climate response in an old-growth hemlock-white pine forest, central Pennsylvania. Journal of the Torrey Botanical Society 132, 103-114.

Booth, J.M., Gullan, P.J., 2006. Synonymy of three pestiferous *Matsucoccus* scale insects (Hemiptera: Coccoidea: Matsucoccidae) based on morphological and molecular evidence. Proceedings of the Entomological Society of Washington 108, 749-760.

Bormann, F.H., 1966. Structure function and ecological significance of root grafts in *Pinus strobus* L. Ecological Monographs 36, 1-26.

Brar, S., Tsui, C.K.M., Dhillon, B., Bergeron, M.J., Joly, D.L., Zambino, P.J., El-Kassaby, Y.A., Hamelin, R.C., 2015. Colonization history, host distribution, anthropogenic influence and landscape features shape populations of white pine blister rust, an invasive alien tree pathogen. Plos One 10.

Carey, J.H., 1993. *Pinus strobus*. In, Fire Effects Information System (FEIS). USDA Forest Service, Rocky Mountain Research Station, Fire Sciences Laboratory.

Chhin, S., 2013. White pine decline in the Great Lakes region. In. USDA-FS Evaluation Monitoring Program, p. 4.

Costanza, K.K.L., Whitney, T.D., McIntire, C.D., Livingston, W.H., Gandhi, K.J.K., 2018. A synthesis of emerging health issues of eastern white pine (*Pinus strobus*) in eastern North America. Forest Ecology and Management 423, 3-17.

Couch, J.N., 1938. The genus *Septobasidium*. University of North Carolina Press, Chapel Hill, North Carolina, USA.

Craig, A.J., 1969. Vegetational history of the Shenandoah Valley, Virginia. Geological Society of America Special Papers, 283-296.

Davis, M.B., 1983. Quaternary history of deciduous forests of eastern North America and Europe. Annals of the Missouri Botanical Garden 70, 550-563.

Delatour, C., 1969. Contribution à l'étude du *Caliciopsis pinea* Peck : résultats complémentaires d'inoculations artificielles. Annales des Science Forestiéres 26, 285-295.

Deng, J., Yu, F., Zhang, T.X., Hu, H.Y., Zhu, C.D., Wu, S.A., Zhang, Y.Z., 2012. DNA barcoding of six *Ceroplastes* species (Hemiptera: Coccoidea: Coccidae) from China. Molecular Ecology Resources 12, 791-796.

Dovciak, M., Frelich, L.E., Reich, P.B., 2005. Pathways in old-field succession to white pine: Seed rain, shade, and climate effects. Ecological Monographs 75, 363-378.

Drooz, A.T., 1985. Insects of Eastern Forests. USDA Forest Service, Asheville, North Carolina, USA. p. 608.

Eckert, A.J., Hall, B.D., 2006. Phylogeny, historical biogeography, and patterns of diversification for *Pinus* (Pinaceae): Phylogenetic tests of fossil-based hypotheses. Molecular Phylogenetics and Evolution 40, 166-182.

Elowe, K.D., Dodge, W.E., 1989. Factors affecting black bear reproductive success and cub survival. Journal of Wildlife Management 53, 962-968.

Farjon, A., 2010. A Handbook of the World's Conifers. Koninklijke Brill NV, Leiden, The Netherlands.

Fitzpatrick, H.M., 1920. Monograph of the Coryneliaceae. Mycologia 12, 206-237.

Fitzpatrick, H.M., 1942. Revisionary studies in the Coryneliaceae. II. The genus *Caliciopsis*. Mycologia 34, 489-514.

Foldi, I., 2004. The Matsucoccidae in the Mediterranean basin with a world list of species (Hemiptera: Sternorrhyncha: Coccoidea). Annales De La Societe Entomologique De France 40, 145-168.

Foster, D.R., 1992. Land-use history (1730-1990) and vegetation dynamics in central New England, USA. Journal of Ecology 80, 753-772.

Frelich, L.E., 1992. The relationship of natural disturbances to white pine stand development. In: Stine, R.A. (Ed.), White Pine Symposium, Duluth, Minnesota, USA, pp. 27-37.

Funk, A., 1963. Studies in the genus *Caliciopsis*. Canadian Journal of Botany 41, 503-543.

Gernandt, D.S., Lopez, G.G., Garcia, S.O., Liston, A., 2005. Phylogeny and classification of *Pinus*. Taxon 54, 29-42.

Gill, R.J., 1993. The scale insects of California Part 2. Analysis and Identification Branch, Division of Plant Industry, California Department of Food and Agriculture.

Greathead, D.J., 1972. Dispersal of sugar-cane scale, *Aulacaspis tegalensis* (Zhut) (Hemiptera: Diaspididae) by air currents. Bulletin of Entomological Research 61, 547-558.

Greathead, D.J., 1990. Crawler behaviour and dispersal. In: Ben-Dov, Y., Hodgson, C. J. (Eds.), Soft Scale Insects: their Biology, Natural Enemies and Control., pp. 305-308.

Green, J.C., 1992. Ecological features of white pine stands for wildlife. In: Stine, R.A. (Ed.), White Pine Symposium, Duluth, Minnesota, USA, pp. 44-53.

Gullan, P.J., Kosztarab, M., 1997. Adaptations in scale insects. Annual Review of Entomology 42, 23-50.

Hadincova, V., Munzbergova, Z., Wild, J., Sajtar, L., Maresova, J., 2008. Dispersal of invasive *Pinus strobus* in sandstone areas of the Czech Republic. In: Tokarska-Guzik, B., Brock, J.H., Brundu, G., Child, L., Daehler, C.C., Pysek, P. (Eds.), Plant Invasions: Human Perception, Ecological Impacts and Mangagement. Backhuys Publishers, Leiden, The Netherlands, pp. 117-132.

Hamid, 1995. White Pine Weevil. USDA Forest Service, Forest Insect and Disease Leaflet. p. 10.

Harlow, W.M., Harrar, E.S., White, F.M., 1979. Textbook of Dendrology. McGraw-Hill, New York, USA.

Hazard, E.B., 1982. The Mammals of Minnesota. University of Minnesota Press, Minneapolis, Minnesota, USA.

Hepp, T.E., Vimmerstedt, J.P., Smalley, G.W., McNab, W.H., 2015. Estimating yields of unthinned eastern white pine plantations from current stocking in the Southern Appalachians. Forest Science 61, 114-122.

Hooker, T.D., Compton, J.E., 2003. Forest ecosystem carbon and nitrogen accumulation during the first century after agricultural abandonment. Ecological Applications 13, 299-313.

Jacobson Jr., G.L., 1992. A 7000-year history of white pine. In: Stine, R.A. (Ed.), White Pine Symposium, Duluth, Minnesota, USA, pp. 19-26.

Jacobson Jr., G.L., Webb III, T., Grimm, E.C., 1987. Patterns and rates of vegetation change during the deglaciation of eastern North America. In: Ruddman, W.F., Wright Jr., H.E. (Eds.), North America and Adjacent Oceans during the Last Deglaciation, the Geology of North America. Geological Society of America, Boulder, Colorado, USA, pp. 277-288.

Kerdelhue, C., Boivin, T., Burban, C., 2014. Contrasted invasion processes imprint the genetic structure of an invasive scale insect across southern Europe. Heredity 113, 390-400.

Kingsley, N.P., Ramquist, J., 1993. Timber resource of the Superior National Forest. Resource Bulletin NC 145.

Kinloch, B.B., 2003. White pine blister rust in North America: Past and prognosis. Phytopathology 93, 1044-1047.

Kosztarab, M., 1996. Scale insects of northeastern North America. Identification, biology, and distribution. Virginia Museum of Natural History Special Publication 3, 1-650.

Little, E.L., 1971. Atlas of United States Trees, Volume 1: Conifers and important hardwoods. USDA Forest Service, Washington, DC, USDA.

Livingston, W., 2016. Eastern White Pine Health and Response to Environmental Changes. In. USDA Multistate Research Project Full Proposal, p. 19.

Llewellyn, J., 2013. White Pine Decline. (Accessed October 6, 2017). https://www.isaontario.com/content/white-pine-decline.

Lombard, K., 2003. Caliciopsis canker (pine canker) *Caliciopsis pinea*. University of New Hampshire Cooperative Extension Publication, Durham, NH, USA.

Lovett, G.M., Canham, C.D., Arthur, M.A., Weathers, K.C., Fitzhugh, R.D., 2006. Forest ecosystem responses to exotic pests and pathogens in eastern North America. Bioscience 56, 395-405.

MacDonald, G.M., Cwynar, L.C., Whitlock, C., 1998. The late Quaternary dynamics of pines in northern North America. In: Richardson, D.M. (Ed.), Ecology and Biogeography of *Pinus*. Cambridge University Press, London, United Kingdom, pp. 122-136.

Magsig-Castillo, J., Morse, J.G., Walker, G.P., Bi, J.L., Rugman-Jones, P.F., Stouthamer, R., 2010. Phoretic Dispersal of Armored Scale Crawlers (Hemiptera: Diaspididae). Journal of Economic Entomology 103, 1172-1179.

Mandak, B., Hadincova, V., Mahelka, V., Wildova, R., 2013. European invasion of North American *Pinus strobus* at large and fine scales: High genetic diversity and fine-scale genetic clustering over time in the adventive range. Plos One 8.

Mech, A.M., Asaro, C., Cram, M.M., Coyle, D.R., Gullan, P.J., Cook, L.G., Gandhi, K.J.K., 2013. *Matsucoccus macrocicatrices* (Hemiptera: Matsucoccidae): First report, distribution, and association with symptomatic eastern white pine in the southeastern United States. Journal of Economic Entomology 106, 2391-2398.

Mehes, M., Nkongolo, K.K., Michael, P., 2009. Assessing genetic diversity and structure of fragmented populations of eastern white pine (*Pinus strobu*s) and western white pine (*P. monticola*) for conservation management. Journal of Plant Ecology 2, 143-151.

Mendel, Z., 1998. Biogeography of *Matsucoccus josephi* (Homoptera : Matsucoccidae) as related to host resistance in *Pinus brutia* and *Pinus halepensis*. Canadian Journal of Forest Research-Revue Canadienne De Recherche Forestiere 28, 323-330.

Mendel, Z., Carmi, E., Podoler, H., 1991. Relations between the genera *Matsucoccus* (Homoptera: Margarodidae) and *Elatophilus* (Hemiptera: Anthocoridae) and their significance. Annals of the Entomological Society of America 84, 502-507.

Michigan Department of Natural Resources, 2015. 2015 Forest Health Highlights. p. 50.

Munck, I.A., Livingston, W., Lombard, K., Luther, T., Ostrofsky, W.D., Weimer, J., Wyka, S., Broders, K., 2015a. Extent and severity of Caliciopsis canker in New England, USA: An emerging disease of eastern white pine (*Pinus strobus* L.). Forests 6, 4360-4373.

Munck, I.A., Tanguay, P., Wehner, J., Villani, S.M., Cox, K.D., 2015b. Impact of white pine blister rust on resistant cultivated *Ribes* and neighboring eastern white pine in New Hampshire. Plant Disease 99, 1374-1382.

Nadeau, S., Godbout, J., Lamothe, M., Gros-Louis, M.-C., Isabel, N., Ritland, K., 2015. Contrasting patterns of genetic diversity across the ranges of *Pinus monticola* and *P. strobus*: A comparison between eastern and western North American postglacial colonization histories. American Journal of Botany 102, 1342-1355.

Ostrofsky, W.D., Rumpf, T., Struble, D., Bradbury, R., 1988. Incidence of white pine blister rust in Maine after 70 years of a *Ribes* eradication program. Plant Disease 72, 967-970.

Ostry, M.E., Laflamme, G., Katovich, S.A., 2010. Silvicultural approaches for management of eastern white pine to minimize impacts of damaging agents. Forest Pathology 40, 332-346.

Overholts, L.O., 1930. Mycological notes for 1928-1929. Mycologia 22, 232-246.

Peck, C.H., 1880. Report of the Botanist. Annual Report N.Y. State Museum of Natural History, p. 32.

Price, R.A., Liston, A., Strauss, S.H., 1998. Phylogeny and systematics of *Pinus*. In: Richardson, D.M. (Ed.), Ecology and biogeography of *Pinus*, 49-68.

Ray, C.H., Jr., 1982. Revision of the genus *Matsucoccus* (Homoptera: Coccoidea: Margarodidae) in North America. Dissertation Abstracts International B Sciences and Engineering 42, 4675-4675.

Ray, W.W., 1936. Pathogenicity and cultural experiments with *Caliciopsis pinea*. Mycologia 28, 201-208.

Richards, W.R., 1960. A new species of *Matsucoccus* Cockerell (Homoptera: Coccoidea). Canadian Entomologist 92, 179-181.

Richardson, D.M., Rundel, P.W., 1998. Ecology and biogeography of *Pinus*: An introduction. In: Richardson, D.M. (Ed.), Ecology and biogeography of *Pinus*, 3-46.

Rogers, L.L., Lindquist, E.L., 1992. Supercanopy white pine and wildlife. In: Stine, R.A. (Ed.), White Pine Symposium, Duluth, Minnesota, USA, pp. 39-43.

Rose, J.A., 2011. A decline in state's white pines found. West Virginia Department of Agriculture Market Bulletin.

Schroeder, H.W., 1992. The tree of peace: Symbolic and spiritual values of the white pine. In: Stine, R.A. (Ed.), White Pine Symposium, Duluth, Minnesota, USA, pp. 73-83.

Schulz, A.N., Mech, A.M., Asaro, C., Coyle, D.R., Cram, M.M., Lucardi, R.D., Gandhi, K.J.K., 2018a. Assessment of abiotic and biotic factors associated with eastern white pine (*Pinus strobus* L.) dieback in the Southern Appalachian Mountains. Forest Ecology and Management 423, 59-69.

Schulz, A.N., Mech, A.M., Cram, M.M., Asaro, C., Coyle, D.R., Lucardi, R.D., Lucas, S., Gandhi, K.J.K., 2018b. Association of *Caliciopsis pinea* Peck and *Matsucoccus macrocicatrices* Richards with eastern white pine (*Pinus strobus* L.) seedling dieback. Forest Ecology and Management 423, 70-83.

Service, M.F., 2008. Forest and shade tree insect and disease conditions for Maine: A summary of the 2007 situation. In, Summary Report. Forest Health and Monitoring Division, Augusta, Maine, USA.

Small, M.J., Small, C.J., Dreyer, G.D., 2005. Changes in a hemlock-dominated forest following woolly adelgid infestation in southern New England. Journal of the Torrey Botanical Society 132, 458-470.

Stiell, W.M., Robinson, C.F., Burgess, D., 1994. 20-year growth of white pine following commercial improvement cut in pine mixedwoods. Forestry Chronicle 70, 385-394.

Vermont Department of Forests, P., and Recreation, 2001. 2001 Forest Health Highlights, p. 3.

Vimmerstedt, J.P., 1962. Southern Appalachian white pine plantations site, volume, and yield. USDA Forest Service, Southeastern Forest Experimental Station, p. 15.

Wang, X.B., Deng, J., Zhang, J.T., Zhou, Q.S., Zhang, Y.Z., Wu, S.A., 2015. DNA barcoding of common soft scales (Hemiptera: Coccoidea: Coccidae) in China. Bulletin of Entomological Research 105, 545-554.

Wantuch, H.A., Havill, N.P., Hoebeke, E.R., Kuhar, T.P., Salom, S.M., 2019. Predators associated with the pine bark adelgid (Hemiptera: Adelgidae), a native insect in Appalachian forests, United States of America, in its southern range. Canadian Entomologist 151, 73-84.

Watson, W.Y., Underwood, G.R., Reid, J., 1960. Notes on *Matsucoccus macrocicatrices* Richards (Homoptera: Margarodidae) and its association with *Septobasidium pinicola* Snell in eastern Canada. Canadian Entomologist 92, 662-667.

Wendel, G.W., Smith, H.C., 1990. *Pinus strobus* L. eastern white pine. In: Burns, R.M., Honkala, B.H. (Eds.), Silvics of North America. USDA Forest Service, Washington DC, pp. 476-488.

Werren, J.H., Baldo, L., Clark, M.E., 2008. *Wolbachia*: master manipulators of invertebrate biology. Nature Reviews Microbiology 6, 741-751.

Wilson, R.W., 1978. Silvical Characteristics of Eastern White Pine. USDA Forest Service, Forest Insect and Disease Leaflet.

Wilson, R.W., McQuilkin, W.E., 1963. Silvical characteristics of eastern white pine (*Pinus strobus*). USDA Forest Service Northeastern Forest Experimental Station Research Paper 13, 1-28.

Yardeni, A., Rosen, D., 1990. Wind dispersal and pattern of colonization of *Ceroplastes floridensis* on citrus in Israel (Homoptera: Coccoidea: Coccidae). In, Sixth International Symposium of Scale Insect Studies, Cracow, Poland, pp. 125-128.

Zinck, J.W.R., Rajora, O.P., 2016. Post-glacial phylogeography and evolution of a wide-ranging highly-exploited keystone forest tree, eastern white pine (*Pinus strobus*) in North America: single refugium, multiple routes. Bmc Evolutionary Biology 16.

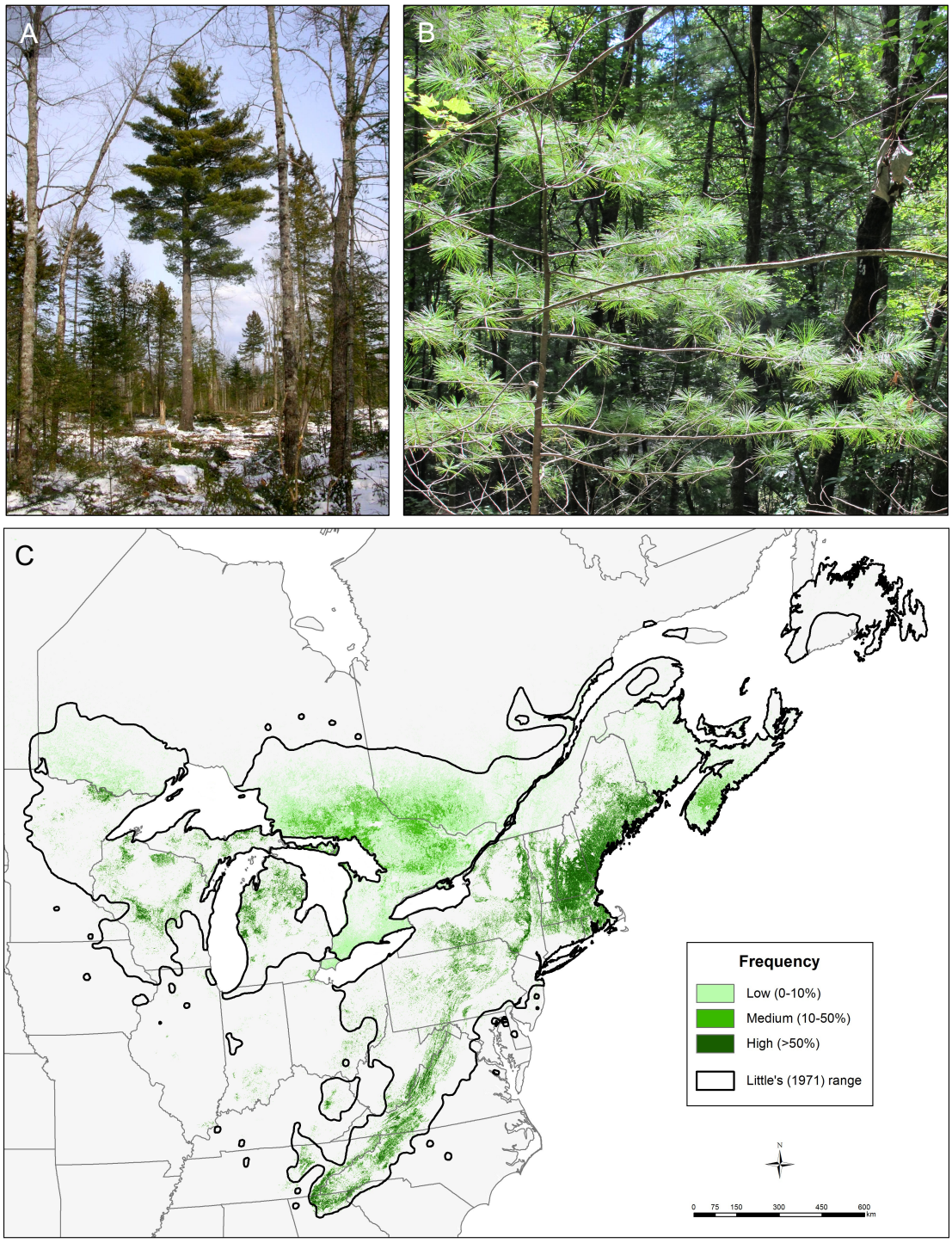

Figure 1.1. A healthy, non-symptomatic (A) sawtimber- and (B) sapling-sized eastern white pine (*Pinus strobus* L.). (C) Native range of *P. strobus* in North America. Frequency refers to the percentage of *P. strobus* comprising total tree species composition for individuals ≥12.7 cm diameter at breast height (250 m resolution). Little's (1971) published range is indicated by bold line. Figure produced by Anthony Elledge (USDA-FS-FHP) using Forest Inventory and Analysis (FIA) data and Canadian Forest Service Inventory data. Photo credit for A: Maine Forest Service.

Figure 1.2. Hallmark symptoms of eastern white pine dieback, including (A) resin exuding from bark fissures, (B) needle flagging, and (C) bottom-up branch dieback, eventually leading to (D) highly reduced crowns. Novel symptomology can lead to mortality in mature trees (E) and in young saplings and seedlings (F-G). Photo credits: Lori Chamberlain (Virginia Department of Forestry, B, E, and G), Joe O'Brien (USDA Forest Service, C and D).

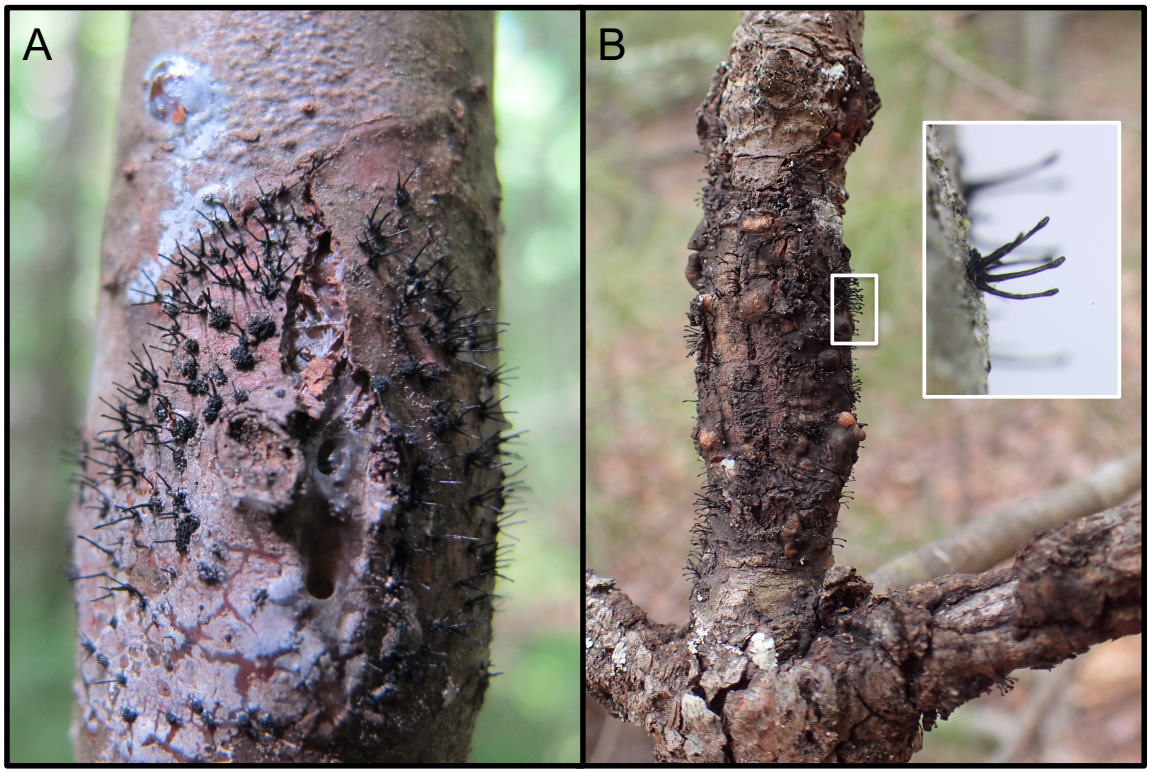

Figure 1.3. (A) Early-stage Caliciopsis canker with characteristic fruiting structures, causing a fissure in the tree bark and resin flow. (B) Late-stage Caliciopsis canker girdling a young *P. strobus* stem. Inset shows *Caliciopsis pinea* ascocarps, which are the characteristic sexual fruiting bodies.

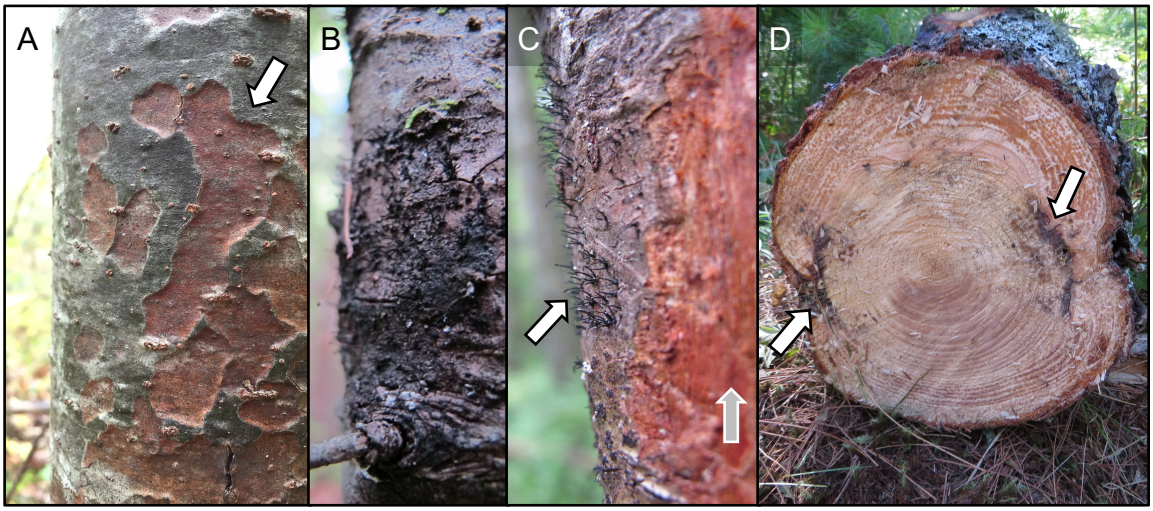

Figure 1.4. Varied appearance of Caliciopsis cankers can include (A) reddish/brownish depressions in the bark, (B) roughened bark above or below branch whorls, and (C) small bark cracks populated with *C. pinea* ascocarps (white arrow), causing necrotic tissue below the exterior bark layer (gray arrow). If the tree calluses over these surface cankers, necrosis persists internally in the xylem tissue (D), which diminishes timber value. Photo credit: Kara Costanza (University of New Brunswick, D).

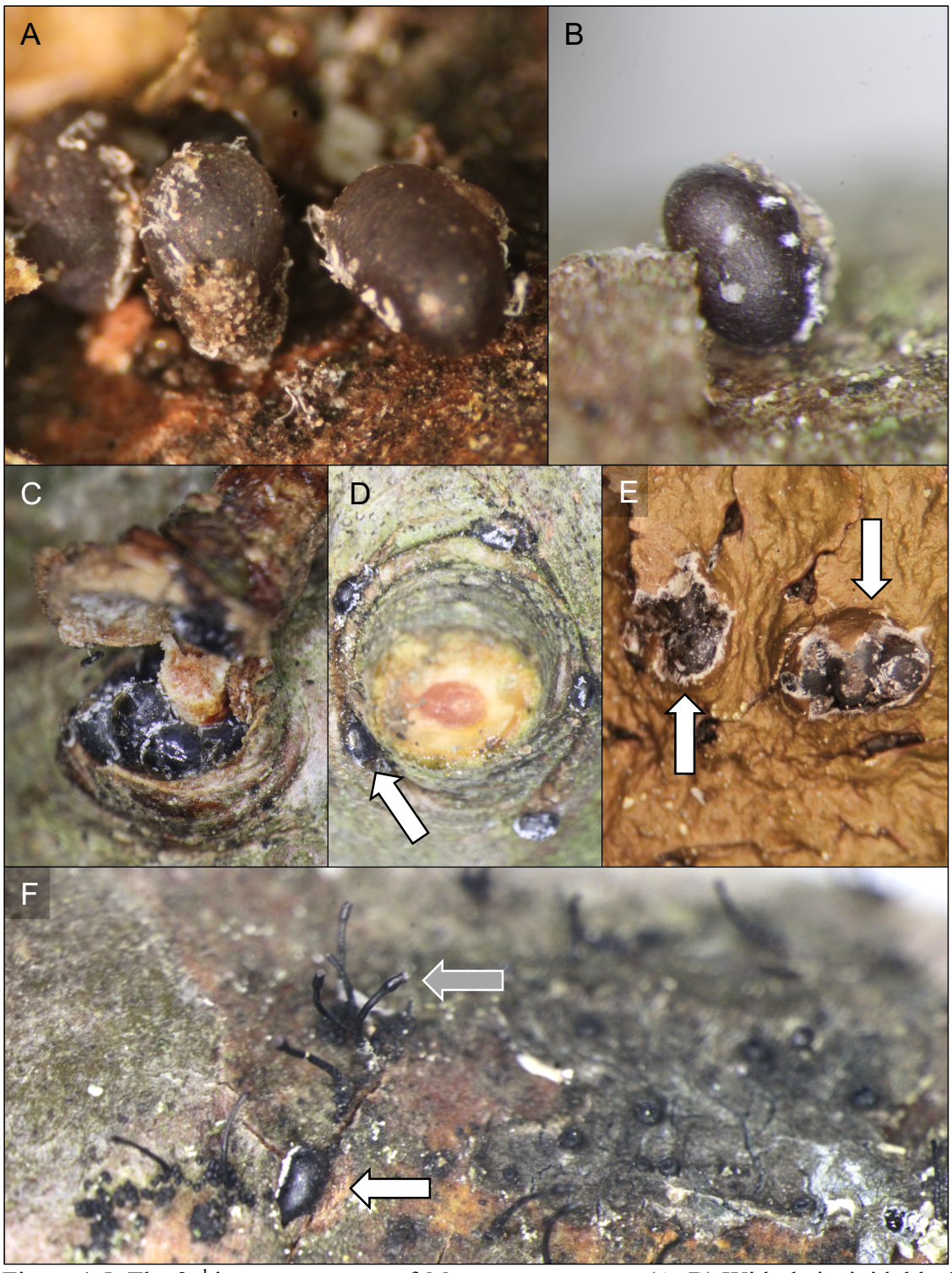

Figure 1.5. The 2nd instar cyst stage of *M. macrocicatrices*. (A, B) With their rigid, black, shiny cuticles, they resemble black pearls. They are extremely difficult to find in the field with the naked eye, only reaching ~1.5 mm at their largest. They colonize hidden, tight spaces on *P. strobus*, such as within nodes (C), the edge of branch crotches (D), under lichen (E), and along the edge/within Caliciopsis cankers (F). The white arrow shows the *M. macrocicatrices* cyst and the gray arrow shows *C. pinea* ascocarps.

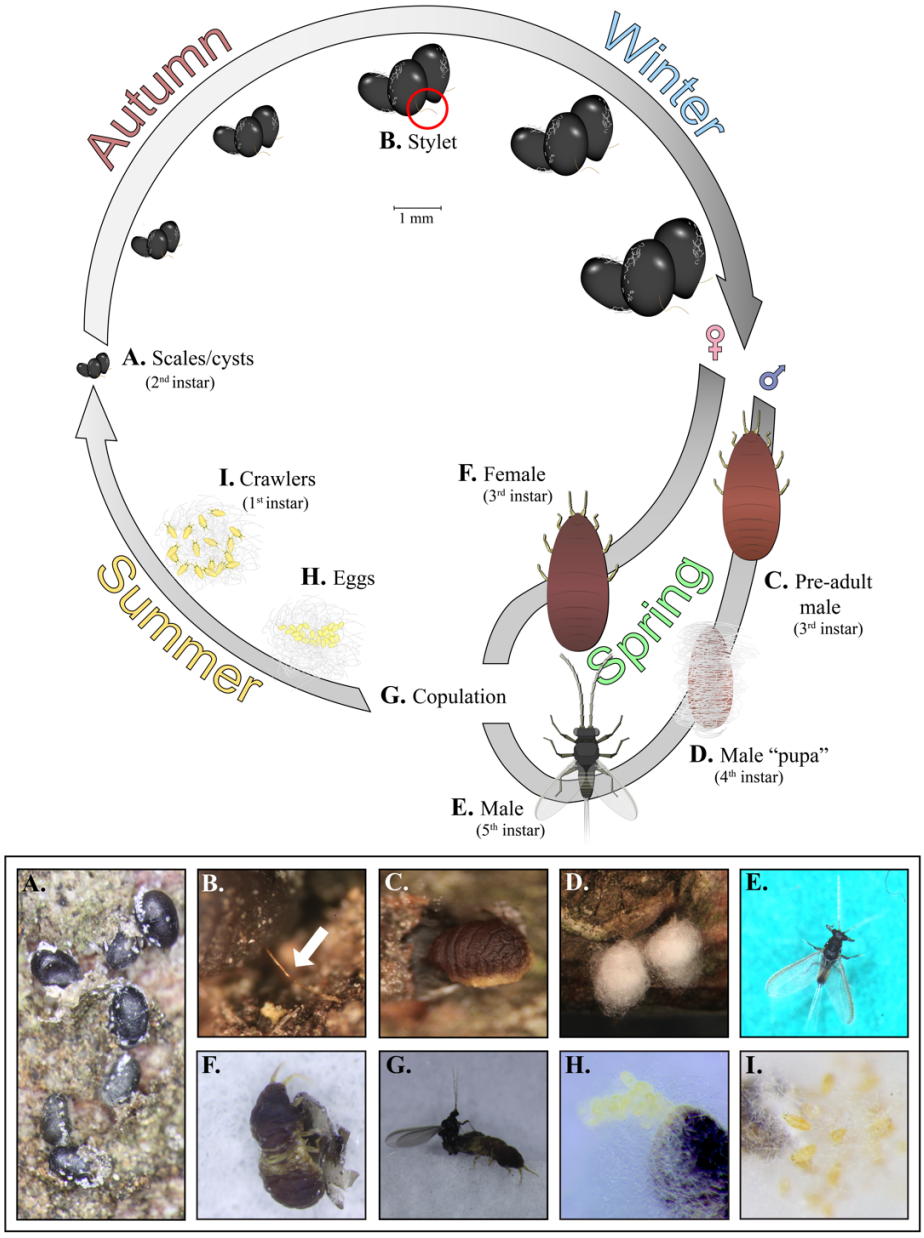

Figure 1.6. Life cycle of the eastern white pine bast scale, *Matsucoccus macrocicatrices* Richards. (A) Second instar nymphs, commonly called scales or cysts, are sessile feeders. Scales extract vascular fluids with (B) long, thin, copper-colored stylets (mouthparts), and beginning in July-September, they steadily grow until spring in Southern Appalachian sites (one-year cycle) or until the following spring at higher latitudes (two-year cycle). Males emerge first as (C) pre-adults and subsequently undergo a metamorphosis resembling that of a holometabolous insect (complete metamorphosis), enclosing themselves in (D) cocoon-like structures to protect their "pupal" instar. They ultimately develop into (E) a winged adult male. The end of adult male development coincides with the emergence of (F) adult females. (G) Mating occurs and females (H) lay eggs in a silken mass. After 2-3 weeks of incubation, (I) first instar "crawlers" hatch. In the summer, crawlers will disperse and settle new feeding sites on *P. strobus*. They subsequently molt and begin their second instar development (Costanza et al. 2018).

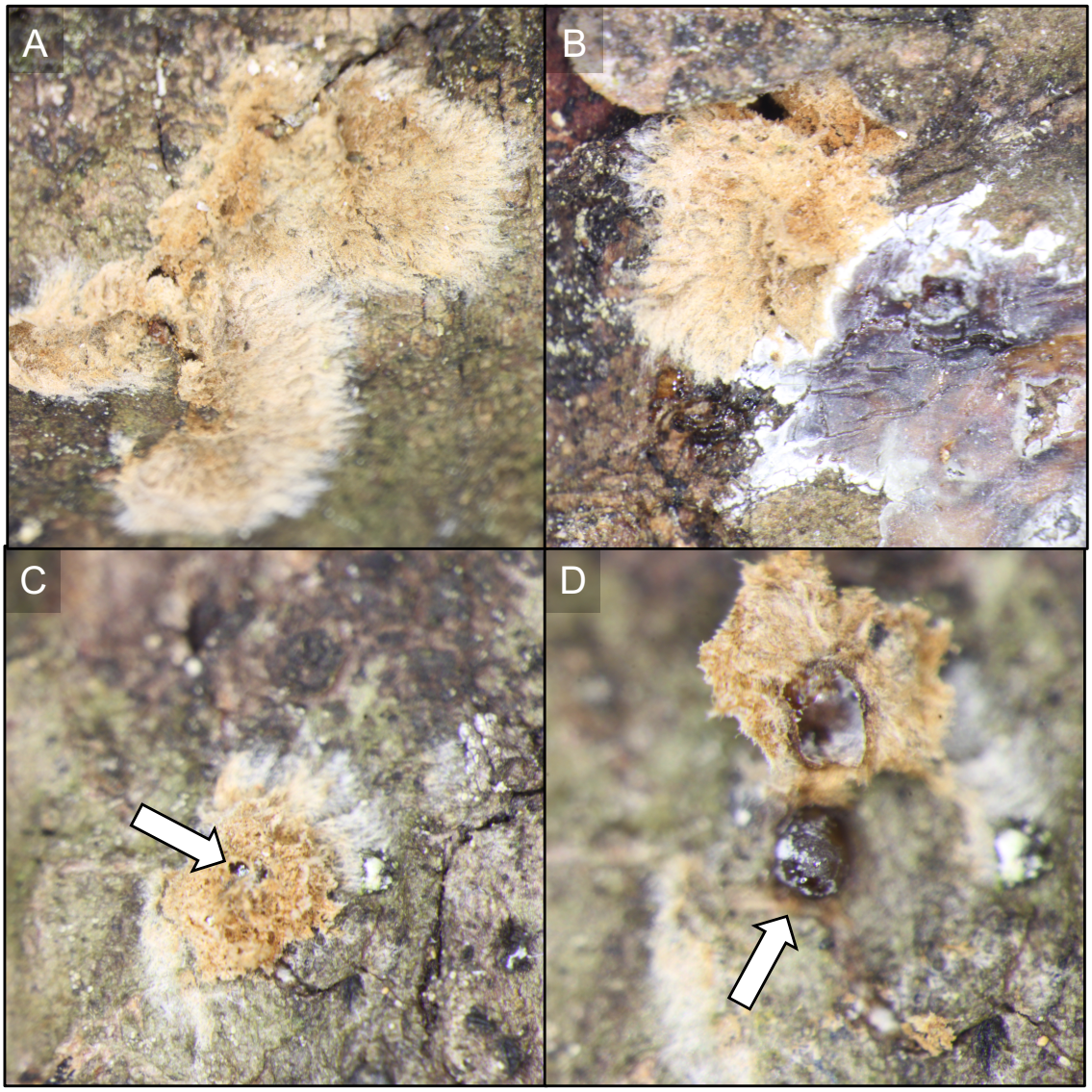

Figure 1.7. The fungus, *Septobasidium pinicola* Snell (A-D), does not harm *P. strobus*. Instead, it develops epiphytic mats on the bark that cover *M. macrocicatrices* cysts (C and D, white arrows), and parasitizes them for its nourishment. Watson et al. (1960) proposed this was a beneficial mutualism, where a proportion of the insects, left unharmed under the mats, gain protection. However, we now find the majority of *M. macrocicatrices* living free of *S. pinicola.* The obligatory nature of this mutualism may have been an overstatement by Watson et al. (1960), or perhaps there has since been a change to this ecological interaction.

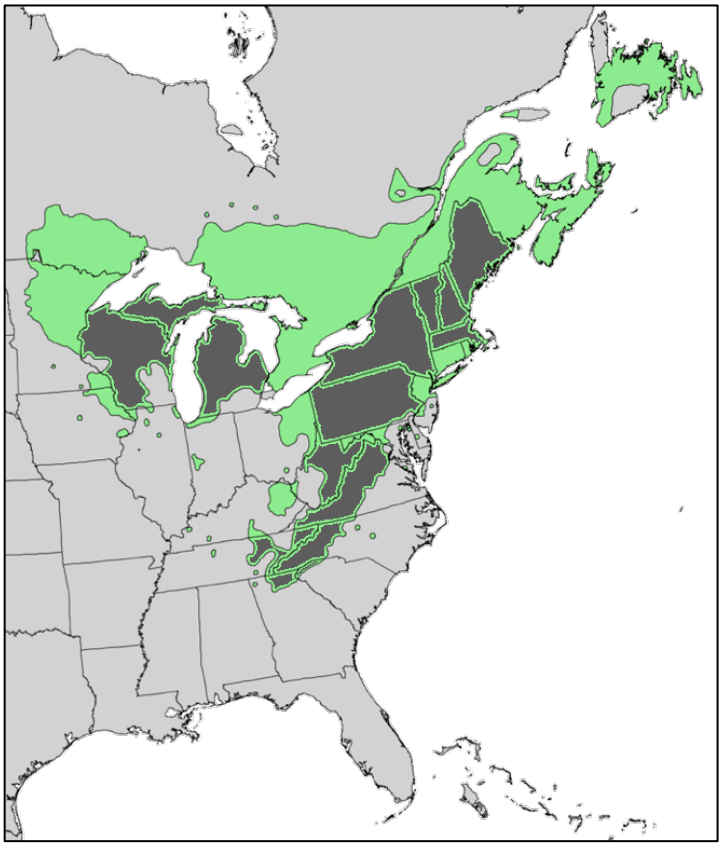

Figure 1.8. The confirmed extent of *M. macrocicatrices* in the USA as of 2018 shaded in gray.

CHAPTER 2

TREE-LEVEL DISTRIBUTION OF A NOVEL INSECT-PATHOGEN COMPLEX AND ITS POTENTIAL CONTRIBUTION TO EASTERN WHITE PINE DIEBACK[2]

[2] Whitney T.D., Cram M.M., Barnes B.F., Yao J., Lucardi R.D., Gandhi K.J.K., 2018. Tree-level distribution of a novel insect-pathogen complex and its potential contribution to eastern white pine dieback. *Forest Ecology & Management* 423, 3-17. DOI: 10.1016/j.foreco.2018.02.002. Reprinted here with permission of the publisher.

Abstract

Bottom-up branch dieback and sapling mortality of eastern white pine (*Pinus strobus* L.) has been observed range-wide during the last two decades. Observational studies thus far implicate these symptoms to an insect, the eastern white pine bast scale (*Matsucoccus macrocicatrices* Richards), and a canker-forming, fungal pathogen, *Caliciopsis pinea* Peck. The scale insect was historically considered an innocuous herbivore of *P. strobus* restricted only to New England and Canada but is now found in high densities on symptomatic trees, in close association with Caliciopsis canker, and in almost every region where the host grows. We sampled branches and boles of *P. strobus* in the Southern Appalachians to better understand the distribution of the insect-pathogen complex on individual trees and among size classes. Results indicate distinct patterns, as branches of poletimber, boles of saplings, and branches lowest in the canopy harbored the greatest numbers of *M. macrocicatrices* and had the highest proportional Caliciopsis canker area. The incidence of scales and cankers was generally highest on older tissue with high percent lichen cover, but with thinner outer bark. Tree-level distribution of *M. macrocicatrices* and Caliciopsis canker was non-random and in fact mirrored the observed dieback patterns reported for *P. strobus*, indicating that these two organisms may be important contributors to tree dieback and mortality in the southeastern USA.

Keywords: *Caliciopsis pinea*, *Matsucoccus macrocicatrices*, *Pinus strobus*, tree mortality, white pine dieback

2.1 Introduction

Trees are spatially and temporally heterogeneous habitats for animals, with abiotic and biotic conditions varying at multiple strata (e.g., root, bole, branch, twig, cone, and foliage levels). As hosts, structural heterogeneity in trees influences herbivorous communities and populations (Lawton 1983). Associated with varying tree size, age, and zonation are marked differences in physiology (e.g., tissue type, volume, texture, nutritional quality, and secondary defensive compounds), micro-environmental conditions (e.g., temperature and moisture), and natural enemy load (e.g., predators and parasitoids), which directly or indirectly affect the spatial colonization of herbivores (Lawton 1983, Ulyshen 2011, Wardhaugh 2014). Many studies have investigated differences in herbivore communities within and across trees (e.g., Maguire et al. 2014, Plewa et al. 2017, Weiss et al. 2016), but fewer studies are available that assess such distributions of forest pest populations (e.g., Wardhaugh et al. 2006), including those associated with tree-killing pathogens. An understanding of tree-level colonization patterns may elucidate if a link exists between observed tree symptoms and the contributing pests and pathogens.

A novel dieback phenomenon is posing ecological and economic threats to eastern white pine (*Pinus strobus* L.) across its entire range in eastern North America (Asaro 2011, Costanza et al. 2018, Lombard 2003). The current symptoms are unrelated to known threats of *P. strobus*, white pine blister rust (*Cronartium ribicola* Fisch) and white pine weevil (*Pissodes strobi* Peck), but rather include a progressive bottom-up thinning of the crown, where lower branches die steadily until the live canopy is reduced to terminal branches and the leader (Schulz et al. 2018a, b). Accompanying this pattern of

branch dieback is the presence of characteristic cankers on branches and boles of trees. Severe resin outflow from cankers is common, especially on large diameter trees, which leads to decreased wood quality (Costanza 2017). The cankers also cause girdling of stem and branches, thus killing trees (Overholts 1930). Seedlings and saplings are particularly vulnerable to mortality, but the phenomenon has been reported from all diameter classes, thus threatening the resilience and sustainability of *P. strobus* as a canopy species in eastern forests (Asaro et al. 2018, Constanza et al. 2018).

A variety of common fungi have been isolated from cankers on symptomatic *P. strobus*, with *Caliciopsis pinea* Peck as the most ubiquitous and virulent species (Cram et al. 2009, Ray 1936, Schulz et al. 2018b). A native pathogen primarily of *P. strobus*, *C. pinea* has been known throughout the eastern United States to be associated with sapling and branch mortality (Overholts 1930, Ray 1936), but never to an extent that would significantly threaten the health of the host-tree species across its range. Recent studies have established that *C. pinea* is strongly associated with current eastern white pine dieback and is now considered a main contributing factor throughout the host range (Lombard 2003, Munck et al. 2015, Schulz et al. 2018b). The fungal hyphae of *C. pinea* establish under the tree bark, grow into the vascular cambium, and kill host tissue. This infection results in the steady expansion of necrotic lesions from which annual crops of ascocarps, the distinctive hair-like fruiting bodies, are produced on the external surface of trees (Figure 1.3a, b) (Funk 1963, Ray 1936). Cankers sometimes coalesce with each other and will girdle branches and young stems (e.g., Figure 1.3b). The epidemiology of the disease (Caliciopsis canker, hereafter) is not fully understood, but it has long been speculated that *C. pinea* requires an infection court in the form of an old lenticel, a

natural bark crack, or an insect feeding wound to first colonize a host (Funk 1963). In 2006, forest health specialists in Virginia discovered a scale insect, known as the eastern white pine bast scale, *Matsucoccus macrocicatrices* Richards (Hemiptera: Matsucoccidae), along the edge of Caliciopsis cankers; it has been suggested that *M. macrocicatrices* is likely facilitating *C. pinea* in its initial infection stage (Asaro 2011, Mech et al. 2013). Historically believed to be a benign sap-sucking insect of *P. strobus*, *M. macrocicatrices* is now strongly associated with both the dieback symptoms and the presence of Caliciopsis canker range-wide (Schulz et al. 2018b).

Matsucoccus macrocicatrices is native to North America and is a specialist on *P. strobus* (Richards 1960). The vast majority of its lifecycle is spent in its feeding stage as a sessile, 2nd instar juvenile: a black, eyeless, and legless cyst (Figure 1.5). With their relatively long piercing-sucking mouthparts (stylet) inserted into the outer phloem (bast), the insect cysts extract sugar-rich, vascular fluids for nourishment (Figure 1.6b). Adult insect emergence occurs annually during the spring in the Southern Appalachians (Mech et al. 2013, Whitney pers. obs.), but occurs biennially at higher latitudes, presumably because winter is more prolonged (Watson et al. 1960). Winged males mate with quiescent females, and the females lay eggs in early summer. First instar "crawlers" hatch from eggs in summer and act as the main dispersal stage, hypothesized to utilize the wind to move to new trees, as in other scale insects (Bean and Godwin 1995, Gullan and Kosztarab 1997). After locating a suitable feeding site in late summer, often within bark crevices, under lichen, or along the edge of cankers, they insert their mouthparts into the tree and undergo a molt into the 2nd instar cyst (see Figure 1.6 for additional lifecycle details). Mech et al. (2013) hypothesized that the feeding wounds created by *M.*

macrocicatrices during its cyst stage facilitate *C. pinea* infection of its host. This proposed pathway of infection is similar to the beech bark disease complex, where two native fungi (*Neonectria faginata* Lohm. & Watson and *N. ditissima* Tul. & C. Tul.) exploit the minute feeding wounds of the introduced beech scale (*Cryptococcus fagisuga* Lindinger), leading to infection and loss of American beech (*Fagus grandifolia* Ehrh.) (Houston 1994). No mechanistic work has yet investigated if *M. macrocicatrices* feeding behavior aids and/or expedites pathogen infection, and thus it would be premature to attribute this as the cause of the dieback. However, there is correlative evidence to suggest there is an association between the two organisms and that together they contribute to extensive canker formations on branches and bole (Mech et al. 2013, Schulz et al. 2018b).

Due to their historical reputations as negligible damaging agents, there is currently a paucity of substantive research on *M. macrocicatrices*, *C. pinea*, and their association (hereby referred to as the insect-pathogen complex). To assist with effective management or mitigation strategies, the basic biology of these organisms must be elucidated, including mechanisms of dispersal, colonization, and the relationship between *C. pinea* infection and insect feeding. Particularly urgent is the quantitative assessment of their distribution on symptomatic trees to infer where they may be preferentially colonizing. Such information may assist managers to identify particular trees within a stand or specific areas on a tree that are especially susceptible to insect-pathogen colonization for specific control prescriptions.

Our goal was to determine the distribution of both focal organisms, *M. macrocicatrices* and *C. pinea*, within and across *P. strobus* trees. We assessed the

incidence of the insect-pathogen complex on branches and boles according to: (1) tree size class (saplings, poletimber, and sawtimber) and (2) vertical distribution within the canopy (lower, middle, and upper crown). Our field observations indicate that intermediate-sized poletimber experience the most severe branch dieback and small saplings experience the most mortality. If the insect-pathogen complex were a driving force behind these symptoms, then the incidence of both organisms would be highest on branches of poletimber trees and on boles of saplings as compared to other size classes. The pattern of branch dieback across size classes also occurs from the bottom up on individual trees. Thus, higher incidence of *M. macrocicatrices* and Caliciopsis canker was also expected on branches in the lower, rather than the upper canopy. We also compared incidence of the insect-pathogen complex according to: (3) horizontal distribution along the branches (proximal, medial, and distal), (4) vertical distribution on boles (lower, middle-lower, middle-upper, and upper), (5) branch-section diameter (an indication of tissue age), (6) bark thickness, and (7) percent lichen cover. We evaluated these variables in addition to our main objectives to determine if *M. macrocicatrices* and *C. pinea* may colonize branches from the outside in or inside out, if vertical bole level or bole bark thickness is more of a limiting factor, and if the presence of lichen increases the incidence of colonization.

2.2 Methods

2.2.1 Field sites

Sampling was conducted at the southernmost extent of the Appalachian Mountains and range of *P. strobus* within the Chattahoochee National Forest in northern

Georgia, USA (Figure 2.1). Five sites were selected and were separated by at least 10 km (Table 1). In addition to *P. strobus*, forest overstory consisted of American beech (*F. grandifolia*), dogwood (*Cornus florida* L.), eastern hemlock (*Tsuga canadensis* (L.) Carrière), oaks (*Quercus* spp.), red maple (*Acer rubrum* L.), sweetgum (*Liquidambar styraciflua* L.), and tulip poplar (*Liriodendron tulipifera* L.). The understory was dominated mostly by *Rhododendron* spp. and mountain laurel (*Kalmia latifolia* L.). Annual precipitation of this region ranges from 150 to 300 cm, average daily temperature ranges from 8 to 16 °C, and the dominant soil orders are Inceptisols and Ultisols with the general textural class being loamy (USDA 2006).

2.2.2 Sampling of trees

From October 2015 through February 2016, we established three, 1-ha circular plots spaced >100 m apart within each site. Wintertime sampling was necessary for the most reliable observations, because the minute *M. macrocicatrices* cysts are largest during this time (Mech et al. 2013). A sapling- (DBH <12.5 cm), a poletimber- (≥12.5 cm-30 cm), and a sawtimber- (DBH >30 cm) sized *P. strobus* were felled within each plot. We selected 15 trees within each size class from across the five sites for a total of 45 trees in the study. Our selection criteria prioritized size class requirements and felling feasibility, and thus dieback severity of trees ranged from light to severe (dieback rating of 2 to 4, per Schulz et al. 2018a, b).

Samples were taken from both the bole and branches of felled trees (Figure 2.2). First, the vertical length of the living tree canopy was measured and divided into three equal sections for branch sampling: lower, middle, and upper crown. Two live branches

were selected from each of the three sections. After removing axillary twigs, the length of each branch was measured and when >3 m, it was partitioned into three equal parts, representing the horizontal stratification of the canopy. One-meter segments were removed from each third of each branch and were categorized as proximal (adjacent to bole), medial (middle of the branch), or distal (furthest from bole). Some branches were <3 m in length, especially those from saplings and branches of the upper canopy from poletimber-sized trees. In these instances, branch segments were categorized only as proximal or distal. We sampled between 7 to 18 branch segments from each of the 45 trees for a total of 604 samples.

To obtain representative bole samples, the entire height of each tree was measured and then partitioned into four equal sections. A 12.5 × 12.5 cm sample (156.25 cm^2 area) was removed from the bole surface from the middle of each of the four partitions (lower, middle-lower, middle-upper, and upper), providing 180 total bole samples for the assessment of vertical bole distribution. All bole and branch samples were transported to the University of Georgia, Athens, and kept at 4 °C until scale counts and canker measurements were made to ensure that the immature *M. macrocicatrices* cysts remained dormant.

2.2.3 Assessment of Matsucoccus macrocicatrices and Caliciopsis pinea

Counts of live cysts (current generation) and exuviae, the shed exoskeletons from which adults emerged in recent years (previous generations), were made on each bole and branch sample under a microscope. Surface areas of Caliciopsis cankers were visually estimated using a 1-cm^2 gridded plastic transparency. When *C. pinea* fruiting bodies or

asexual structures were present, we used a vegetable skinner to remove the outer bark and reveal the larger extent of necrotic, subcortical tissue, which allowed for a more precise estimate of canker area. Bark thickness was measured and percent lichen cover was estimated in 25% increments on each bole and branch sample.

2.2.4 Non-Caliciopsis canker survey

A subset of cankers without the distinguishing *C. pinea* fruiting bodies were collected from ten of the study trees and then plated on three different types of general media to isolate and identify other potential fungal pathogens in our study. Up to four cankers without the characteristic signs of *C. pinea* were sampled per branch for a total of 43 cankers for isolation and potential identification. The branches were stored at 4 °C until processed. For isolation, the surface bark was removed with a scalpel and the cankered tissue surface was sterilized for 10 s with 95% ethanol, followed by 4 min in 1.08% sodium hypochlorite solution (18 mL of 6% bleach in 100 mL water) (Blodgett and Stanosz 1997). Surface sterilized tissues were then washed in sterile water for 1 min and blotted dry with sterile paper towels. Each canker sample was divided evenly into 5-mm^2 diseased tissue sections and plated on three different media types: (1) modified Nash-Snyder media (Nelson et al. 1983), (2) pine needle agar (PNA) media (Blodgett et al. 2003), and (3) potato dextrose agar with streptomycin and terigitol (PDA+S+T) media (Steiner and Watson 1965).

Plated samples were incubated for four weeks at 20 °C, with weekly observation for identification and/or transfer of isolates to other media. Samples with unidentifiable mycelium isolates were transferred to carnation-leaf water agar (Nelson et al. 1983) or

pine needle agar to induce spore production for identification. Plates from secondary transfers were visually inspected weekly for another four weeks. Frozen samples of isolates representing each genus or unknown morphological types were sent for DNA sequencing of the internal transcribed spacer (ITS) region, which is widely used in fungal barcoding. In a few cases, there were amplification issues and the small subunit was used. We used primers ITS1F (5′-CTTGGTCATTTAGAGGAAGTAA-3′) (Gardes & Bruns 1993) and ITS4 (5′-TCCTCCGCTTATTGATATGC-3) (White et al. 1990) to amplify the ITS rDNA region under the following cycling conditions: 85 °C for 2 min, 95 °C for 95 sec, and then 36 cycles of 58 °C for 1 min, 72 °C for 80 sec, and 95 °C for 70 sec, followed by a 52°C for 1 min and 72°C for 15 min. Sequences were compared to GenBank and other accessions using BLAST searches (NCBI). Species identifications from ITS barcodes were assigned when isolates had ≥99% matching ITS sequences (Gazis et al. 2011). In the cases where an isolate did not meet these criteria, the isolate morphology and relative sequence similarity were combined to infer the most probable genus, family, or order. Of the 43 cankers initially sampled, 36 underwent further identification. Seven samples were removed from the survey, because of mold contaminants (e.g., *Penicillium* spp.) or lichenized fungi (e.g., *Sarea* spp.).

2.2.5 Statistical analyses

Surface areas of branch and bole samples were used to standardize counts of *M. macrocicatrices* and area of cankers. Surface area of each branch segment differed and required estimation using the area of a truncated cone: $A = s\pi (r_1+r_2) + \pi r_1^2 + \pi r_2^2$, where s is the slant length (length of the branch segment) and r_1 and r_2 are the radii of each

circular end. Surface area of all bole samples was ~156.25 cm^2. The adjusted *M. macrocicatrices* counts were analyzed as number of scales per m^2 (i.e., scale density) and the adjusted Caliciopsis canker areas were analyzed as proportional canker area (i.e., the ratio of canker area to branch/bole sample surface area). Measurements of scale density only included counts of the current generation (live cysts), however we included both the current and previous generations (exuviae) of *M. macrocicatrices* in the presence/absence data.

Branch and bole data are discrete sets and thus their analyses were conducted separately. Both branch and bole samples are nested within trees, but it was necessary to treat the sub-factors and whole unit factor as experimental units due to the variables of interest. Size class is a tree-level factor and all other variables we assessed (canopy level, lichen cover, etc.) are sample-level factors. Due to this and the uneven sampling of branch samples among trees, it was necessary to partition our analyses further within branch and bole data sets, explained in detail below: assessment of the whole unit factor (tree) and assessment of the subunit factor (branch or bole sample).

2.2.5.1 Branch sample analyses

To assess if the incidence of the insect-pathogen complex differed on branches across tree size classes, we summed the scale counts, canker areas, and total analyzed branch surface area from each tree, and standardized our response variables as described above (N = 45). This procedure was done to accommodate the uneven sub-sampling of branch segments among trees. To satisfy the assumptions of normality, scales per m^2 were log (+1) transformed, proportional canker area was log (+ 0.005) transformed, and

then an analyses of variance (ANOVA) test was performed, blocking for site. Further, generalized linear regression was conducted with scale density as the response variable with both tree DBH and DBH^2 (to investigate a possible non-linear trend) as the explanatory variables, using the negative binomial distribution to avoid overdispersion (referred to as negative binomial regression herein). Since canker area was a proportion, the beta distribution was used for this generalized linear regression (referred to as beta regression herein) using the R package 'betareg' (Zeileis et al. 2016). To accommodate zero values, the response variable was first transformed per the adjustment recommended in Smithson and Verkuilen (2006): $y'' = [y\ (N - 1) + 0.5] / N$, where N is the sample size.

To determine the effect of canopy level (lower, middle, and upper) on scale density and canker area on branch segments, we conducted negative binomial regressions of scale counts (offset by branch segment area) and beta regressions of proportional canker area, using both site and tree size class as blocking factors (N = 604). Distal diameter of branch segments was also included in these tests as a proxy for age of tissue to elucidate any interactive effects between tree-level distribution and maturity of the host resource on incidence of the organisms. To determine the effect of branch segment level (proximal, medial, and distal), it was necessary to conduct separate negative binomial and beta regressions, because there was a deficiency in medial branch segments of the upper canopy, especially in saplings. Beta regressions required the Smithson and Verkuilen (2006) adjustment to proportional canker area. We assessed whole model fit using likelihood ratio tests with the *lrtest* function from the R package 'lmtest' (Hothorn et al. 2017), and Wald tests assessed significance of the factor levels, using *middle* as a reference for canopy level (vertical distribution) and *medial* as a reference for segment

level (horizontal distribution). We selected these references to make intuitive contrasts between the intermediate factor level (e.g. *middle*) and the extreme factor levels (e.g. *lower* and *upper*). We also conducted individual logistic regressions, blocking for site and size class, to determine whether lichen cover affected the likelihood of *M. macrocicatrices* and Caliciopsis canker presence on branch segments.

2.2.5.2 Bole sample analyses

Scales and cankers were absent from a number of bole samples, and thus simple transformations failed to meet the assumption of normality. To compensate for the skewed distributions and high overdispersion, we conducted hurdle regressions for scale count data and logistic regressions for canker presence/absence. Hurdle regressions are two-part models that test zero counts and positive counts of a dataset separately. Using the *hurdle* function from R package 'pscl' (Jackman 2017), our hurdle regressions first performed a logistic regression on the presence/absence of scales, and second performed a negative binomial regression on the scale counts (offset by bole sample surface area) using the truncated dataset of only positive values.

To determine if the incidence of *M. macrocicatrices* and Caliciopsis canker differed on tree boles as based on DBH, scale counts (offset by sampled bole area) and canker counts were summed per tree (N = 45). We performed a hurdle model for scale insects and a logistic regression for canker presence/absence using DBH as the explanatory variable and site as a blocking factor. To determine if scale presence and density varied with vertical bole level, bark thickness, and/or percent lichen cover of individual bark samples (N = 180), we included these variables in a single negative

binomial hurdle regression, blocking for site and tree size class. We similarly included bole level, bark thickness, and percent lichen cover into a logistic regression with canker presence/absence to evaluate which of these variables best predicted *C. pinea* presence on bole samples. All regressions were assessed with likelihood ratio tests to determine whole model fit and Wald tests were used to determine factor significance.

2.3 Results

2.3.1 Branch sample analysis

Scale densities (number of scales per m^2) on assessed branches per tree ranged from 0-3.42 per m^2 on saplings, 1.03-12.96 per m^2 on poletimber, and 0.24-3.83 per m^2 on sawtimber-sized trees. The mean density on branches of poletimber trees was 3.8 times greater than on saplings and 2.4 times greater than on sawtimber trees ($F_{2,38}$ = 13.58, $P < 0.001$) (Figure 2.3a). The quadratic DBH term of the generalized linear regression was also significant ($Z = -3.23$, $df = 1$, $P < 0.001$), indicating higher scale insect density on trees of intermediate size (poletimber). Proportional canker area (the ratio of Caliciopsis canker area to branch surface area) on branches ranged from 0-12% on saplings, 1-12% on poletimber, and 0.1-26% on sawtimber trees. Proportional canker area on branches was greater on poletimber trees than on saplings, but neither poletimber trees nor saplings significantly differed from sawtimber trees ($F_{2,38} = 4.81$, $P = 0.014$) (Figure 2.3b). Beta regressions revealed proportional canker area was not significantly correlated with tree DBH ($Z = 1.86$, $df = 1$, $P = 0.06$) or DBH^2 ($Z = -1.70$, $df = 1$, $P = 0.08$).

Scale insect density on branch segments differed across canopy levels (χ^2 = 206.43, df = 5, P < 0.001), with up to 30.68, 24.18, and 9.92 per m^2 in the lower, middle, and upper canopy, respectively. With the middle canopy as reference, we found the highest and lowest scale densities were respectively, on the branches of the lower and upper canopy (Z = 1.04, df = 1, P < 0.001; Z = -0.64, df = 1, P = 0.03). There was no interaction between branch segment diameter and canopy level, but scale density increased as the diameter of branch segments also increased (Z = 0.65, df = 1, P < 0.001). Proportional canker area on branch segments also varied with canopy level (χ^2 = 169.05, df = 5, P < 0.001), with up to 76%, 15%, and 13% in the lower, middle canopy, and upper canopy, respectively. There was an interaction between canopy level and branch segment diameter. With the middle canopy as reference, proportional canker area was greatest in the lower canopy, where branches are generally greater in diameter (Z = 0.41, df = 1, P < 0.001), and lowest in the upper canopy, where branches are generally smallest in diameter (Z = -0.30, df = 1, P = 0.04).

Scale insect density (χ^2 = 56.66, df = 2, P < 0.001) and proportional canker area (χ^2 = 23.51, df = 2, P < 0.001) also differed horizontally along each branch. With medial branch segments as a reference, lower scale density was observed on distal branch segments, furthest from the bole (Z = -4.45, df = 1, P < 0.001), whereas higher scale density was found on the proximal branch segments, adjacent to the bole (Z = 2.31, df = 1, P = 0.02). Similarly, less proportional canker area was found on distal branch segments (Z = -3.24, df = 1, P < 0.001), but there were no differences between the proximal and medial segments (Z = 0.41, df = 1, P = 0.68). Percent lichen cover was significantly correlated with the probability of scale (χ^2 = 117.39, df = 1, P < 0.001) and canker (χ^2 =

182.88, df = 1, P < 0.001) presence on all branch segments. Both the scale insect (Z = 7.11, df = 1, P < 0.001) and Caliciopsis canker (Z = 8.57, df = 1, P < 0.001) were more likely to be present on branch segments with greater percent lichen cover.

2.3.2 Bole sample analysis

On the bole, tree DBH was a significant predictor of *M. macrocicatrices* presence according to the hurdle model (χ^2 = 6.29, df = 2, P = 0.04) and Caliciopsis canker presence according to the logistic regression (χ^2 = 5.30, df = 1, P = 0.02). Scale insect presence on the bole increased as tree DBH decreased (Z = -2.03, df = 1, P = 0.04) (Figure 2.4a), but DBH had no effect on scale density (Z = -1.32, df = 1, P = 0.19). Caliciopsis canker presence on boles also increased as DBH decreased (Z = -2.12, df = 1, P = 0.03) (Figure 2.4b).

The hurdle model that tested vertical bole level, bark thickness, lichen cover, and their interactive effects on the presence and density of scale insects fit the data better than the null model (χ^2 = 33.03, df = 16, P = 0.007). The logistic regression revealed no interactive effects, and bark thickness was the only significant variable in the model, revealing that scales were more often found on thinner bark (Z = -2.16, df = 1, P = 0.03) (Figure 2.5a). The binomial regression on the truncated dataset revealed no significant variable for predicting scale insect densities on bole samples. The logistic regression that tested vertical bole level, bark thickness, lichen cover, and their interactive effects on Caliciopsis canker presence did not perform better than the null model (χ^2 = 11.24, df = 8, P = 0.19). When modeled individually, the logistic regression including vertical bole position (χ^2 = 8.31, df = 3, P = 0.04) and the logistic regression including bark thickness

(χ^2 = 5.44, df = 1, P = 0.02) outperformed the null model (Figure 2.5b). Overall, cankers were most often found on the upper-middle sections of boles (Z = 2.83, df = 1, P = 0.005) and on the thinnest bark (Z = -2.30, df = 1, P = 0.02). No scales or cankers were found on bole samples with bark thicker than 1 cm and 0.8 cm, respectively.

2.3.3 Fungal isolations

Of the 36 isolates from cankers without *C. pinea* fruiting structures, two were assigned to the genus *Caliciopsis*. One was identified as *C. pinea* based on the stromata fruiting structures that formed in culture on the canker tissue. The other could not be confidently given a species-level identification based on the ITS region (DQ471039), so it was instead denoted only as a member of the genus *Caliciopsis*. Seventeen isolates belonged to one of two species from the genus *Pestalotiopsis* based on 99% ITS region matches: *P. maculans* (Corda) (KX610327) and *P. chamaeropis* Maharachch (KM199326). Eleven isolates belonged to the genus *Phaeomoniella,* based on spore morphology and 98% ITS region matches to a species within Eurotiomycetes (KM519288) (Damm et al. 2010). Two isolates were identified as *Coniochaeta velutina* (Funkel) (GQ154624) and one isolate was identified as *Clonostahys rosea* (Link) (LT576164). Three isolates without confident ITS region matches were unidentifiable.

2.4 Discussion

Since European settlement in the 18th century, *P. strobus* has experienced a significant increase in many abiotic (e.g., logging and burning) and biotic (exotic and native herbivores and pathogens) stressors (Abrams 2001, Wendel and Smith 1990). The

current dieback and mortality of *P. strobus* is unprecedented and is associated with a previously unknown insect-pathogen complex that has been reported throughout its North American range with varying severity. Field observations indicate that poletimber-sized trees are experiencing the most severe branch dieback, while saplings have the highest levels of mortality (Asaro et al. 2018, Costanza et al. 2018, Schulz et al. 2018a). Hence, we expected to observe greater scale density and canker area on the branches of poletimber and the boles of saplings. In evidence, we found that the distribution of both the scale insects and cankers were consistent with our expectations and hypotheses. These data suggest a link between the insect-pathogen complex and the observed tree dieback and mortality patterns relating to tree size class. Such distributional variation of insects and pathogens based on tree size has also been reported, such as for the scale insect, *Ultracoelostoma assimile* Maskell, on New Zealand beeches (*Nothofagus* spp.) (Wardhaugh et al. 2006) and for stem cankers on a Panamanian canopy tree (*Ocotea whitei* Woodsen) (Gilbert et al. 1994). It is noteworthy that although sawtimber-sized trees currently appear resilient, a higher pest load and/or longer period of time is required for the insect-pathogen complex to girdle and overcome larger diameter trees.

Matsucoccus macrocicatrices and Caliciopsis cankers were also distributed in a bottom-up gradient within trees, consistent with the observed pattern of branch dieback. The lower branches of symptomatic *P. strobus* are commonly the first to die, and subsequent branch deaths occur in an upward progression (Constanza et al. 2018a, Schulz et al. 2018b). We expected and found that the incidence of the insect-pathogen complex reflected this pattern, as there were higher scale densities and a higher proportional canker area in the lower canopy. Such bottom-up canopy distribution patterns have also

been observed in two other *Matsucoccus* species, both of which are important pests. McClure (1976) found that 81% of red pine scales, *M. resinosae* [now *M. matsumurae* (Kuwana)], inhabited the lower canopy of Connecticut red pine (*Pinus resinosae* Aiton), whereas only 7% inhabited the upper canopy. Additionally, Jactel et al. (1996) found the maritime pine scale, *M. feytaudi* Ducasse, was completely absent from the upper third of maritime pine (*Pinus pinaster* Aiton) canopies in southern France, with maximum densities inhabiting intermediate tree heights. Other pathogens can also infect trees from the bottom up, such as the non-native fungus, *Cronartium ribicola* J.C. Fisch. (white pine blister rust), which tends to form cankers in the lower canopy of North American five-needle pines (Schwandt et al. 2013).

Heterogeneity in within- and across-tree distribution may be the result of factors that facilitate insects and pathogens to more easily develop in certain size classes and canopy levels, such as mode of dispersal (e.g., Brown et al. 1997) and/or climatic conditions (e.g., Rowe and Potter 1996). For example, wind-dispersed *M. macrocicatrices* crawlers are most likely to land on the ground and may opt to settle sites within a short crawling distance on lower branches and smaller tree boles. Cooler and moist conditions at low strata under a closed canopy can also be optimal for pathogen establishment. In evidence, Munck et al. (2016) reported a greater incidence of *C. pinea* cankers with increasing stand density, indicating that the amount of host material and specific, micro-habitat conditions may be significant predictors of the insect-pathogen complex occurrence and contribution to dieback symptoms.

Our study suggests that tree bark thickness influences the distribution and patchiness of *M. macrocicatrices* and Caliciopsis canker formations. We found that *M.*

macrocicatrices and cankers were more frequent on thinner than thicker bark, and they were both completely absent on bole samples with bark thicknesses greater than one centimeter. Bark thickness was less variable on branches, as the bark on approximately 90% of the branch samples were less than 0.1 cm thick. We posit that the bark of *P. strobus* becomes an effective barrier to herbivory when it nears and exceeds this 1-cm threshold. The bark is likely too thick at this point for a bast scale's stylet to effectively penetrate and reach the outer phloem cells. Similarly, Wardhaugh et al. (2006) reported that *U. assimile* colonization is a function of host bark thickness, and Jactel et al. (1996) found that *M. feytaudi* was also unable to settle on portions of host trees with the thickest bark. If *M. macrocicatrices* is absent and there are no feeding wounds on these thicker-barked surfaces of the tree, then we speculate that *C. pinea* may also lack the infection court needed to establish and form cankers.

Despite having thin bark, new growth also seemed to be unsuitable for *M. macrocicatrices* settlement, as significantly fewer scale insects were found on younger, upper canopy branches and distal branch segments (which include axillary meristems) when compared to older, lower canopy branches and proximal branch segments. This was further supported by the finding that scale density increased with branch segment diameter (a reflection of the branch tissue's age). Bark texture, rather than bark thickness, may better explain this trend, as it is an important determinant of an insect's ability to colonize host trees (Ferrenberg and Mitton 2014). *Matsucoccus macrocicatrices* typically embed themselves in bark crevices, along canker edges, and under lichen (Mech et al. 2013, Michigan DNR 2015, Schulz et al. 2018b), but the smooth, young bark on new growth generally lacks these ideal settlement sites. We found that the presence of scale

insects and canker formations were positively correlated with percent lichen cover on branch segments. No such pattern was found on bole samples, however, because percent lichen cover covaried with bark thickness. *Matsucoccus macrocicatrices* settlement and *C. pinea* infection appears to occur most often on intermediate bark surfaces: old enough to provide appropriate refugia, but thin enough to provide access to nutritive fluids in the outer phloem (Wardhaugh et al. 2014).

We assessed the current distribution of the *M. macrocicatrices* and Caliciopsis canker on individual trees. In previous years, scale insects may have settled on thinner-barked parts of trees that are now too thick for recolonization. If *C. pinea* infected and produced cankers on these areas following prior scale infestation, they may have gone undetected in our study, given that we only examined surface cankers. Internal necrosis of the xylem by *C. pinea*, resulting from the bark callusing over surface cankers, is also common, especially in stems of larger-diameter trees (Costanza et al. 2018). Although internal cankers decrease wood quality of sawtimber trees (Costanza et al. 2017), surface cankers likely pose the greater threat of girdling limbs and small stems. The vast majority of the cankers we observed appeared to have been caused by *C. pinea*. By identifying the characteristic fruiting structures of *C. pinea*, we were confidently able to distinguish between Caliciopsis and non-Caliciopsis surface cankers. From our isolation work of fungi present in cankers without *C. pinea* fruiting structures, the incidence of *Caliciopsis* spp. was very low (5.6%). The other 94.4% of fungi isolated from these cankers are likely endophytic opportunists, which are common on trees but are not considered highly virulent (Damm et al. 2010, Sieber 1989).

In summary, this is the first quantitative study to elucidate the distribution of *M. macrocicatrices* and Caliciopsis canker within individual trees and across tree size classes. The organisms in this insect-pathogen complex seem to favor thin-barked, but mature, *P. strobus* surfaces, including poletimber branches, sapling boles, and the lower canopy branches across all examined size classes. The incidence of *M. macrocicatrices* and Caliciopsis canker is consistent with, and we suggest is strongly associated with, observed dieback patterns in *P. strobus*, which include: (1) highest severity of branch dieback occurring in poletimber-sized trees, (2) highest mortality occurring in saplings via stem girdling, and (3) a pattern of crown thinning occurring from the bottom up in all sized trees (Asaro et al. 2018, Costanza et al. 2018, Schulz et al. 2018a, b). Although this research was conducted exclusively at the southernmost extent of the host-tree's range in Georgia, the dieback symptoms in *P. strobus* do increase in severity with increasing latitude (Schulz et al. 2018a). Thus, it is reasonable to suspect that the insect-pathogen complex would have a similar tree-level distribution elsewhere in their respective ranges.

There are a few management implications from our study. To reduce pest populations, managers may consider the use of registered insecticidal sprays on lower branches in the spring to coincide with adult emergence. Pruning of lower branches in the fall season, once *M. macrocicatrices* cysts have settled on trees, may have similar effects. This practice is suggested for white pine blister rust (Schwandt et al. 2013) and may also have the potential to slow the rate of infection by *C. pinea*. Tree density has been found to be an important factor in dieback and mortality patterns as open-grown trees tend to be more resilient (Munck et al. 2015, Schulz et al. 2018a), and so selective thinning of infected trees may slow the spread of the insect-pathogen complex. Although sawtimber

dieback appears less severe than the dieback and mortality in smaller diameter trees, especially in the Southern Appalachians, they should continue to be monitored for symptom advancement. Mortality in younger trees is a significant and immediate concern, but if the loss of reproductive trees occurs, this dieback phenomenon will invariably result in the loss of seed-source and regeneration, threatening its sustainability (Asaro et al. 2018). Overall, a greater focus on the management and conservation efforts for *P. strobus* may result in alleviating the ecological and economic impacts of this native insect-pathogen complex in eastern North America.

2.5 Acknowledgements

We are especially grateful to Arya Aghdasi and Jimmy Dale Porterfield (University of Georgia) for field and laboratory assistance. Dr. Tyler Dreaden (USDA Forest Service, Southern Research Station, Lexington, KY) conducted fungal DNA analyses. We are also grateful to the anonymous reviewer who provided valuable comments on a previous version of this manuscript. This research was funded by the USDA Forest Service, Forest Health Protection, Evaluation Monitoring Grant, the USDA Forest Service, Southern Research Station (13-CA-11330129056 and 16-CS-11330129045), and D. B. Warnell School of Forestry and Natural Resources, University of Georgia, Athens.

2.6 References

Abrams, M.D., 2001. Eastern white pine versatility in the presettlement forest. Bioscience 51, 967-979. DOI: 10.1641/00063568(2001)051[0967:EWPVIT]2.0.CO;2

Asaro, C., 2011. Forest Health Review. In, Virginia Department of Forestry, Charlottesville, Virginia, USA.

Asaro, C., Chamberlin L.A., Rose, J.A., Mooneyham, K., Rose, A.K. 2018. Mortality of understory white pine (*Pinus strobus* L.) in association with a novel scale insect-pathogen complex in VA and WV. Forest Ecology and Management.

Bean, J.L., Godwin, P.A., 1955. Description and bionomics of a new red pine scale, *Matsucoccus resinosae*. Forest Science 1, 164-176.

Blodgett, J.T., Bonello, P., Stanosz, G.R., 2003. An effective medium for isolating *Sphaeropsis sapinea* from asymptomatic pines. Forest Pathology 33, 395-404. DOI: 10.1046/j.1437-4781.2003.00342.x

Blodgett, J.T., Stanosz, G.R., 1997. *Sphaeropsis sapinea* morphotypes differ in aggressiveness, but both infect nonwounded red or jack pines. Plant Disease 81, 143-147. DOI: 10.1094/PDIS.1997.81.2.143

Costanza, K.K.L., Whitney, T.D., McIntire, C.D., Livingston, W.H., Gandhi, K.J.K. 2018. Health, ecology, and management of eastern white pine (*Pinus strobus* L.) in eastern North America. Forest Ecology and Management.

Costanza, K.K.L. 2017. Biological, ecological, and economic impact of *Caliciopsis pinea* on *Pinus strobus* (eastern white pine) in the northeastern United States. PhD Dissertation, University of Maine, Orono, Maine, USA. 160 p.

Cram, M.M., Asaro, C., Edgerton, T., 2009. Field inoculations of *Caliciopsis pinea* and *Diplodia scrobiculata* on eastern white pine. In, Southwide Forest Disease Workshop. Department of Agriculture Forest Service, Forest Health Protection, Ashville, North Carolina, Greenville, South Carolina, USA. p. 13.

Damm, U., Fourie, P.H., Crous, P.W., 2010. *Coniochaeta* (Lecythophora), *Collophora* gen. nov and *Phaeomoniella* species associated with wood necroses of *Prunus* trees. Persoonia 24, 60-80. DOI: 10.3767/003158510X500705

Ferrenberg, S., Mitton, J.B., 2014. Smooth bark surfaces can defend trees against insect attack: resurrecting a 'slippery' hypothesis. Functional Ecology 28, 837-845.

Funk, A., 1963. Studies in the genus *Caliciopsis*. Canadian Journal of Botany 41, 503-543. DOI: 10.1139/b63-044

Gardes, M., Bruns, T.D. 1993. ITS primers with enhanced specificity for basiomycetes-application to the identification of mycorrhizae and rusts. Molecular Ecology 2, 113-118. DOI: 10.1111/j.1365-294X.1993.tb00005.x

Gazis, R., Rehner, S., Chaverri, P., 2011. Species delimitation in fungal endophyte diversity studies and its implications in ecological and biogeographic inferences. Molecular Ecology 20, 3001-3013. DOI: 10.1111/j.1365-294X.2011.05110.x

Gilbert, G.S., Hubbell, S.P., Foster, R.B., 1994. Density and distance-to-adult effects of a canker disease of trees in a moist tropical forest. Oecologia 98, 100-108. DOI: 10.1007/BF00326095

Gullan, P.J., Kosztarab, M., 1997. Adaptations in scale insects. Annual Review of Entomology 42, 23-50. DOI: 10.1146/annurev.ento.42.1.23

Hothorn, T., Zeileis A., Farebrother, R.W. Cummins, C., Giovanni, M., Mitchell, D., 2017. Package 'lmtest'. https://cran.r-project.org/package=lmtest.

Houston, D.R., 1994. Major new tree disease epidemics: beech bark disease. Annual Review of Phytopathology 32, 75-87. DOI: 10.1146/annurev.py.32.090194.000451

Jackman, S., Zeileis, C.M., Fearon, J., Jackman, M.S., 2017. Package 'pscl'. https://cran.r-project.org/package=pscl.

Jactel, H., Perthuisot, N., Menassieu, P., Raise, G., Burban, C., 1996. A sampling design for within-tree larval populations of the maritime pine scale, *Matsucoccus feytaudi* Duc. (Homoptera: Margarodidae), and the relationship between larval population estimates and male catch in pheromone traps. Canadian Entomologist 128, 1143-1156. DOI: 10.4039/Ent1281143-6

Lawton, J.H., 1983. Plant architecture and the diversity of phytophagous insects. Annual Review of Entomology 28, 23-39. DOI: 10.1146/annurev.en.28.010183.000323

Little, E.L., 1971. Atlas of United States Trees, Volume 1, Conifers and Important Hardwoods. USDA Miscellaneous Publication 1146, 313.

Lombard, K., 2003. Caliciopsis canker (pine canker) *Caliciopsis pinea*. University of New Hampshire Cooperative Extension Publication, Durham, New Hampshire, USA.

Maguire, D.Y., Robert, K., Brochu, K., Larrivee, M., Buddle, C.M., Wheeler, T.A., 2014. Vertical stratification of beetles (Coleoptera) and flies (Diptera) in temperate forest canopies. Environmental Entomology 43, 9-17. DOI: 10.1603/EN13056

McClure, M.S., 1976. Colonization and establishment of red pine scale, *Matsucoccus resinosae* (Homoptera: Margarodidae) in a Connecticut plantation. Environmental Entomology 5, 943-947. DOI: 10.1093/ee/5.5.943

Mech, A.M., Asaro, C., Cram, M.M., Coyle, D.R., Gullan, P.J., Cook, L.G., Gandhi, K.J.K., 2013. *Matsucoccus macrocicatrices* (Hemiptera: Matsucoccidae): First report, distribution, and association with symptomatic eastern white pine in the southeastern United States. Journal of Economic Entomology 106, 2391-2398. DOI: 10.1603/EC13251

Michigan Department of Natural Resources. 2015. 2015 Forest Health Highlights. https://fhm.fs.fed.us/fhh/fhh_15/MI_FHH_2015.pdf. Accessed October 2, 2017.

Munck, I.A., Livingston, W., Lombard, K., Luther, T., Ostrofsky, W.D., Weimer, J., Wyka, S., Broders, K., 2015. Extent and severity of Caliciopsis canker in New England, USA: an emerging disease of eastern white pine (*Pinus strobus* L.). Forests 6, 4360-4373. DOI: 10.3390/f6114360

Munck, I.A., Luther, T., Wyka, S., Keirstead, D., McCracken, K., Ostrofsky, W., Searles, W., Lombard, K., Weimer, J., Allen, B., 2016. Soil and Stocking Effects on Caliciopsis Canker of *Pinus strobus* L. Forests 7, 269-278. DOI: 10.3390/f7110269

Nelson, P.E., Tousson, T.A., O., M.W.F., 1983. Fusarium species: an illustrated manual for identification. Pennsylvania State University Press, University Park, PA.

Overholts, L.O., 1930. Mycological notes for 1928-1929. Mycologia 22, 232-246. DOI: 10.2307/3753745

Plewa, R., Jaworski, T., Hilszczanski, J., Horak, J., 2017. Investigating the biodiversity of the forest strata: The importance of vertical stratification to the activity and development of saproxylic beetles in managed temperate deciduous forests. Forest Ecology and Management 402, 186-193. DOI: 10.1016/j.foreco.2017.07.052

Ray, W.W., 1936. Pathogenicity and cultural experiments with *Caliciopsis pinea*. Mycologia 28, 201-208. DOI: 10.2307/3754267

Richards, W.R., 1960. A new species of *Matsucoccus* Cockerell (Homoptera: Coccoidea). Canadian Entomologist 92, 179-181. DOI: 10.4039/Ent92179-3

Rowe, W.J., Potter, D.A., 1996. Vertical stratification of feeding by Japanese beetles within linden tree canopies: Selective foraging or height per se? Oecologia 108, 459-466. DOI: 10.1007/BF00333722

Schulz, A.N., Mech A.M., Asaro C.A., Coyle D.R., Cram M.M., Lucardi R.D., Gandhi K.J.K., 2018a. Temperature and biotic factors contribute to a novel eastern white pine (*Pinus strobus* L.) dieback phenomenon in the Southern Appalachian Mountains. Forest Ecology and Management.

Schulz, A.N., Mech A.M., Cram M.M., Coyle D.R., Harrington, T.C., Lucardi R.D., Lucas, S., Gandhi K.J.K., 2018b. *Caliciopsis pinea* Peck and *Matsucoccus macrocicatrices* Richards (Hemiptera: Matsucoccidae): An update on associations

with symptomatic eastern white pine (*Pinus strobus* L.) in the eastern United States. Forest Ecology and Management.

Schwandt, J.H., Kearns, H., Byler, J., 2013. White pine blister rust general ecology and management. Forest Health and Protection USDA Forest Service 14.

Sieber, T.N., 1989. Endophytic fungi in twigs of healthy and diseased Norway spruce and white fir. Mycological Research 92, 322-326. DOI: 10.1016/S0953-7562(89)80073-5

Smithson, M., Verkuilen, J., 2006. A better lemon squeezer? Maximum-likelihood regression with beta-distributed dependent variables. Psychological Methods 11, 54-71. DOI: 10.1037/1082-989X.11.1.54

Steiner, G.W., Watson, R.D., 1965. Use of surfactants in soil dilution and plate count method. Phytopathology 55, 728-730.

Ulyshen, M.D., 2011. Arthropod vertical stratification in temperate deciduous forests: Implications for conservation-oriented management. Forest Ecology and Management 261, 1479-1489. DOI: 10.1016/j.foreco.2011.01.033

United States Department of Agriculture., 2006. Georgia Soil Survey 130B – Southern Blue Ridge. https://www.nrcs.usda.gov/wps/portal/nrcs/detail/ga/soils/?cid=nrcs144p2_021881. Accessed September 1, 2017.

Wardhaugh, C.W., 2014. The spatial and temporal distributions of arthropods in forest canopies: uniting disparate patterns with hypotheses for specialisation. Biological Reviews 89, 1021-1041. DOI: 10.1111/brv.12094

Wardhaugh, C.W., Blakely, T.J., Greig, H., Morris, P.D., Barnden, A., Rickard, S., Atkinson, B., Fagan, L.L., Ewers, R.M., Didham, R.K., 2006. Vertical stratification in the spatial distribution of the beech scale insect (*Ultracoelostoma assimile*) in *Nothofagus* tree canopies in New Zealand. Ecological Entomology 31, 185-195. DOI: 10.1673/2006_06_29.1

Watson, W.Y., Underwood, G.R., Reid, J., 1960. Notes on *Matsucoccus macrocicatrices* Richards (Homoptera: Margarodidae) and its association with *Septobasidium pinicola* Snell in eastern Canada. Canadian Entomologist 92, 662-667. DOI: 10.4039/Ent92662-9

Weiss, M., Prochazka, J., Schlaghamersky, J., Cizek, L., 2016. Fine-scale vertical stratification and guild composition of saproxylic beetles in lowland and montane forests: Similar patterns despite low faunal overlap. Plos One 11. DOI: 10.1371/journal.pone.0149506

Wendel, G.W., Smith, H.C., 1990. *Pinus strobus* L. eastern white pine. In: Burns, R.M., Honkala, B.H. (Eds.), Silvics of North America. USDA Forest Service, Washington, DC, USA 476-488.

White, T.J., Bruns, T.D., Lee, S.J., Taylor, J.L. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. PCR protocols: a guide to methods and applications, 18, 315-322.

Zeileis, A., Cribari-Neto, B., Gruen, B., Kosmidis, I., Simas, A.B., Rocha, A.V., 2016. Package 'betareg'. http://betareg.r-forge.r-project.org/

Table 2.1. Study site information. All sites were located within the Chattahoochee National Forest in northern Georgia, USA.

Site	Latitude	Longitude	Elevation (m)	Ranger District
Boggs Creek	34.70083	-83.88603	560	Blue Ridge
Canada Creek	34.67940	-84.04239	860	Blue Ridge
Panther Creek	34.69949	-83.41986	460	Chattooga
Princess Pine	34.66688	-83.78466	540	Chattooga
Raper Creek	34.74417	-83.57211	530	Chattooga

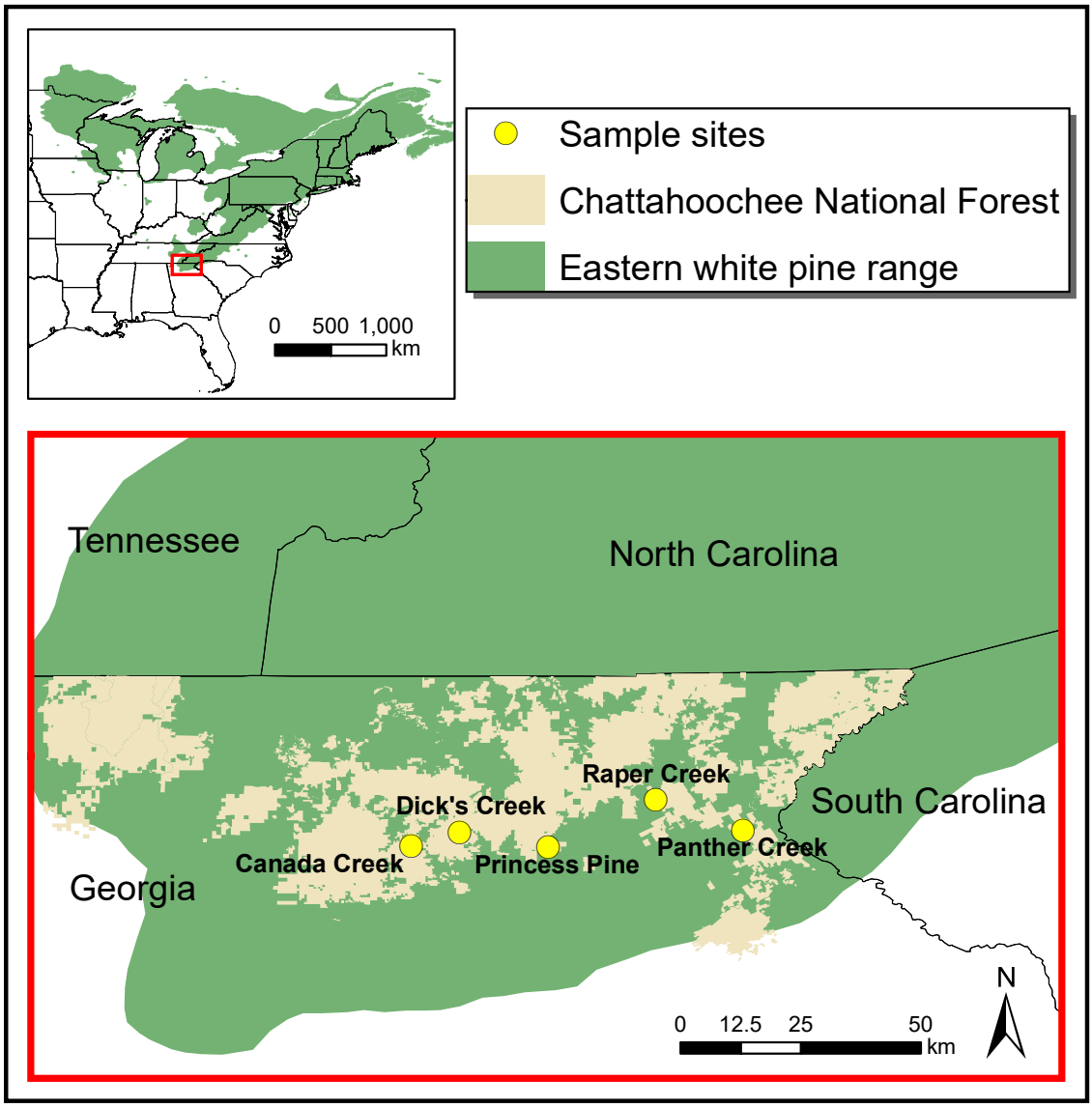

Figure 2.1. Range map (Little 1971) of *P. strobus* in North America with focal inset displaying geographic locations of the five study sites within the Chattahoochee National Forest, Georgia, USA.

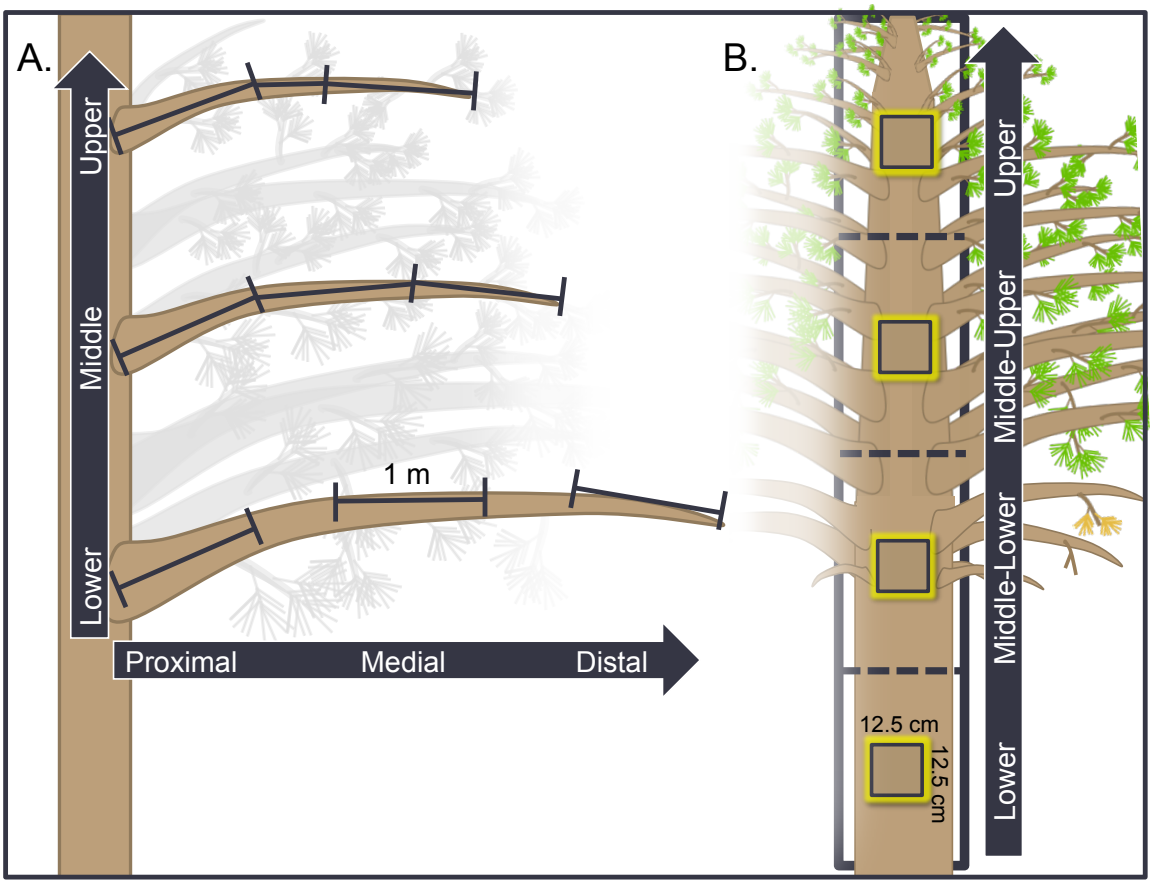

Figure 2.2. Sampling strategy on each tree (N = 45) for scale insect counts and Caliciopsis canker measurements. (A) Six live branches were analyzed per tree, two from each vertical third of the canopy (lower, middle, and upper). Each branch was partitioned into horizontal thirds and a 1-m section was removed from each (proximal, medial, and distal) (N = 604). (B) Four 12.5 × 12.5 cm area bole samples were sampled per tree (N = 180), one from each vertical fourth of the total bole height (lower, lower-middle, upper-middle, and upper).

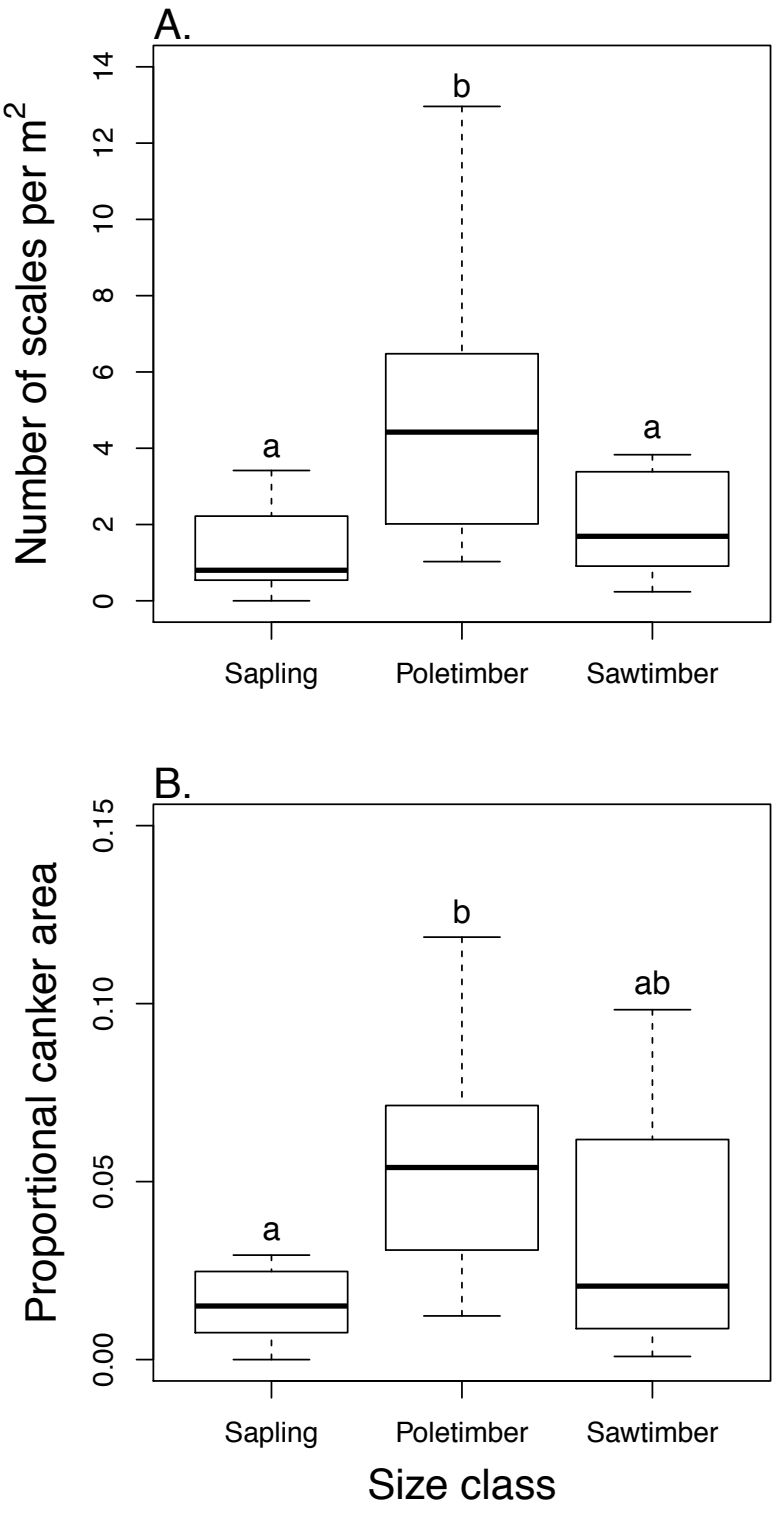

Figure 2.3. Number of (A) *M. macrocicatrices* per m^2 and (B) proportional Caliciopsis canker area found on branches analyzed according to size class: saplings (DBH <12.5 cm), poletimber (DBH = 12.5-30 cm), and sawtimber (DBH >30 cm) trees.

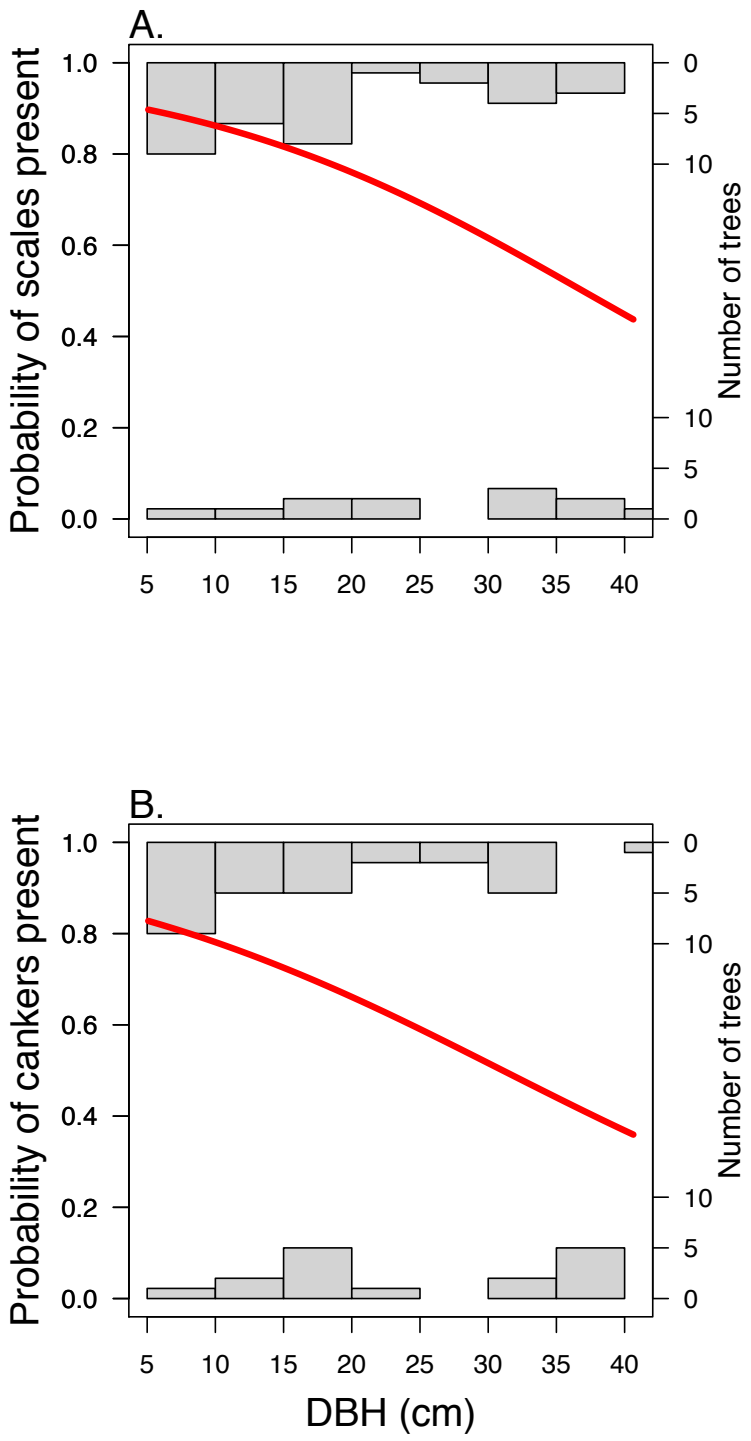

Figure 2.4. Probability of (A) *M. macrocicatrices* and (B) Caliciopsis canker presence on the boles of trees (line, left y-axis) according to tree DBH. Superimposed histogram shows the frequency of trees with and without each organism present on the bole according to DBH (bars, right y-axis).

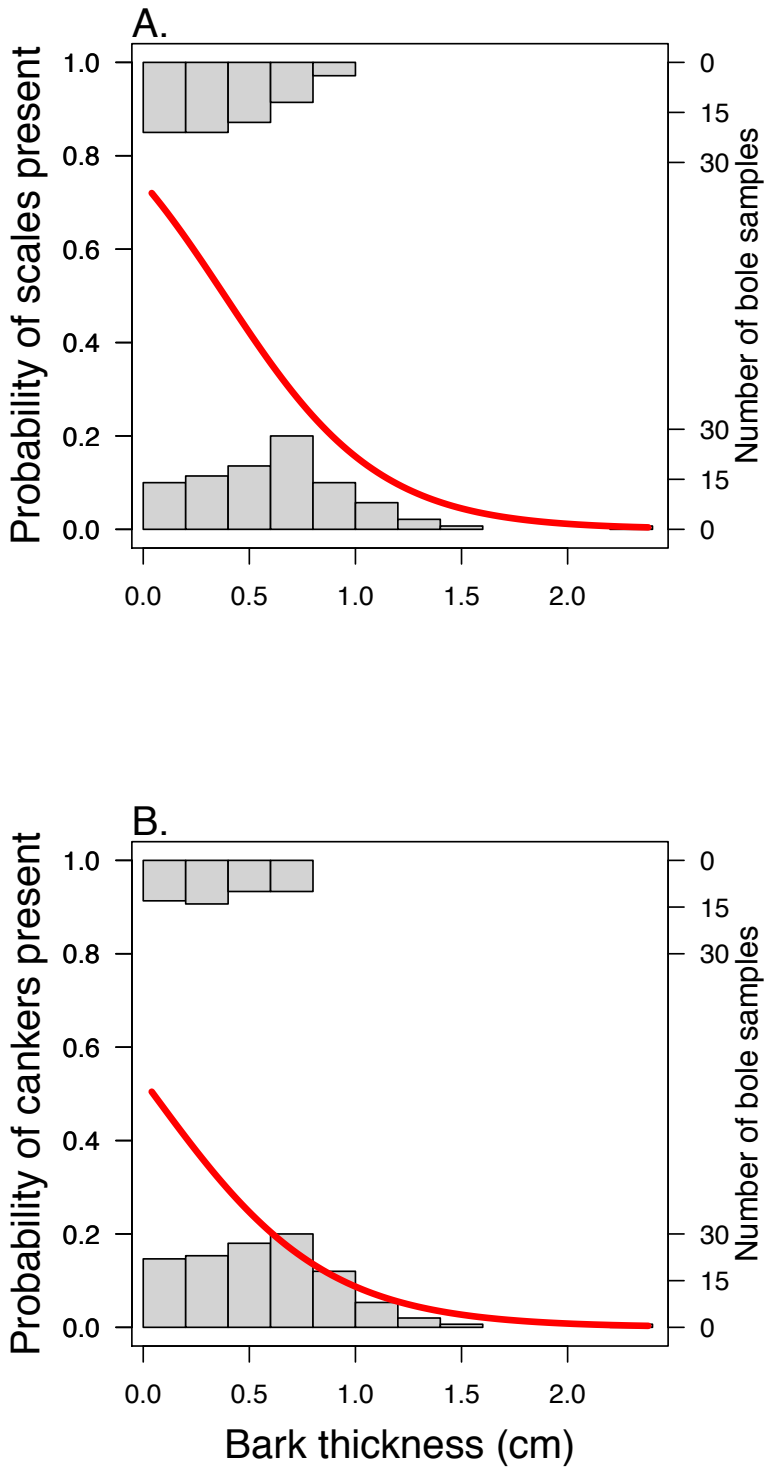

Figure 2.5. Probability of (A) *M. macrocicatrices* (B) Caliciopsis canker present on bole samples (line, left y-axis) according to bark thickness. Superimposed histogram shows the frequency of bole samples with and without each organism present according to bark thickness (bars, right y-axis).

CHAPTER 3

NATIVE OR NON-NATIVE? HISTORICAL BIOGEOGRAPHY OF AN EMERGENT FOREST PEST, *MATSUCOCCUS MACROCICATRICES*[3]

[3] Whitney T.D., Gandhi K.J.K., Lucardi R.D., 2018. In preparation for *Journal of Biogeography.*

ABSTRACT

Aim: A historically benign insect herbivore, *Matsucoccus macrocicatrices,* has recently been linked to dieback and mortality of eastern white pine (*Pinus strobus* L.). Previous reports indicated its native range was restricted to New England, USA and southeastern Canada. Now, the insect occurs throughout an area extending from the putative native range, southward to Georgia, and westward to Wisconsin. Our goal was to evaluate whether its current distribution was due to recent introductions consistent with invasion processes. We considered two hypotheses: (1) if recent expansion into adventive regions occurred, those populations would have reduced genetic diversity due to founder effect(s); alternatively (2) if *M. macrocicatrices* is native and historically co-occurred with its host tree throughout the North American range, then populations would have greater overall genetic diversity and population structure indicative of past biogeographic influences.

Location: Eastern North America

Methods: We developed nine *M. macrocicatrices*-specific microsatellite markers *de novo* and genotyped 390 individuals from 22 populations sampled across the range of *P. strobus* in the USA. We assessed genetic variability, relatedness, and population structure.

Results: Overall genetic diversity metrics were higher than expected for this species. No signatures of genetic bottlenecking were observed. The number of rare alleles and observed heterozygosity had an inverse relationship with latitude. Analyses of population structure indicated three distinct genetic clusters separated by the Great Lakes and the Blue Ridge Mountains.

Main Conclusions: The seemingly sudden ecological shift from benign herbivore to significant pest led us to suspect *M. macrocicatrices* was non-native. However, our findings suggest this insect is native and has likely co-occurred with its host tree since the last glacial maximum. Our study demonstrates the importance of historical biogeographical reconstruction to inform how to approach an emergent pest.

Keywords: Blue Ridge Mountains, eastern white pine bast scale, Great Lakes, forest health, microsatellites, native pest, *Pinus strobus*, population genetics, Southern Appalachians

3.1 Introduction

Non-native species lack the long evolutionary history that native species have within a local community, and hence communities can suffer greater damage from non-natives due to naiveté (Paolucci et al. 2013, Salo et al. 2007, Simberloff et al. 2012). However, endemism does not preclude a species from developing pestiferous behaviors. Although rarer, native species can become pests within their native ranges similar to non-native species through expansion into adventive ranges (Dodds et al. 2018, Hassan and Ricciardi 2014, Simberloff et al. 2012).

Unifying all organisms causing serious ecological and economic damage is the release from evolutionary constraints and/or the exploitation of new niche opportunities (Carey et al. 2012). For instance, the absence of co-evolved natural enemies (Keane and Crawley 2002) or host/prey defenses (Gandhi and Herms 2010, Paolucci et al. 2013) can allow non-natives to establish and thrive in novel environments, but for a native species, these constraints on their populations generally remain intact (Tong et al. 2018). Instead, the reasons certain native species elevate to pest status are often multi-faceted, sometimes involving positive population responses to climate change (Nackley et al. 2017), anthropogenic habitat alterations (Carrete et al. 2010), and/or host-shifts following other non-native introductions (Lefort et al. 2014). Reconstructing the historical origin of an emergent pest species can provide an evolutionary context to its contemporary interactions (Richardson and Ricciardi 2013, Sakai et al. 2001), an important first step in control and conservation efforts.

In this study, we evaluated the population genetic variability and distribution of the eastern white pine bast scale, *Matsucoccus macrocicatrices* Richards (Matsucoccidae:

Hemiptera), a small sap-sucking insect currently associated with the novel dieback phenomenon of eastern white pine (*Pinus strobus* L.) in North America (Costanza et al. 2018, Mech et al. 2013) (Figure 3.1a). This insect creates deep feeding wounds during its second-instar cyst stage, which may facilitate subcortical infection of trees by pathogens, primarily the native *Caliciopsis pinea* Peck (Schulz et al. 2018b). This fungus contributes to the formation of cankers on the bark (Figure 3.1b), causing hallmark symptoms, including the girdling of stems in young trees and the bottom-up branch dieback in older trees (Figure 3.1c) (Asaro et al. 2018, Costanza et al. 2018).

Prior to 2011, *M. macrocicatrices* was considered a benign herbivore with a limited distribution. The only recorded specimens were collected in the northeastern USA (Massachusetts, New Hampshire, and Vermont) and southeastern Canada (New Brunswick, Nova Scotia, Ontario, and Quebec) (Mech et al. 2013, Richards 1960, Watson et al. 1960). However, it is now found throughout the North American range of *P. strobus* and linked to host-tree damage and mortality (Mech et al. 2013, Schulz et al. 2018a). Other *Matsucoccus* spp. have become pests outside their native ranges, such as the Japanese pine bast scale (*M. matsumurae* Kuwana), maritime pine bast scale (*M. feytaudi* Ducasse), and Israeli pine bast scale (*M. josephi* Bodenheimer et Harpaz) (Bean and Godwin 1971, Kerdelhue et al. 2014, Mendel 1998). In these cases, release from natural enemies and/or host defenses were attributed as causes for invasion (Jactel et al. 2006, Mendel 1998). Whether *M. macrocicatrices* has similarly expanded its range to enemy-free areas with naïve host provenances or has become pestiferous within its native range due to abiotic or biotic shifts, remains unknown.

Microsatellites are frequently used in population genetic studies to identify the origin of pest arthropods (e.g. Havill et al. 2016, Khamis et al. 2009, Virgilio et al. 2010, Zemanova et al. 2016, Zhang et al. 2012). As *M. macrocicatrices* is now well-established south of Massachusetts and west of Lake Erie, where no records existed prior to 2011 and 2015, respectively (Mech et al. 2013, Michigan DNR 2015), we developed microsatellites *de novo* to learn if this insect species was new to regions outside its putative native range. We tested two competing hypotheses: (A) if populations of *M. macrocicatrices* established outside of its purported native range (New England) are the result of recent introduction(s) and colonization, then we expected to observe reductions in genetic diversity consistent with founder events. Alternatively, (B) if *M. macrocicatrices* has historically co-occurred with its host outside its purported native range, then we expected to observe similar levels of genetic diversity. This hypothesis assumes that, like its *P. strobus* host, the insect existed in Southern Appalachian refugial populations during the last glacial maximum, recolonized northward as glaciers receded, and re-accumulated genetic diversity over thousands of years (Nadeau et al., 2015). Further, we also expected to observe prominent population structure where geographic barriers, such as the Great Lakes, may have limited *M. macrocicatrices* gene flow over time.

3.2 Methods

3.2.1 Sample collection

We sampled 22 sites throughout the range of *P. strobus* for *M. macrocicatrices* in the USA (Figure 3.2, Table 3.1). Immature cysts were collected between 2015-2018

during the winters and springs when the insects are near the end of their second-instar, relatively large (0.5-1.0 mm), and easiest to locate when sessile and embedded in tree bark (Figure 3.1a). Sampling occurred in one of two ways: (1) for 14 of the 22 sites, branches and stems of *P. strobus* trees [1-12 cm diameter at breast height (DBH)] were shipped overnight to the University of Georgia (Athens, Georgia, USA) (Table 3.1). Individual scale insects were then located with a stereo microscope and immediately preserved in 95% ethanol at -20 °C. (2) For the remaining 8 of 22 sites, individual cysts were located and removed *in situ* from *P. strobus* bark, preserved in 95% ethanol immediately, and stored at -20 °C within three days of collection (Table 3.1). Sites were separated by ≥25 km. We sampled 11-20 individual *M. macrocicatrices* from between one and nine trees per site (referred to as population, hereafter). For one Michigan population (n = 8) and the Wisconsin population (n = 6), sampling was conducted in June 2018, narrowly after most of the insects had already molted. We instead collected the voided cuticles (i.e. exoskeletons) in lieu of live cysts. Cuticles that produced adequate genomic DNA purity and yield were used for microsatellite analyses.

3.2.2 Molecular Analysis

3.2.2.1 DNA Extraction

All DNA extractions of individual *M. macrocicatrices* cysts and cuticles were performed with the Qiagen DNEasy® Blood and Tissue Extraction Kit (Qiagen Inc., Chatsworth, California, USA) following the manufacturer's protocol with two minor modifications: (1) we pierced each *M. macrocicatrices* cyst with a flame-sterilized insect pin and proceeded with overnight lysis, which allowed us to retain the cuticles for

vouchering and still achieve adequate genomic DNA yield; and (2), we decreased final DNA volumes for each sample to 100 µL total (two elution steps of 50 µL).

3.2.2.2 Species identification

The only species within the genus *Matsucoccus* to have been reported on *P. strobus* in North America is *M. macrocicatrices*. However, there are congeneric species, both native (e.g. *M. gallicolus*) and exotic (e.g. *M. matsumurae*), that occur on other pines sympatrically with *M. macrocicatrices*, and thus we wished to confirm our identifications. Adult specimens are required for morphological identification to species; instead we utilized molecular barcoding to confirm that our samples, all juvenile cysts, were *M. macrocicatrices.* We conducted Sanger sequencing on the polymerase chain reaction (PCR) products of one individual per sampled site (N = 22), with the DNA barcodes of the D2-D3 domains of the nuclear rDNA 28S gene with the following primers: S3360 (Dowton and Austin 1998) and A335 (Whiting et al. 1997). This gene region has demonstrated successful identification of *Matsucoccus* specimens to species (Booth and Gullan 2006, Mech et al. 2013).

Reaction conditions for sequencing this barcode region were: 25 µL reactions consisting of 2.5 µL 10x Takara buffer (Takara Bio Inc., Shiga, Japan), 0.16 mM dNTPs, 0.4 µM per primer, 0.625 U Takara *Ex Taq*™, and 1-10 ng of template gDNA. Reactions were conducted in Mastercycler Pro S Thermal Cyclers (Eppendorf, Hamburg, Germany) under the following amplification: 94 °C for 4 min; 35 cycles of 94 °C for 30 s, 55 °C for 30 s, and 72 °C for 60 s; and a final extension of 72 °C for 3 min. Consensus sequences were assembled in GENEIOUS version 10.2.3 (Biomatters, Auckland, New Zealand). We

then used BLAST searches (NCBI) to reference our returned sequences to those accessions (KF040554 – KF040572) submitted to GenBank by Mech et al. (2013); presently, these are the only existing DNA barcodes of *M. macrocicatrice*s.

3.2.2.3 Microsatellite development and selection

Novel microsatellite markers were developed using shotgun sequence reads generated with Illumina MiSeq sequencing (Illumina, San Diego, California, USA). Genomic DNA was extracted from a single individual *M. macrocicatrices* cyst collected from Georgia (population GA3; Table 3.1) using an Omega Mag-Bind blood and tissue HDQ 96 Kit (Omega Bio-Tek, Norcross, Georgia, USA). Libraries were prepared with the Kapa Hyper Prep Kit KR0961-v1.14 (Kapa Biosystems, Wilmington, Massachusetts, USA) following the manufacturer's protocol with the following modifications: (1) half-reaction volumes were used, (2) universal iTru y-yoke adapters were ligated onto the genomic DNA, and (3) iTru5 and iTru7 primers were used to produce Illumina TruSeqHT compatible libraries (Glenn et al. 2016). Products were purified with Sera-Mag Speedbeads (Thermo-Scientific, Waltham, MA, USA), resulting in a total of 3,998,7582 reads paired by name in GENEIOUS 10.0.9 (Biomatters) and archived in the GenBank Sequence Read Archive (SRP159008). Illumina TruSeq adapters and bases with an error probability above 0.04 were trimmed, and *de novo* assembly was performed on sequences greater than 150 bp. Unassembled reads (mean of 256 bp) totaled 2,778,623 and were exported to MSATCOMMANDER 1.0.8*beta* (Faircloth 2008), Where the program searched for microsatellite loci with perfect repeats of 3 to 6 nucleotides. Dinucleotide repeats were avoided due to the propensity to stutter (Flores-Renteria and

Krohn 2013). We used default minimum lengths and melting temperatures. As well as combining loci 50 bp apart, MSATCOMMANDER identified 2566 candidate loci, of which 530 met our criteria for primer design. We selected 96 primer pairs, all with a variety of repeat lengths and nucleotide motifs, and each with a pair penalty assigned by PRIMER3 (Rozen and Skaletsky 2000) of less than five. For each locus, a M13 (GGAAACAGCTATGACCAT) or CAG (CAGTCGGGCGTCATCA) sequence tag was added to the 5' end of one primer and a GTTT pig-tail was tagged to the corresponding primer (Brownstein et al. 1996).

From the suite of 96 primers tested on one individual *M. macrocicatrices* (from population GA4) and another from Michigan (44.59429, -84.1246), 36 primer pairs successfully amplified and underwent further evaluation for consistency and polymorphism using extracted DNA from 20 individuals collected from a single population (GA3). Of the 36 primer pairs, 12 microsatellite loci met the assumptions of Hardy-Weinberg equilibrium and had a null allele rate below 0.1, with consistent amplification (Table 3.2). The 12 loci were then tested on all 390 individuals from the 22 populations of the study.

3.2.2.3 Microsatellite amplification

Amplification procedures were uniform for each of the 12 loci on all 390 total individuals. Polymerase chain reactions (PCRs) were in 10-µL volumes consisting of 1 µL 10x Takara buffer (Takara Bio Inc.), 0.125 mM dNTPs, 0.05 µM forward M13- or CAG-tagged locus-specific primer, 0.50 µM reverse GTTT pig-tailed locus-specific primer (Integrated DNA Technologies, Coralville, Iowa, USA), 0.45 µM fluorescently

labeled M13- or CAG-tagged primer, 0.50 U Takara *Ex Taq*™, and 1-10 ng template DNA. The fluorophore labels for each primer included FAM (Integrated DNA Technologies), PET, VIC, or NED (Applied Biosystems, Foster City, California, USA). All reactions were conducted in Mastercycler Pro S Thermal Cyclers (Eppendorf, Hamburg, Germany), under the following PCR touchdown protocol: 95 °C for 2 min; 20 cycles of 95 °C for 60s, 67-57 °C for 30 s decreasing 0.5 °C every cycle, and 72 °C for 45 s; 20 cycles of 95 °C for 60s, 57 °C for 30 s, and 72 °C for 45 s; and a final extension of 72 °C for 5 min. Amplicon sizes were determined on a 3730 capillary sequencer (Applied Biosystems) at the Arizona State University DNA Core Lab using GeneScan LIZ 500 size standard (Applied Biosystems). Allele sizes were scored using the microsatellite plugin for GENEIOUS version 10.2.3 (Biomatters).

3.2.3 Statistical analysis

3.2.3.1 Genetic diversity

All microsatellite loci for each population were tested for the following in GENEPOP version 4.2 (Raymond and Rousset 1995): linkage disequilibrium with the probability test, deviations from Hardy-Weinberg equilibrium with exact tests, and null allele frequency with the Brookfield (1996) method. *Matsucoccus macrocicatrices* reproduces sexually (Costanza et al. 2018) but is purportedly capable of parthenogenesis (Foldi, 2004). So, to test the null hypothesis of a random association of alleles among different loci consistent with random mating, we calculated the index of association (I_A) (Brown et al. 1980) and the standardized index of multilocus linkage disequilibrium ($\bar{r}_d$) (Agapow and Burt 2001) for each population using the 'poppr' package (Kamvar et al.

2014) in R version 3.5.1 (R Core Team 2018). We also sought to determine if this species has a haplodiploid sex-determination system, where males develop from unfertilized eggs and are haploid. We assessed the potential for haplodiploidy by observing if any individuals were homozygous for every locus.

Genetic diversity was estimated using effective number of alleles (A_E), mean frequency of private alleles (A_P), mean number of locally common alleles ($\geq$5%) occurring in $\leq$50% of populations (A_{LC}), observed heterozygosity (H_O), and unbiased expected heterozygosity (H_E) in GENALEX version 6.503 (Peakall and Smouse 2006, 2012). Rarefied allelic richness (A_R) and inbreeding coefficients (F_{IS}) were calculated in the R package 'hierfstat' (Goudet 2005). Generalized linear models were conducted in R to evaluate the association between latitude and longitude with genetic diversity. Latitude and longitude were included in the models simultaneously as covariates. To specifically test for clinal decays in genetic diversity according to geographic distance from the putative native range, we created a Euclidean distance (km) matrix among all populations from one population in New Hampshire (NH1). We assigned a value of 0 km to the four populations located within the insect's putative native range (NH1, NH2, NH3 and ME). Regressions were then conducted on the indices as stated above. All models included the number of genotyped individuals as a covariate to control for uneven population sizes. All indices met the assumptions of normality except for A_R, F_{IS} and A_{LC}, which were log-transformed.

The program BOTTLENECK (Cornuet and Luikart 1996, Piry et al. 1999) was used to detect if signals of recent bottleneck event(s) existed within in our dataset for *M. macrocicatrices*. This program tests for deviations from mutation-drift equilibrium with

the assumption that allelic richness decreases faster than heterozygosity in shrinking populations. All 22 populations were tested separately. We also tested three pooled populations utilizing the genetic clusters informed from Bayesian clustering analysis (see below Methods and Results). All BOTTLENECK tests were permuted 1000 times. We used the single-step mutation model (SMM) and the two-phase model (TPM) with 95% single-step mutations and 5% multi-step mutations (Piry et al. 1999). Significant excesses in heterozygosity for each population were determined with the one-tailed Wilcoxon signed-rank test.

3.2.3.2 Population structure

We used the Bayesian clustering algorithm STRUCTURE version 2.3.4 (Pritchard et al. 2000) to infer subgroup assignments for *M. macrocicatrices*. For all simulations we did not use a location prior, and we assumed an admixture model with allele frequencies correlated among groups (Falush et al. 2003). Each run utilized 25,000 burn-in, followed by 50,000 Markov Chain Monte Carlo (MCMC) iterations, replicated 20 times for each number of clusters assumed (*K*).

Hierarchical groupings of individual *M. macrocicatrices* were simulated in separate STRUCTURE runs as follows: (1) the entire dataset of 22 populations, with *K* ranging from 1 to 22; (2) simulations to evaluate substructure within resulting major clusters, including (2a) the Southern Appalachians ("*SApps*"; 7 populations in Georgia, Tennessee and North Carolina) with *K* = 1 through 7, (2b) the Northeast ("*NEast*"; 11 populations in Virginia, West Virginia, Pennsylvania, New Hampshire, and Maine) with *K* = 1 through 11, and (2c) the Great Lakes ("*GLakes*"; 4 populations in Michigan and

Wisconsin) with K = 1 through 5. Optimal K, or the most likely number of clusters for each grouping, was determined by the Evanno et al. (2005) method implemented in STRUCTURE HARVESTER (Earl and Vonholdt 2012). Populations were assigned to the cluster with the highest corresponding mean posterior probability of ancestry.

We also inferred optimal population structure from analyses of molecular variance (AMOVA) using ARLEQUIN version 3.5 (Excoffier and Lischer 2010) to determine the hierarchical partitioning of genetic variance using pre-defined population structure from STRUCTURE results. We conducted six AMOVAs with 10,000 permutations to test: (A) no genetic structure, (B) genetic structure where K = 2, (C) genetic structure where K = 3, (D) *NEast* populations only, (E) *SApps* populations only, and (F) *GLakes* populations only.

We performed multiple principal coordinates analyses (PCoA) using Nei's unbiased genetic distances (Nei 1978) in GENALEX. We also calculated pairwise F_{ST} (Weir and Cockerham 1984) and Slatkin's (1995) linearized pairwise F_{ST} values in ARLEQUIN to evaluate genetic differentiation between populations. The linearized F_{ST} matrix and a pairwise matrix of log-transformed geographic distances (km) were also used in a Mantel test (Mantel 1967) to detect isolation-by-distance (IBD). Mantel tests may falsely detect IBD in instances of hierarchical structure with distinct barriers to gene flow (Meirmans 2012), so we conducted additional Mantel tests within clusters. We also performed a partial Mantel test controlling for cluster assignment with a covariate matrix containing binary values for each pairwise relationship: 0 for pairs of populations belonging to the same cluster and 1 for those belonging to separate clusters. In another partial Mantel test, we examined the association between genetic distance and cluster

assignment, using the geographic distance matrix as a covariate. All Mantel and partial Mantel tests were performed with 100,000 permutations in the R package 'vegan' (Oksanen et al. 2018).

We used BARRIER version 2.2 (Manni 2004), which implements Monmonier's maximum difference algorithm, to detect where abrupt changes in pairwise F_{ST} occur and to identify where barriers to gene flow exist on the landscape. Based on the finding of a potential barrier existing in the Blue Ridge mountains (see Results), we further evaluated the link between *P. strobus* density and the genetic connectivity of *M. macrocicatrices* by assessing the least-cost paths between populations in Georgia, North Carolina, Tennessee, Virginia, and West Virginia. We sought to test tree-host connectivity through the application of circuit theory (McRae and Beier 2007) for the purpose of comparing pairwise genetic distances but not for modeling gene flow. Remote sensing data of *P. strobus* from Forest Inventory and Analysis (FIA, USDA Forest Service) were used to create a relative density raster in R where each pixel (size = 250 m) holds a value equal to the percentage of *P. strobus* comprising the total composition of trees ≥12.7 cm DBH. We created a cost-surface raster with each pixel holding a resistance value based on its corresponding tree density value. Pixels with 0% *P. strobus* were assigned a high resistance value of 200 and all other pixels were assigned resistance values of 1-100, inversely proportional to their relative density of host trees (100-1%). We assessed the least-cost paths of the cost-surface raster between the 11 populations adjacent to the Blue Ridge geographic barrier in the R package 'gdistance' (van Etten 2017). We conducted another partial Mantel test, controlling for cluster assignment as above, to assess the correlation between linearized F_{ST} and pairwise least-cost distance.

3.3 Results

3.3.1 Species identification

We were able to form contigs from all 22 returned sequence pairs except for one from Michigan (population MI3) likely due to DNA degradation of the insect cuticle from which we used for extraction. All consensus sequences had ≥99.4% matching 28S sequences to *M. macrocicatrices*. GenBank accession numbers for our samples are MK402259-MK402280.

3.3.2 Microsatellite loci quality

All loci had null allele rates of less than 0.1 averaged across all populations except for three, which we then removed from all further analyses (Table 3.2). Of the 198 locus-population combinations, exact tests revealed a significant departure from Hardy-Weinberg equilibrium in 47 pairs, but with no clear concentration in any particular locus or population. None of the nine remaining loci showed significant linkage disequilibrium. Every individual had a unique multilocus genotype, and only two populations (GA1 and ME) deviated from the null hypothesis of random allelic association among loci according to I_A and $\bar{r}_d$ statistics (Table 3.3). No individual was homozygous across all loci, which negates the possibility *M. macrocicatrices* has haplodiploid sex-determination.

3.3.3 Genetic diversity

Estimated indices of genetic diversity for each population (A_E, A_R, A_P, A_{LC}, F_{IS}, H_O and H_E) are summarized in Table 1. Mean number of alleles per population ranged

from 2.78 to 6.57 and mean expected heterozygosity (H_E) ranged from 0.40 to 0.69. Our analyses found some evidence of latitudinal, but not longitudinal clines, with mean number of locally common alleles (A_{LC}) ($R^2 = 0.41$, df = 18, t = -2.78, P = 0.01) and observed heterozygosity (H_O) ($R^2 = 0.33$, df = 18, t = -2.57, P = 0.02) both decreasing as latitude increased. Linear models showed no correlation between distance from the putative native range (the four populations in New England) and genetic diversity, except for A_{LC}, which was observed to increase with distance from our New Hampshire reference population (NH1) ($R^2 = 0.25$, df = 19, t = 2.12, P = 0.047).

No signatures of bottleneck were detected in any of the populations we sampled according to one-tailed Wilcoxon tests (Table 3.4). Likewise, we found no support for recent bottleneck events when pooling *NEast*, *SApps*, and *GLakes* populations. Heterozygote deficiency, however, which is indicative of population expansion, was detected in 13 of the 22 populations and in all three pooled groups according to two-tailed Wilcoxon tests for at least one of the two models (SMM and TPM).

3.3.4 Population structure

3.3.4.1 Range-wide population structure

The optimal number of clusters for all *M. macrocicatrices* (N = 390) in the USA was $K = 3$ (Figure 3.3a). Although ΔK indicated optimal $K = 2$ from STRUCTURE results, the plateau in the Ln Pr(X|K) curve was strongest when $K = 3$ (Figure 3.4a). Furthermore, the clear clustering visualized in the PCoA (Figure 3.5), as well as results from AMOVAs (Table 3.5a-c), Mantel tests (Table 3.6), and BARRIER analysis, all support a three-cluster model ($K = 3$). The three clusters were regionally distinct and

were designated as follows: populations located in (1) Virginia and northward were defined as the "*NEast*" cluster, (2) those located in North Carolina and southward were defined as the "*SApps*" cluster, and (3) those from Michigan and Wisconsin were defined as the "*GLakes*" cluster (Figure 3.3a).

Analyses of molecular variance (AMOVAs) resulted in significant genetic differentiation between defined groups (Table 3.5). Range-wide tests using cluster assignment priors inferred from STRUCTURE accounted for more overall differentiation (F_{ST} = 0.325 for K = 2, 0.324 for K = 3) than the test assuming no population structure (F_{ST} = 0.253 for K = 1). The two-cluster and three-cluster AMOVA tests resulted in nearly the same F_{ST}, but genetic variation was partitioned differently. Twice as much variation was partitioned between groups (F_{CT} = 0.235) as among populations within groups (F_{SC} = 0.116) for the three-cluster model, whereas there was only 1.3 times as much genetic variation partitioned between groups (F_{CT} = 0.201) as among populations within groups (F_{SC} = 0.155) for the two-cluster model. Standard and partial mantel tests (Table 3.6) revealed that the association between overall pairwise genetic and geographic distance was significant (R = 0.509, P < 0.001) but not when controlling for clustering assignments (R = 0.046, P < 0.30). There was also an association between genetic distance and cluster assignment, controlling for genetic distance (R = 0.653, P < 0.001).

BARRIER analysis identified two main regions with abrupt changes in genetic distances, consistent with the barriers that separate the three clusters identified previously. The first barrier isolated all four Great Lakes populations (*GLakes*) from the rest of the dataset, and the second barrier was located in between the North Carolina population, NC2, and the Virginia population, VA1, in the Blue Ridge mountains. Host-

tree connectivity could not explain the presence of this second barrier, as pairwise genetic distance (linearized F_{ST}) between populations in Georgia, North Carolina, Tennessee, Virginia, and West Virginia was not associated with pairwise least-cost distance based on *P. strobus* density, according to a partial Mantel test controlling for cluster assignment (R = 0.179, P = 0.882).

3.3.4.2 Regional population structure

Results from within-cluster STRUCTURE runs are shown in Figure 3.3b. Within the *NEast* cluster, we found the optimal $K = 2$ (Figure 3.4b), where the three Virginia and single West Virginia populations comprised one subgroup and populations from Pennsylvania, New Hampshire and Maine comprised the other subgroup. Within the *SApps* cluster, we found the optimal $K = 3$ (Figure 3.4c), with the first subgroup consisting of the three populations in northeastern Georgia and western North Carolina (populations GA3, GA4 and NC1), the second subgroup consisting of the three populations in the northwestern Georgia and southeastern Tennessee (populations: GA1, GA2 and TN), and the third subgroup consisting solely of individuals from the population NC2. Within the *GLakes* cluster, we found the optimal $K = 4$ (Figure 3.4d), but every population appeared to be of a mostly mixed ancestry; only one population (WI) had a mean posterior probability of assignment greater than 50% for one of the four clusters.

The AMOVAs conducted within each cluster revealed that genetic differentiation was partitioned similarly and was overall comparable in both the *NEast* cluster (F_{ST} = 0.150, F_{CT} = 0.091, F_{SC} = 0.065) and the *SApps* cluster (F_{ST} = 0.158, F_{CT} = 0.095, F_{SC} = 0.069). The *GLakes* cluster had comparatively lower genetic differentiation (F_{ST} = 0.071).

Among pairwise populations, F_{ST} values were all statistically significant except for four: two among New Hampshire populations and two among Michigan populations (Table 3.7). The NC2 population from North Carolina was the most highly differentiated, with pairwise F_{ST} values ranging from 0.230 to 0.533. Mantel tests revealed significant isolation-by-distance within each cluster (*NEast*: R = 0.335, P = 0.005; *SApps*: R = 0.833, P < 0.001; and *GLakes*: R = 0.835, P = 0.04) (Table 3.6, Figure 3.6).

3.4 Discussion

Non-native and native species that become pestiferous often do so by escaping different evolutionary constraints, such as exploiting an enemy-free or defense-free space (Gandhi and Herms 2010, Richardson and Ricciardi 2013). Determining whether *Matsucoccus macrocicatrices* is new outside its purported native range in New England may offer perspective into its sudden association with novel dieback symptoms and mortality of its host tree. Based on evidence presented herein, we propose the insect is native throughout the North American range of *P. strobus* and that the two organisms have likely co-occurred since the LGM.

3.4.1 Evidence for nativity

The genetic landscape of *M. macrocicatrices* was not consistent with that of an exotic species recently introduced to a new range. Source populations are usually genetically rich, whereas founder populations are usually genetically depauperate (Nei et al. 1975). We found high levels of global genetic diversity (e.g., H_E = 0.43-0.68), especially when compared to the congener, *M. feytaudi,* where there are both source

populations and recent, non-native founder populations in Europe (e.g., H_E = 0.25-0.58) (Kerdelhue et al. 2014). There was no evidence of a recent bottleneck in any population, nor was there a longitudinal cline in genetic diversity despite the most easterly populations being within the purported native range (New Hampshire and Maine). We also did not find the expected decay in genetic diversity when assessing Euclidean distance from these populations. In fact, the most southerly populations tended to be the most genetically rich, as both the mean number of locally common alleles (i.e. rare alleles) and the observed heterozygosity per population were negatively associated with latitude. Glacial history may provide some context to our resulting patterns. Both palynological and molecular phylogeographic evidence indicate that refugial populations of *P. strobus* survived in the mid-Atlantic and at the southernmost portion of the Appalachian Mountain range during the last glacial maximum (Davis 1983, Nadeau et al. 2015, Zinck and Rajora 2016). If *M. macrocicatrices,* being a specialist of its host tree, co-occurred during northward recolonization following glacial thaw, then the southernmost populations would likely have retained more ancestral genetic variation (Hewitt 1999).

The patterns of genetic differentiation also failed to substantiate that *M. macrocicatrices* is non-native outside of New England. In addition to restricted genetic exchange and genetic drift, sufficient time is required for populations to differentiate, and thus, our results consistently suggest *M. macrocicatrices* is well-established within its entire, current distribution. Populations were highly structured overall and delineated into three distinct, regional and genetic clusters: *NEast, SApps, and GLakes*. STRUCTURE had high support for clustering *SApps* and *GLakes* together, but with only 44 total

individuals analyzed from four populations in the Great Lakes region, uneven sampling may have influenced the results. STRUCTURE analysis tends to merge distinct, but small, subpopulations together when sampling is biased (Puechmaille 2016). However, results from AMOVA (Table 3.5) and pairwise genetic distances (Table 3.7) strongly suggest *SApps* and *GLakes* are separate groups.

3.4.2 Barriers to gene flow

The limited dispersal ability of *M. macrocicatrices* may help to explain their overall high genetic differentiation. The main dispersal stage for this species is the first instar "crawler", capable of walking short distances and being carried by wind for longer distances (Costanza et al. 2018). Congeners have been observed to stay airborne for up to 0.5 km (Bean and Godwin 1955) and could theoretically exceed 85 km in passive flight given optimal conditions (Hanks and Denno 1998). Arthropods evolving in heterogenous landscapes with frequent patches of unsuitable habitat tend to avoid passive, aerial dispersal, because it presents a high risk of mortality (Bonte et al. 2012; Bonte et al. 2003). *Matsucoccus* spp. are no different, with only <20% of individuals observed to disperse from their natal trees (McClure 1977, Stephens and Aylor 1978; Unruh and Luck 1987). As a host tree, *P. strobus* only occasionally grows in pure stands and is more commonly found as a highly scattered supercanopy tree (Abrams 2001). We found significant isolation-by-distance within each population cluster according to Mantel tests (Table 3.6), indicating that long-distance dispersal of *M. macrocicatrices* between patches is rare (McClure 1976).

Between the three clusters, there were sudden increases in genetic distance not simply explained by geographic distance. We identified two main barriers to gene flow that were likely responsible. One barrier isolates the *GLakes* cluster, suggesting the Great Lakes act as a physical barrier to successful dispersal (Figure 5a). Large water bodies present a high risk of mortality for passive, aerial dispersed arthropods and can lead to vicariance (Hawes et al. 2007, Kuntner and Agnarsson 2011). Genetic divergence between USA populations located in the Great Lakes and the northeastern states have also been observed in active-dispersing terrestrial animals (e.g. Bagley et al. 2017, Hapeman et al. 2017).

The second barrier to *M. macrocicatrices* gene flow is located in between North Carolina (population NC2) and Virginia (population VA1) where the Blue Ridge Mountains and the Ridge and Valley ecoregions meet (U.S. Environmental Protection Agency 2013) (Figure 3.7a). Population NC2 was the most genetically isolated in our study, perhaps because it lies in the French Broad River basin in North Carolina with imposing mountains to its southwest and northeast. It is noteworthy that the least-cost paths (based on *P. strobus* density) from this population to the nearest-neighbor *SApps* populations are roughly equal to its least-cost path to the southernmost *NEast* population (VA1). However, there was a much greater genetic divide between *M. macrocicatrices* in North Carolina and Virginia. Unique features of the area near the North Carolina-Virginia border, other than host-tree density, must therefore be contributing to the restriction of gene flow. Geologic attributes of the Blue Ridge Mountains, such as their irregularity and precipitous changes in elevation from 450 m to over 2000 m, may be factors contributing to hindered gametic exchange (Figure 3.7b). Significant genetic

structure in the Blue Ridge, and especially among populations on either side of the French Broad River in North Carolina, has been observed in several other taxa, such as snakes (Fontanella et al. 2008), salamanders (Crespi et al. 2003), centipedes (Garrick et al. 2018), and harvestmen (Hedin and McCormack 2017). It stands to reason that the terrain would also make it difficult for passive, wind-dispersed animals to be carried freely among suitable habitat patches of its host. The long, parallel mountains within the Ridge and Valley ecoregion channel wind along their axes (Whiteman and Doran 1993), whereas in the Blue Ridge ecoregion, prevailing winds travel perpendicular to mountain axes (Raichle and Carson 2009). Thus, impeding winds and irregular terrain, rather than just host plant distribution and density (Figure 3.7c), may explain why the Blue Ridge Mountains appear to significantly block geneflow in the area between the *NEast* and *SApps* clusters.

3.5 Conclusions

With high range-wide genetic diversity, no signatures of recent founder events, and clear genetic clusters separated by distinct geographic barriers, we reject the hypothesis that *M. macrocicatrices* is a non-native invader within the North American range of *P. strobus*. Hence, host trees currently experiencing dieback symptoms and mortality have likely co-evolved with this insect. Its small size, sessile nature, and seemingly benign impacts probably allowed it to remain undetected until the recent emergence of eastern white pine dieback symptoms in the mid-2000s. Costanza et al. (2018) reviews several ecological disturbance factors contributing to this phenomenon – such as climate change, land use, site conditions, and forest management – which may be

contributing to sudden *M. macrocicatrices* population growth. Assumed to be just one of over 250 innocuous herbivores of *P. strobus* (Wendel and Smith 1990), it remains a mystery why this native species has recently been associated with severe tree injury and mortality. The pathogenic fungus thought to exploit *M. macrocicatrices* feeding wounds and drive canker formation, *Caliciopsis pinea*, is also native (Ray 1936). This system presents a unique opportunity to understand how a native insect-pathogen complex, perhaps nonexistent or rare in the past, can become a transregional forest health concern. Excluding the possibility of a non-native invasion, as we have done in this study, narrows the search for why and how a species might become pestiferous. Our work has demonstrated the utility of establishing the origin of a pest in guiding ecosystem conservation.

3.6 Acknowledgements

We are especially grateful to those who assisted in sampling for this project: Kim Adams (SUNY-ESF), Arya Aghdasi (University of Georgia), Brittany Barnes (University of Georgia), Aaron Bergdahl (Maine Forest Service), Joe Bither (Maine Forest Service), Allison Kanoti (Maine Forest Service), Kyle Lombard (New Hampshire Division of Forests and Lands), Kaitlin Mooneyham (Virginia Department of Forestry), Joe O'Brien (USDA Forest Service), Jill Rose (West Virginia Department of Agriculture), Ashley Schulz (Arkansas State University), Tim Tomon (Pennsylvania DCNR), Jen Weimer (New Hampshire DNCR), and Jiangming Yao (Guangxi University). We also thank Nathan Havill (USDA Forest Service), Troy Kieran (University of Georgia), and Joseph Nairn (University of Georgia) for assistance with microsatellite development and Lyn

Cook (University of Queensland) and Penny Gullan (Australian National University) for suggestions on DNA extraction. This research was supported by the USDA Forest Service, Southern Research Station (13-CA-11330129-056 and 16-CS-11330129-045), Forest Health Protection, USDA Agriculture and Food Research Initiative, Foundation Grant, and D.B. Warnell School of Forestry and Natural Resources, University of Georgia.

3.7 References

Abrams, M.D., 2001. Eastern white pine versatility in the presettlement forest. Bioscience 51, 967-979.

Agapow, P.M., Burt, A., 2001. Indices of multilocus linkage disequilibrium. Molecular Ecology Notes 1, 101-102.

Asaro, C., Chamberlin, L.A., Rose, J.A., Mooneyham, K., Rose, A.K., 2018. Mortality of eastern white pine (*Pinus strobus* L.) in association with a novel scale insect-pathogen complex in Virginia and West Virginia. Forest Ecology and Management 423, 37-48.

Bagley, R.K., Sousa, V.C., Niemiller, M.L., Linnen, C.R., 2017. History, geography and host use shape genomewide patterns of genetic variation in the redheaded pine sawfly (*Neodiprion lecontei*). Molecular Ecology 26, 1022-1044.

Bean, J.L., Godwin, P.A., 1955. Description and bionomics of a new red pine scale, *Matsucoccus resinosae*. Forest Science 1, 164-176.

Bean, J.L., Godwin, P.A., 1971. Red Pine Scale. Forest Pest Leaflet 10. USDA Forest Service, p. 6.

Bonte, D., Van Dyck, H., Bullock, J.M., Coulon, A., Delgado, M., Gibbs, M., Lehouck, V., Matthysen, E., Mustin, K., Saastamoinen, M., Schtickzelle, N., Stevens, V.M., Vandewoestijne, S., Baguette, M., Barton, K., Benton, T.G., Chaput-Bardy, A., Clobert, J., Dytham, C., Hovestadt, T., Meier, C.M., Palmer, S.C.F., Turlure, C., Travis, J.M.J., 2012. Costs of dispersal. Biological Reviews 87, 290-312.

Bonte, D., Vandenbroecke, N., Lens, L., Maelfait, J.P., 2003. Low propensity for aerial dispersal in specialist spiders from fragmented landscapes. Proceedings of the Royal Society B-Biological Sciences 270, 1601-1607.

Booth, J.M., Gullan, P.J., 2006. Synonymy of three pestiferous *Matsucoccus* scale insects (Hemiptera: Coccoidea: Matsucoccidae) based on morphological and molecular evidence. Proceedings of the Entomological Society of Washington 108, 749-760.

Brookfield, J.F.Y., 1996. A simple new method for estimating null allele frequency from heterozygote deficiency. Molecular Ecology 5, 453-455.

Brown, A.H.D., Feldman, M.W., Nevo, E., 1980. Multilocus structure of natural populations of *Hordeum spontaneum*. Genetics 96, 523-536.

Brownstein, M.J., Carpten, J.D., Smith, J.R., 1996. Modulation of non-templated nucleotide addition by taq DNA polymerase: Primer modifications that facilitate genotyping. Biotechniques 20, 1004-1010.

Carey, M.P., Sanderson, B.L., Barnas, K.A., Olden, J.D., 2012. Native invaders: challenges for science, management, policy, and society. Frontiers in Ecology and the Environment 10, 373-381.

Carrete, M., Lambertucci, S.A., Speziale, K., Ceballos, O., Travaini, A., Delibes, M., Hiraldo, F., Donazar, J.A., 2010. Winners and losers in human-made habitats: interspecific competition outcomes in two neotropical vultures. Animal Conservation 13, 390-398.

Cornuet, J.M., Luikart, G., 1996. Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. Genetics 144, 2001-2014.

Costanza, K.K.L., Whitney, T.D., McIntire, C.D., Livingston, W.H., Gandhi, K.J.K., 2018. A synthesis of emerging health issues of eastern white pine (*Pinus strobus*) in eastern North America. Forest Ecology and Management 423, 3-17.

Crespi, E.J., Rissler, L.J., Browne, R.A., 2003. Testing Pleistocene refugia theory: phylogeographical analysis of *Desmognathus wrighti*, a high-elevation salamander in the southern Appalachians. Molecular Ecology 12, 969-984.

Davis, M.B., 1983. Quaternary history of deciduous forests of eastern North America and Europe. Annals of the Missouri Botanical Garden 70, 550-563.

Dodds, K.J., Aoki, C.F., Arango-Velez, A., Cancelliere, J., D'Amato, A.W., DiGirolomo, M.F., Rabaglia, R.J., 2018. Expansion of southern pine beetle into northeastern forests: management and impact of a primary bark beetle in a new region. Journal of Forestry 116, 178-191.

Dowton, M., Austin, A.D., 1998. Phylogenetic relationships among the microgastroid wasps (Hymenoptera: Braconidae): Combined analysis of 16S and 28S rDNA genes and morphological data. Molecular Phylogenetics and Evolution 10, 354-366.

Earl, D.A., Vonholdt, B.M., 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources 4, 359-361.

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

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

Faircloth, B.C., 2008. MSATCOMMANDER: detection of microsatellite repeat arrays and automated, locus-specific primer design. Molecular Ecology Resources 8, 92-94.

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

Flores-Renteria, L., Krohn, A., 2013. Scoring Microsatellite Loci. In: Kantartzi, S.K. (Ed.), Microsatellites: Methods and Protocols, Methods in Molecular Biology. Springer, Berlin, Germany.

Foldi, I., 2004. The Matsucoccidae in the Mediterranean basin with a world list of species (Hemiptera: Sternorrhyncha: Coccoidea). Annales De La Societe Entomologique De France 40, 145-168.

Fontanella, F.M., Feldman, C.R., Siddall, M.E., Burbrink, F.T., 2008. Phylogeography of *Diadophis punctatus*: Extensive lineage diversity and repeated patterns of historical demography in a trans-continental snake. Molecular Phylogenetics and Evolution 46, 1049-1070.

Gandhi, K.J.K., Herms, D.A., 2010. Direct and indirect effects of alien insect herbivores on ecological processes and interactions in forests of eastern North America. Biological Invasions 12, 389-405.

Garrick, R.C., Newton, K.E., Worthington, R.J., 2018. Cryptic diversity in the southern Appalachian Mountains: genetic data reveal that the red centipede, *Scolopocryptops sexspinosus*, is a species complex. Journal of Insect Conservation 22, 799-805.

Glenn, T.C., Nilsen, R., Kieran, T.J., Finger, J.W., Pierson, T.W., Bentley, K.E., Hoffberg, S., Louha, S., Garcia-De-Leon, F.J., del Rio Portilla, M.A., Reed, K., Anderson, J.L., Meece, J.K., Aggery, S., Rekaya, R., Alabady, M., Belanger, M., Winker, K., Faircloth, B.C., 2016. Adapterama I: Universal stubs and primers for thousands of dual-indexed Illumina libraries (iTru & iNext). bioRxiv 049114.

Goudet, J., 2005. HIERFSTAT, a package for R to compute and test hierarchical F-statistics. Molecular Ecology Notes 5, 184-186.

Hanks, L.M., Denno, R.F., 1998. Dispersal and adaptive deme formation in sedentary coccoid insects. In, Genetic structure and local adaptation in natural insect populations. Springer, Boston, MA, pp. 239-262.

Hapeman, P., Latch, E.K., Rhodes, O.E., Swanson, B., Kilpatrick, C.W., 2017. Genetic population structure of fishers (*Pekania pennanti*) in the Great Lakes region: remnants and reintroductions. Canadian Journal of Zoology 95, 869-876.

Hassan, A., Ricciardi, A., 2014. Are non- native species more likely to become pests? Influence of biogeographic origin on the impacts of freshwater organisms. Frontiers in Ecology and the Environment 12, 218-223.

Havill, N.P., Shiyake, S., Galloway, A.L., Foottit, R.G., Yu, G.Y., Paradis, A., Elkinton, J., Montgomery, M.E., Sano, M., Caccone, A., 2016. Ancient and modern colonization of North America by hemlock woolly adelgid, *Adelges tsugae* (Hemiptera: Adelgidae), an invasive insect from East Asia. Molecular Ecology 25, 2065-2080.

Hawes, T.C., Worland, M.R., Convey, P., Bale, J.S., 2007. Aerial dispersal of springtails on the Antarctic Peninsula: implications for local distribution and demography. Antarctic Science 19, 3-10.

Hedin, M., McCormack, M., 2017. Biogeographical evidence for common vicariance and rare dispersal in a southern Appalachian harvestman (Sabaconidae, *Sabacon cavicolens*). Journal of Biogeography 44, 1665-1678.

Hewitt, G.M., 1999. Post-glacial re-colonization of European biota. Biological Journal of the Linnean Society 68, 87-112.

Jactel, H., Menassieu, P., Vetillard, F., Gaulier, A., Samalens, J.C., Brockerhoff, E.G., 2006. Tree species diversity reduces the invasibility of maritime pine stands by the bast scale, *Matsucoccus feytaudi* (Homoptera: Margarodidae). Canadian Journal of Forest Research-Revue Canadienne De Recherche Forestiere 36, 314-323.

Kamvar, Z.N., Tabima, J.F., Grunwald, N.J., 2014. Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. Peerj 2, e281.

Keane, R.M., Crawley, M.J., 2002. Exotic plant invasions and the enemy release hypothesis. Trends in Ecology & Evolution 17, 164-170.

Kerdelhue, C., Boivin, T., Burban, C., 2014. Contrasted invasion processes imprint the genetic structure of an invasive scale insect across southern Europe. Heredity 113, 390-400.

Khamis, F.M., Karam, N., Ekesi, S., de Meyer, M., Bonomi, A., Gomulski, L.M., Scolari, F., Gabrieli, P., Siciliano, P., Masiga, D., Kenya, E.U., Gasperi, G., Malacrida, A.R., Guglielmino, C.R., 2009. Uncovering the tracks of a recent and rapid invasion: the case of the fruit fly pest *Bactrocera invadens* (Diptera: Tephritidae) in Africa. Molecular Ecology 18, 4798-4810.

Kuntner, M., Agnarsson, I., 2011. Phylogeography of a successful aerial disperser: the golden orb spider *Nephila* on Indian Ocean islands. Bmc Evolutionary Biology 11.

Lefort, M.C., Boyer, S., De Romans, S., Glare, T., Armstrong, K., Worner, S., 2014. Invasion success of a scarab beetle within its native range: host range expansion versus host-shift. Peerj 2, e262.

Little, E.L., 1971. Atlas of United States Trees, Volume 1: Conifers and important hardwoods. In. USDA Forest Service, Washington, DC.

Manni, F., Guerard, E., Heyer, E., 2004. Geographic patterns of (genetic, morphologic, linguistic) variation: How barriers can be detected by using Monmonier's algorithm. Human Biology 76, 173-190.

Mantel, N., 1967. Detection of disease clustering and a generalized regression approach. Cancer Research 27, 209-220.

McClure, M.S., 1976. Colonization and establishment of red pine scale, *Matsucoccus resinosae* (Homoptera: Margarodidae) in a Connecticut plantation. Environmental Entomology 5, 943-947.

McClure, M.S., 1977. Population-dynamics of red pine scale, *Matsucoccus resinosae* (Homoptera: Margarodidae): influence of resinosis. Environmental Entomology 6, 789-795.

McRae, B.H., Beier, P., 2007. Circuit theory predicts gene flow in plant and animal populations. Proceedings of the National Academy of Sciences of the United States of America 104, 19885-19890.

Mech, A.M., Asaro, C., Cram, M.M., Coyle, D.R., Gullan, P.J., Cook, L.G., Gandhi, K.J.K., 2013. *Matsucoccus macrocicatrices* (Hemiptera: Matsucoccidae): First report, distribution, and association with symptomatic eastern white pine in the southeastern United States. Journal of Economic Entomology 106, 2391-2398.

Meirmans, P.G., 2012. The trouble with isolation by distance. Molecular Ecology 21, 2839-2846.

Mendel, Z., 1998. Biogeography of *Matsucoccus josephi* (Homoptera: Matsucoccidae) as related to host resistance in *Pinus brutia* and *Pinus halepensis*. Canadian Journal of Forest Research-Revue Canadienne De Recherche Forestiere 28, 323-330.

Michigan Department of Natural Resources, 2015. 2015 Forest Health Highlights. p. 50.

Nackley, L.L., West, A.G., Skowno, A.L., Bond, W.J., 2017. The nebulous ecology of native invasions. Trends in Ecology & Evolution 32, 814-824.

Nadeau, S., Godbout, J., Lamothe, M., Gros-Louis, M.-C., Isabel, N., Ritland, K., 2015. Contrasting patterns of genetic diversity across the ranges of *Pinus monticola* and *P. strobus*: A comparison between eastern and western North American postglacial colonization histories. American Journal of Botany 102, 1342-1355.

Nahrung, H.F., Swain, A.J., 2015. Strangers in a strange land: do life history traits differ for alien and native colonisers of novel environments? Biological Invasions 17, 699-709.

Nei, M., 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 89, 583-590.

Nei, M., Maruyama, T., Chakraborty, R., 1975. Bottleneck effect and genetic variability in populations. Evolution 29, 1-10.

Oksanen, J., Blanchet, F.G., Friendly, M., Roeland Kindt, Legendre, P., McGlinn, D., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H.H., Szoecs, E., Wagner, H., 2018. vegan: community ecology package. R package.

Paolucci, E.M., MacIsaac, H.J., Ricciardi, A., 2013. Origin matters: alien consumers inflict greater damage on prey populations than do native consumers. Diversity and Distributions 19, 988-995.

Peakall, R., Smouse, P.E., 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Molecular Ecology Notes 6, 288-295.

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

Piry, S., Luikart, G., Cornuet, J.M., 1999. BOTTLENECK: A computer program for detecting recent reductions in the effective population size using allele frequency data. Journal of Heredity 90, 502-503.

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

Raichle, B.W., Carson, W.R., 2009. Wind resource assessment of the Southern Appalachian ridges in the southeastern United States. Renewable & Sustainable Energy Reviews 13, 1104-1110.

R Core Team, 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Ray, W.W., 1936. Pathogenicity and cultural experiments with *Caliciopsis pinea*. Mycologia 28, 201-208.

Raymond, M., Rousset, F., 1995. GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. Journal of Heredity 86, 248-249.

Richards, W.R., 1960. A new species of *Matsucoccus* Cockerell (Homoptera: Coccoidea). Canadian Entomologist 92, 179-181.

Richardson, D.M., Ricciardi, A., 2013. Misleading criticisms of invasion science: a field guide. Diversity and Distributions 19, 1461-1467.

Rozen, S., Skaletsky, H., 2000. Primer3 on the WWW for general users and for biologist programmers. Methods in molecular biology (Clifton, N.J.) 132, 365-386.

Rugman-Jones, P.F., Hoddle, M.S., Phillips, P.A., Jeong, G., Stouthamer, R., 2012. Strong genetic structure among populations of the invasive avocado pest *Pseudacysta perseae* (Heidemann) (Hemiptera: Tingidae) reveals the source of introduced populations. Biological Invasions 14, 1079-1100.

Sakai, A.K., Allendorf, F.W., Holt, J.S., Lodge, D.M., Molofsky, J., With, K.A., Baughman, S., Cabin, R.J., Cohen, J.E., Ellstrand, N.C., McCauley, D.E., O'Neil, P., Parker, I.M., Thompson, J.N., Weller, S.G., 2001. The population biology of invasive species. Annual Review of Ecology and Systematics 32, 305-332.

Salo, P., Korpimaki, E., Banks, P.B., Nordstrom, M., Dickman, C.R., 2007. Alien predators are more dangerous than native predators to prey populations. Proceedings of the Royal Society B-Biological Sciences 274, 1237-1243.

Schulz, A.N., Mech, A.M., Asaro, C., Coyle, D.R., Cram, M.M., Lucardi, R.D., Gandhi, K.J.K., 2018a. Assessment of abiotic and biotic factors associated with eastern white pine (*Pinus strobus* L.) dieback in the Southern Appalachian Mountains. Forest Ecology and Management 423, 59-69.

Schulz, A.N., Mech, A.M., Cram, M.M., Asaro, C., Coyle, D.R., Lucardi, R.D., Lucas, S., Gandhi, K.J.K., 2018b. Association of *Caliciopsis pinea* Peck and *Matsucoccus macrocicatrices* Richards with eastern white pine (*Pinus strobus* L.) seedling dieback. Forest Ecology and Management 423, 70-83.

Simberloff, D., Souza, L., Nunez, M.A., Barrios-Garcia, M.N., Bunn, W., 2012. The natives are restless, but not often and mostly when disturbed. Ecology 93, 598-607.

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

Spurgin, L.G., Illera, J.C., Jorgensen, T.H., Dawson, D.A., Richardson, D.S., 2014. Genetic and phenotypic divergence in an island bird: isolation by distance, by colonization or by adaptation? Molecular Ecology 23, 1028-1039.

Stephens, G.R., Aylor, D.E., 1978. Aerial dispersal of red pine scale, *Matsucoccus resinosae* (Homoptera: Margarodidae). Environmental Entomology 7, 556-563.

Tong, X., Wang, R., Chen, X.-Y., 2018. Expansion or invasion? A response to Nackley et al. Trends in Ecology & Evolution 33, 234-235.

Unruh, T.R., Luck, R.F., 1987. Deme formation in scale insects: a test with the pinyon needle scale and a review of other evidence. Ecological Entomology 12, 439-449.

van Etten, J., 2017. R Package gdistance: Distances and Routes on Geographical Grids. Journal of Statistical Software 76, 1-21.

Virgilio, M., Delatte, H., Backeljau, T., De Meyer, M., 2010. Macrogeographic population structuring in the cosmopolitan agricultural pest *Bactrocera cucurbitae* (Diptera: Tephritidae). Molecular Ecology 19, 2713-2724.

Watson, W.Y., Underwood, G.R., Reid, J., 1960. Notes on *Matsucoccus macrocicatrices* Richards (Homoptera: Margarodidae) and its association with *Septobasidium pinicola* Snell in eastern Canada. Canadian Entomologist 92, 662-667.

Weir, B.S., Cockerham, C.C., 1984. Estimating F-statistics for the analysis of population structure. Evolution 38, 1358-1370.

Whiteman, C.D., Doran, J.C., 1993. The relationship between overlying synoptic-scale flows and winds within a valley. Journal of Applied Meteorology 32, 1669-1682.

Whiting, M.F., Carpenter, J.C., Wheeler, Q.D., Wheeler, W.C., 1997. The strepsiptera problem: Phylogeny of the holometabolous insect orders inferred from 18S and 28S ribosomal DNA sequences and morphology. Systematic Biology 46, 1-68.

Zemanova, M.A., Knop, E., Heckel, G., 2016. Phylogeographic past and invasive presence of *Arion* pest slugs in Europe. Molecular Ecology 25, 5747-5764.

Zhang, B., Edwards, O.R., Kang, L., Fuller, S.J., 2012. Russian wheat aphids (*Diuraphis noxia*) in China: native range expansion or recent introduction? Molecular Ecology 21, 2130-2144.

Zinck, J.W.R., Rajora, O.P., 2016. Post-glacial phylogeography and evolution of a wide-ranging highly-exploited keystone forest tree, eastern white pine (*Pinus strobus*) in North America: single refugium, multiple routes. Bmc Evolutionary Biology 16.

Table 3.1. Population information for microsatellite analysis of *M. macrocicatrices*, including location data and estimates of genetic diversity.

Site ID	Site Name	County, state[a]	N	Latitude	Longitude	Elev. (m)	A_N	A_R	A_E	A_P	A_{LC}	F_{IS}	H_O	H_E
GA1*	Mill Creek	Murray, GA	12	34.8729	-84.7236	287	4.56	2.88	2.70	0.00	2.44	0.02	0.63	0.64
GA2*	Rock Creek	Gilmer, GA	20	34.7806	-84.3312	591	5.78	3.17	3.32	0.11	3	0.21	0.55	0.68
GA3†	Boggs Creek	Lumpkin, GA	20	34.7008	-83.8860	558	6.11	2.97	2.87	0.11	3	0.04	0.61	0.64
GA4†	Panther Creek	Habersham, GA	20	34.6996	-83.4195	454	4.78	2.76	2.73	0.00	2.33	0.08	0.56	0.61
TN*	Tellico Plains	Monroe, TN	20	35.3362	-84.1471	573	6.56	3.16	3.51	0.00	3.33	0.14	0.58	0.66
NC1*	Glenn Falls	Macon, NC	20	35.0359	-83.2352	1113	4.78	2.73	2.54	0.00	2.67	0.05	0.56	0.59
NC2*	Silver Mine	Madison, NC	16	35.9101	-82.7923	679	3.78	2.34	2.08	0.11	1.67	0.29	0.33	0.45
VA1*	Price Ridge	Bland, VA	20	37.1567	-80.9578	710	5.44	2.77	3.28	0.11	2.22	0.15	0.45	0.53
VA2†	Deerfield	Augusta, VA	20	38.2034	-79.3351	612	5.33	2.87	3.15	0.22	2.11	0.20	0.46	0.58
VA3†	Falling Springs	Alleghany, VA	19	38.1028	-78.7652	784	5.44	2.88	3.13	0.11	2.44	0.07	0.54	0.58
WV†	Watoga State Park	Pocahontas, WV	20	38.1110	-80.1076	840	4.11	2.29	2.46	0.00	1.56	0.05	0.38	0.40
PA1†	Cook Forest State Park	Clarion, PA	19	41.3535	-79.2193	453	4.11	2.55	2.59	0.22	1.44	0.16	0.44	0.52
PA2†	Rothrock State Forest	Centre, PA	20	40.7764	-77.6195	460	6.00	2.79	3.03	0.33	2.44	0.06	0.51	0.54
PA3†	Tioga State Forest	Tioga, PA	20	41.5272	-77.4506	283	4.78	2.45	2.93	0.00	1.78	0.11	0.38	0.43
NH1†	Bear Brook State Park	Merrimack, NH	20	43.1645	-71.3902	102	4.78	2.62	2.51	0.11	1.67	0.13	0.46	0.53
NH2†	Litchfield State Forest	Hillsborough, NH	20	42.8592	-71.4603	65	5.89	2.69	2.87	0.11	2.67	0.12	0.44	0.49
NH3†	Mast Yard State Forest	Merrimack, NH	20	43.2392	-71.6528	112	5.33	2.75	2.93	0.11	2.22	0.10	0.48	0.53
ME†	University of Maine Forest	Penobscot, ME	20	44.9307	-68.6850	38	5.44	2.72	3.30	0.22	2	0.04	0.50	0.52
MI1†	Grouse Trail	Crawford, MI	11	44.7028	-84.4086	332	3.56	2.45	2.35	0.00	1.56	0.10	0.45	0.50
MI2†	Frederic	Crawford, MI	19	44.7862	-84.7387	382	3.89	2.27	2.03	0.00	1.56	0.15	0.40	0.47
MI3*	Hiawatha National Forest	Delta, MI	8	45.9095	-86.8392	214	4.11	3.11	2.98	0.11	1.67	0.39	0.41	0.64
WI*	Lake Owen	Bayfield, WI	6	46.3243	-91.2372	410	2.78	2.45	1.99	0.11	1.33	0.48	0.25	0.46

A_N: mean number of alleles; A_E: effective number of alleles; A_R: rarefied allelic richness; A_P: number of private alleles per locus; A_{LC}: mean number of locally common alleles (≥5%) occurring in ≤50% of populations; F_{IS}: inbreeding coefficient; H_O: observed heterozygosity; H_E: unbiased expected heterozygosity.

* Sampled from shipped wood material.

† Sampled *in situ*.

[a] USA states: GA = Georgia; ME = Maine; MI = Michigan; NC = North Carolina; NH = New Hampshire; PA = Pennsylvania; TN = Tennessee; VA = Virginia; WI = Wisconsin; WV = West Virginia.

Table 3.2. The 12 novel microsatellite markers developed for *M. macrocicatrices*. For all statistical analyses, genotypes for the loci *Mmac-28, Mmac-68, and Mmac-93* were removed due to high null allele frequency.

Locus[a]	Sequence (5' – 3')	Repeat motif	Allele size range (bp)	GenBank Accession
Mmac-13[F]	F: [d]ACGCACTGTCAGACCATAAC	$(AATGCG)_6$	188-218	MK205304
	R: [b]ACAGCGATCGAATTATGCGC			
Mmac-15[V]	F: [d]TGTATCGTGTTCCTCTTGCAC	$(ACCCAT)_6$	178-226	MK205305
	R: [b]GGCGCTGAAAGACGTG			
Mmac-28[V]	F: [d]ACCCACTTCCAATTACAAGCC	$(AAGAG)_{10}$	226-296	MK205306
	R: [b]ACTGCTTTATGTGCCGACTTC			
Mmac-33[P]	F: [b]TAAAGGCCACGTCTTAGGGC	$(AATGT)_9$	178-236	MK205307
	R: [d]AAGTAATCGGGTCTGGCGG			
Mmac-36[N]	F: [b]GCATCGAACAGGTATGCGAG	$(ATCCC)_8$	181-211	MK205308
	R: [c]TCATTCATCCCACCACGTCC			
Mmac-56[N]	F: [b]GCAGTCGTATTTGGCCAATTC	$(ATCC)_{12}$	134-178	MK205309
	R: [c]TCGTGTGTCAGTCCCTAG			
Mmac-68[F]	F: [b]CTGGTACGAATGTGTCTGCG	$(AACC)_8$	130-175	MK205310
	R: [c]ACTCGTTGGTTTCGTTCTGC			
Mmac-70[P]	F: [b]GGCTAGATCATGATGCGTGC	$(ACTC)_8$	158-262	MK205311
	R: [d]TGGGATCTAGAGTGCTTCGTG			
Mmac-71[N]	F: [c]ATCAGCAGTTCCGGAGATCTC	$(ACAT)_8$	199-280	MK205312
	R: [b]GCGTTTGTTCGGATGGTAGG			
Mmac-76[P]	F: [b]GACGGCCACATGTGAAGATG	$(ACAG)_8$	230-250	MK205313
	R: [d]TGCGCCAGGAGATACTCG			
Mmac-93[V]	F: [d]CGTTTCAATGGCGTGATGTC	$(AAT)_{11}$	275-325	MK205314
	R: [b]CACGTCTGTTTCCCTCATCG			
Mmac-96[F]	F: [b]ACGCTTATCTACGTGGAC	$(AAT)_{10}$	106-127	MK205315
	R: [c]GCGATCGTAATTGGCAATCAG			

[a] Fluorophore used to label primers: F = FAM, V = VIC, N = NED, P = PET.
[b] GTTT tag added to 5′ end.
[c] CAG tag (CAGTCGGGCGTCATCA) added to 5′ end.
[d] M13 tag (GGAAACAGCTATGACCAT) added to 5′ end.

Table 3.3. Test statistics and p-values for both the index of association (I_A) (Brown et al. 1980) and the standardized index of multilocus linkage disequilibrium ($\bar{r}_d$) (Agapow and Burt 2001) of each population, calculated in the r package 'poppr' (Kamvar et al. 2014). Asterisk indicates a significant result.

Population	Ia	p.Ia	rbarD	p.rD
GA1	-0.207	0.797	-0.026	0.795
GA2	-0.117	0.814	-0.015	0.815
GA3	0.418	0.011*	0.053	0.011*
GA4	-0.103	0.759	-0.013	0.759
TN	0.042	0.378	0.005	0.378
NC1	0.03	0.359	0.004	0.358
NC2	-0.162	0.954	-0.021	0.954
VA1	-0.089	0.726	-0.013	0.727
VA2	0.031	0.567	0.004	0.567
VA3	0.001	0.446	0	0.446
WV	0.036	0.341	0.008	0.337
PA1	-0.045	0.6	-0.007	0.602
PA2	-0.056	0.615	-0.007	0.615
PA3	0.297	0.312	0.047	0.309
NH1	-0.152	0.949	-0.019	0.948
NH2	-0.274	0.961	-0.035	0.954
NH3	0.105	0.402	0.014	0.398
ME	0.219	0.05*	0.032	0.048*
MI1	-0.139	0.718	-0.02	0.722
MI2	0.173	0.146	0.024	0.131
MI3	-0.257	0.986	-0.033	0.986
WI	-0.071	0.672	-0.011	0.676

Table 3.4. Results from BOTTLENECK analysis (Luikart and Cornuet 1996) for each population and pooled group, using one- and two-tailed Wilcoxon sign rank tests of the single mutation model (SMM) and the two-phase model (TPM). Asterisk indicates a significant result.

				SMM		TPM	
Population	N	A_N	H_E	P (1t)	P (2t)	P (1t)	P (2t)
Grouped by sampled population							
GA1	23.11	4.56	0.643	0.98	0.13	0.85	0.36
GA2	38.22	5.78	0.685	0.98	0.13	0.33	0.65
GA3	40	6.11	0.638	1.00	0.01*	1.00	0.01*
GA4	39.78	4.78	0.606	0.99	0.02*	0.98	0.13
TN	39.78	6.56	0.662	1.00	0.00*	1.00	0.01*
NC1	39.78	4.78	0.590	0.88	0.30	0.79	0.50
NC2	29.11	3.78	0.454	0.99	0.03*	0.98	0.05*
VA1	39.78	5.44	0.534	0.98	0.05*	0.98	0.05*
VA2	39.11	5.33	0.576	0.94	0.16	0.82	0.43
VA3	36.44	5.44	0.578	0.99	0.03*	0.98	0.05*
WV	39.78	4.11	0.398	0.98	0.08*	0.98	0.08*
PA1	37.56	4.11	0.523	0.53	1.00	0.47	0.95
PA2	42	6	0.535	1.00	0.01*	1.00	0.01*
PA3	36.22	4.78	0.423	0.99	0.03*	0.99	0.03*
NH1	38.89	4.78	0.529	1.00	0.00*	1.00	0.01*
NH2	39.11	5.89	0.495	1.00	0.00*	1.00	0.00*
NH3	38.89	5.33	0.534	0.99	0.03*	0.92	0.20
ME	38.89	5.44	0.522	0.90	0.25	0.84	0.38
MI1	21.56	3.56	0.500	0.47	0.95	0.47	0.95
MI2	37.56	3.89	0.468	0.94	0.16	0.90	0.25
MI3	12.67	4.11	0.641	0.75	0.57	0.71	0.65
WI	9.78	2.78	0.465	0.99	0.04*	0.99	0.04*
Grouped by STRUCTURE clusters							
SApps	249.78	8.89	0.700	1.00	0.01*	0.94	0.16
NEast	426.67	10.44	0.569	1.00	0.00*	1.00	0.00*
GLakes	81.56	6.67	0.535	1.00	0.00*	1.00	0.00*

Table 3.5. Analyses of Molecular Variance (AMOVA). Significant *F*-statistics are bold (P < 0.05).

Test	Group Structure	Source of variation	Degrees of freedom	Sum of squares	Variance components	Percent variation	Fixation indices
	Range-wide						
A	No structure	Among populations	21	632	0.78	25.3	F_{ST} = **0.253**
		Within populations	758	1760	2.32	74.7	
		Total	779	2392			
B	*K* = 2	Between groups	1	284	0.69	20.1	F_{CT} = **0.201**
		Among populations	20	348	0.43	12.4	F_{SC} = **0.155**
		Within populations	758	1760	2.32	67.5	F_{ST} = **0.325**
		Total	779	2392	3.44		
C	*K* = 3	Between groups	2	381	0.81	23.4	F_{CT} = **0.235**
		Among populations	19	250	0.3	8.9	F_{SC} = **0.116**
		Within populations	758	1760	2.32	67.7	F_{ST} = **0.324**
		Total	779	2392	3.43		
	Within cluster						
D	NEast	Among subgroups	1	56	0.24	9.1	F_{CT} = **0.091**
		Among populations	9	75	0.15	5.9	F_{SC} = **0.065**
		Within populations	425	937	2.21	85	F_{ST} = **0.150**
		Total	435	1068	2.6		
E	SApps	Among subgroups	2	65	0.3	9.5	F_{CT} = **0.095**
		Among populations	4	40	0.2	6.2	F_{SC} = **0.069**
		Within populations	249	660	2.65	84.3	F_{ST} = **0.158**
		Total	255	765	3.15		
F*	GLakes	Among populations	3	15	0.15	7.1	F_{ST} = **0.071**
		Within populations	84	163	1.94	92.9	
		Total	87	178	2.09		

*No priors were set, because the *K* = 4 result from STRUCTURE implied a largely mixed ancestry among populations.

Table 3.6. Summary of Mantel tests. Covariate matrices were used in partial Mantel tests. Matrices that were subset by cluster indicated below. Asterisk indicates a significant result.

Matrix A	Matrix B	Covariate	R	P-value
Genetic	Geographic	--	0.509	$< 0.001^{*}$
Genetic	Geographic	Clusters	0.046	0.298
Genetic	Clusters	Geographic	0.653	$< 0.001^{*}$
Genetic: NEast	Geographic: NEast	--	0.335	0.005^{*}
Genetic: SApps	Geographic: SApps	--	0.833	$< 0.001^{*}$
Genetic: GLakes	Geographic: GLakes	--	0.835	0.042^{*}

Table 3.7. Population pairwise F_{ST} values. All pairwise relationships showed significant differentiation ($P < 0.05$) except those indicated with †.

Population	GA1	GA2	GA3	GA4	TN	NC1	NC2	VA1	VA2	VA3	WV	PA1	PA2	PA3	NH1	NH2	NH3	ME	MI1	MI2	MI3	WI
GA1																						
GA2	0.027																					
GA3	0.151	0.105																				
GA4	0.168	0.118	0.028																			
TN	0.102	0.082	0.053	0.056																		
NC1	0.185	0.153	0.049	0.091	0.091																	
NC2	0.264	0.253	0.242	0.230	0.251	0.232																
VA1	0.326	0.288	0.285	0.259	0.253	0.311	0.418															
VA2	0.306	0.273	0.259	0.260	0.209	0.282	0.352	0.112														
VA3	0.316	0.275	0.271	0.263	0.233	0.302	0.402	0.055	0.058													
WV	0.409	0.371	0.393	0.389	0.335	0.417	0.533	0.135	0.161	0.125												
PA1	0.323	0.306	0.320	0.325	0.258	0.351	0.388	0.206	0.079	0.151	0.295											
PA2	0.330	0.299	0.304	0.308	0.252	0.334	0.383	0.143	0.032	0.093	0.210	0.042										
PA3	0.412	0.356	0.362	0.376	0.308	0.388	0.476	0.212	0.072	0.167	0.269	0.079	0.035									
NH1	0.331	0.302	0.312	0.325	0.254	0.334	0.397	0.171	0.041	0.110	0.184	0.080	0.024	0.055								
NH2	0.361	0.321	0.327	0.348	0.275	0.350	0.428	0.190	0.044	0.118	0.213	0.064	0.025	0.020	0.001†							
NH3	0.325	0.287	0.302	0.315	0.252	0.330	0.387	0.177	0.044	0.114	0.212	0.056	0.027	0.024	0.017	0.007†						
ME	0.336	0.293	0.335	0.356	0.293	0.363	0.393	0.229	0.100	0.168	0.272	0.078	0.063	0.063	0.047	0.038	0.032					
MI1	0.301	0.297	0.259	0.248	0.222	0.311	0.434	0.254	0.263	0.247	0.314	0.360	0.336	0.426	0.332	0.370	0.348	0.400				
MI2	0.301	0.303	0.281	0.277	0.242	0.327	0.438	0.290	0.307	0.296	0.350	0.382	0.372	0.445	0.365	0.397	0.375	0.423	0.016†			
MI3	0.211	0.208	0.198	0.207	0.165	0.253	0.377	0.161	0.170	0.146	0.257	0.246	0.222	0.359	0.226	0.264	0.249	0.281	0.076	0.055		
WI	0.295	0.247	0.229	0.238	0.212	0.259	0.432	0.176	0.211	0.183	0.307	0.309	0.288	0.401	0.289	0.309	0.288	0.336	0.181	0.123	0.042†	
Mean	0.275	0.246	0.241	0.247	0.209	0.272	0.367	0.222	0.173	0.195	0.293	0.224	0.196	0.253	0.200	0.213	0.199	0.233	0.285	0.303	0.208	0.255

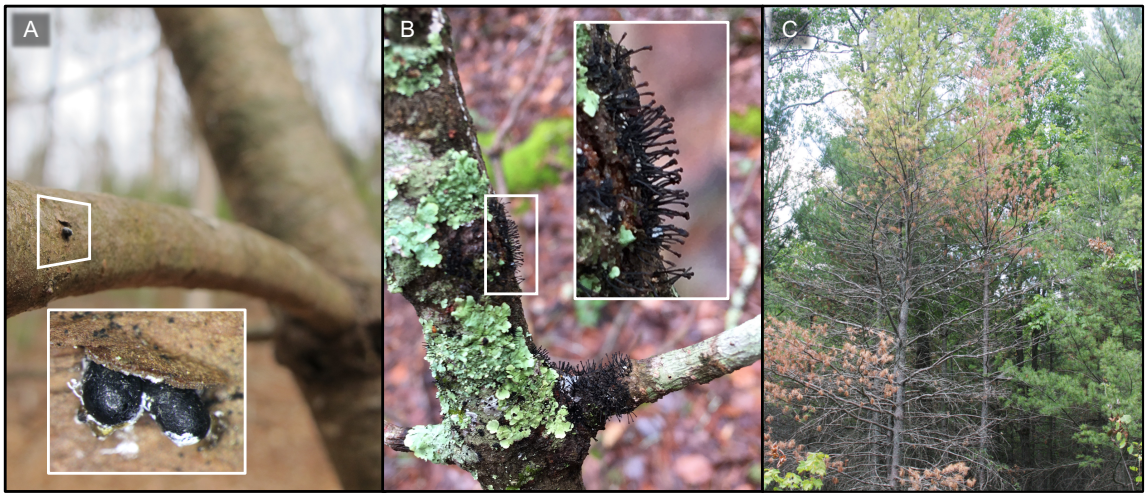

Figure 3.1. The insect-pathogen complex associated with eastern white pine dieback. (A) The eastern white pine bast scale (*Matsucoccus macrocicatrices*) will colonize in branch nodes, under lichen, and in bark crevices (inset), where feeding wounds may facilitate infection by (B) *Caliciopsis pinea* (inset shows the characteristic "eyelash-like" fruiting bodies), which drive canker development and leads to bottom-up branch dieback and mortality (C). Photo credit: Joe O'Brien (USDA Forest Service, C).

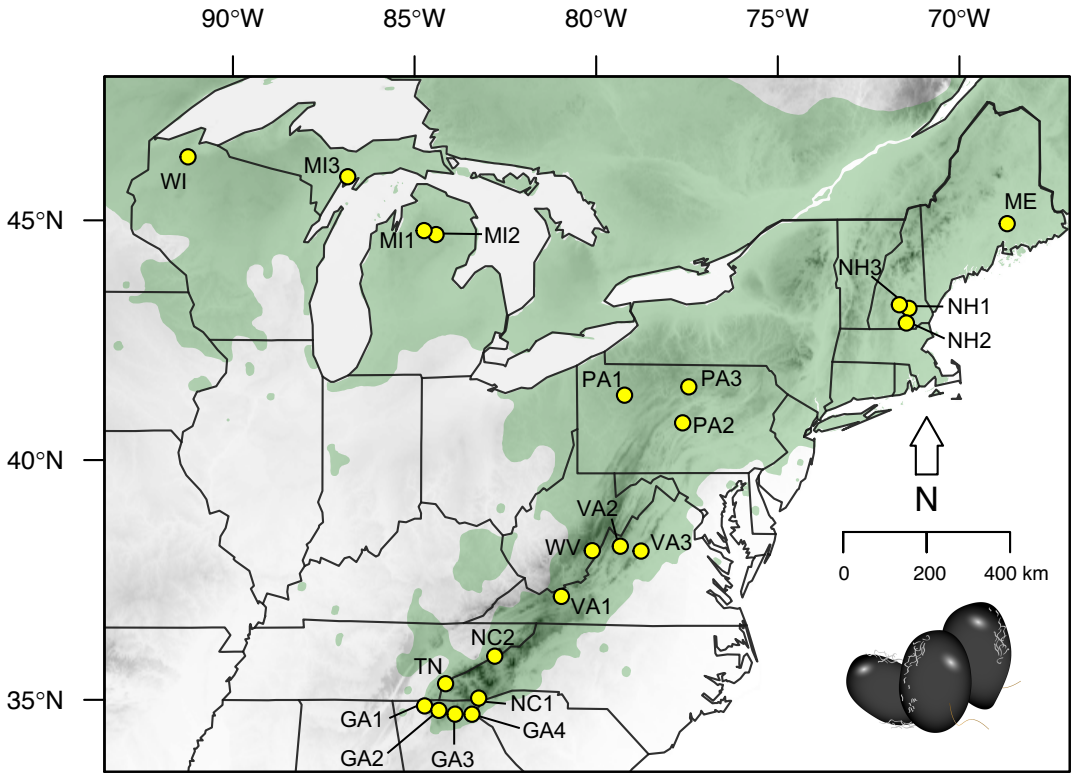

Figure 3.2. Sampled population sites of *M. macrocicatrices* included in molecular analyses. Shaded green area represents the range of its host, *P. strobus* (Little 1971).

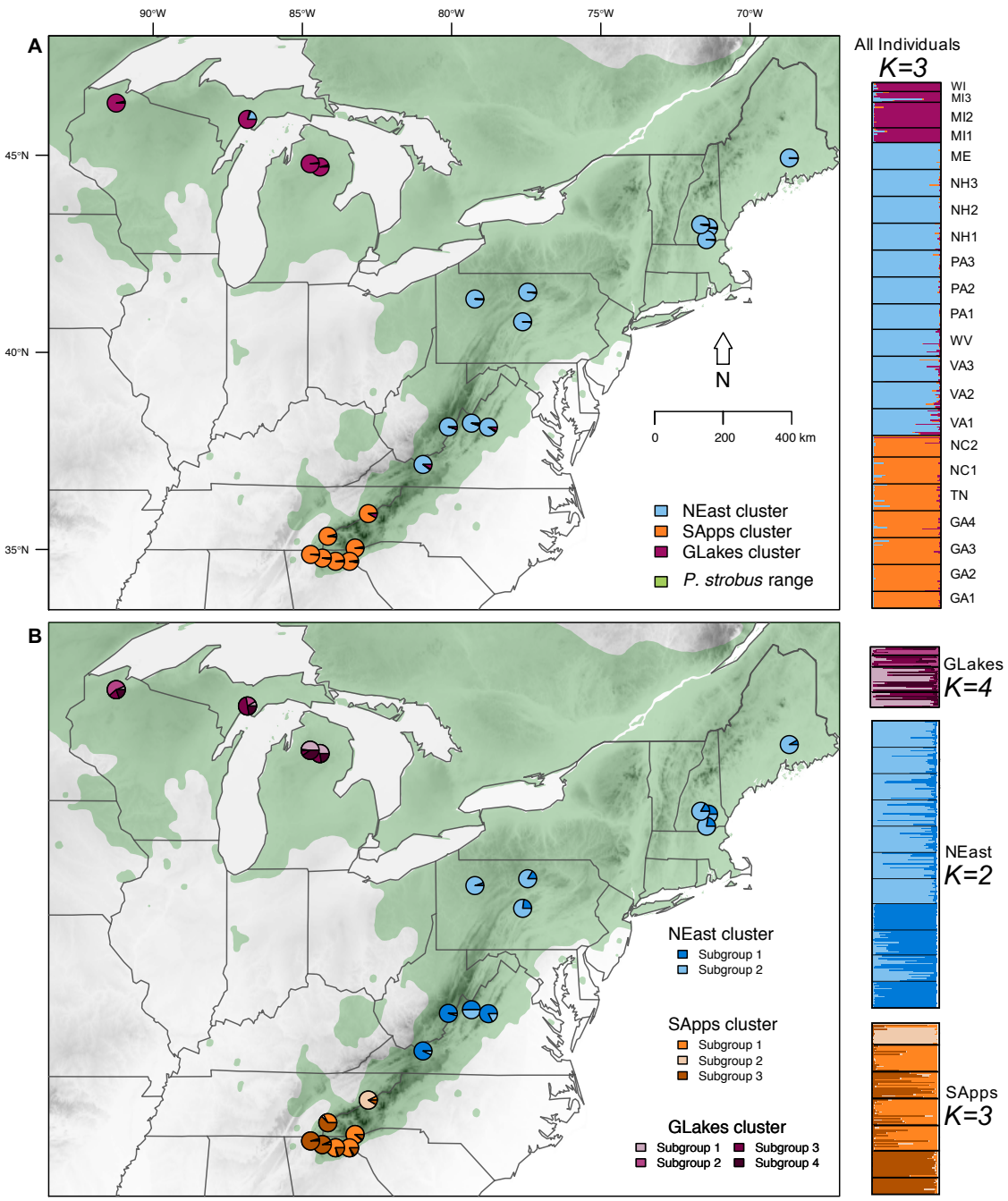

Figure 3.3. STRUCTURE results for *M. macrocicatrices*: (A) run including all samples; (B) runs including only populations from each inferred cluster ("*NEast*," "*SApps*" and "*GLakes*").

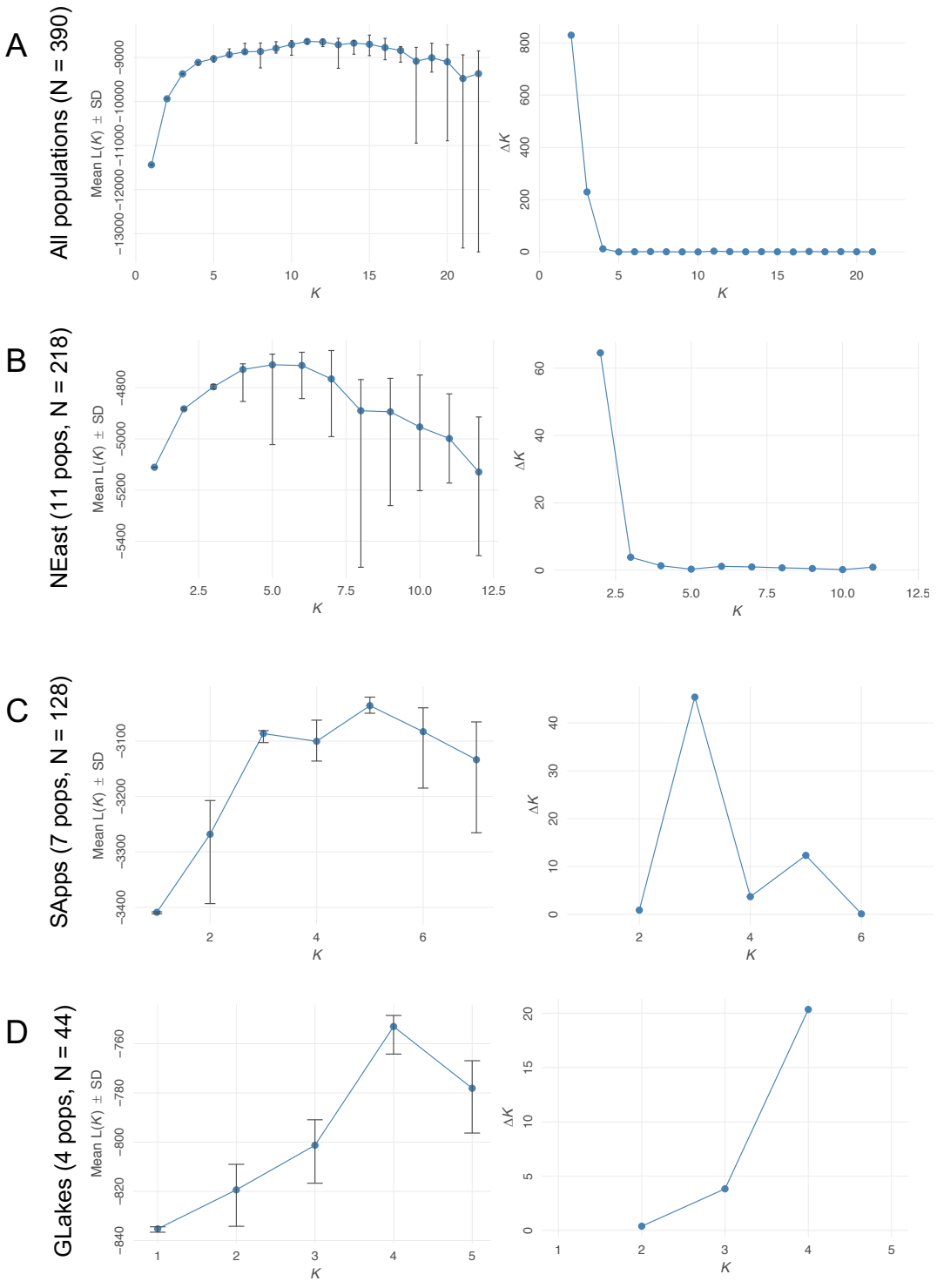

Figure 3.4. Mean L(*K*) and Δ*K* (Evanno et al. 2005) for the STRUCTURE runs conducted on (A) all populations and populations within the (B) *NEast* cluster, (B) *SApps* cluster, and (C) *GLakes* cluster. Data was calculated in STRUCTURE HARVESTER (Earl and Vonholdt 2012).

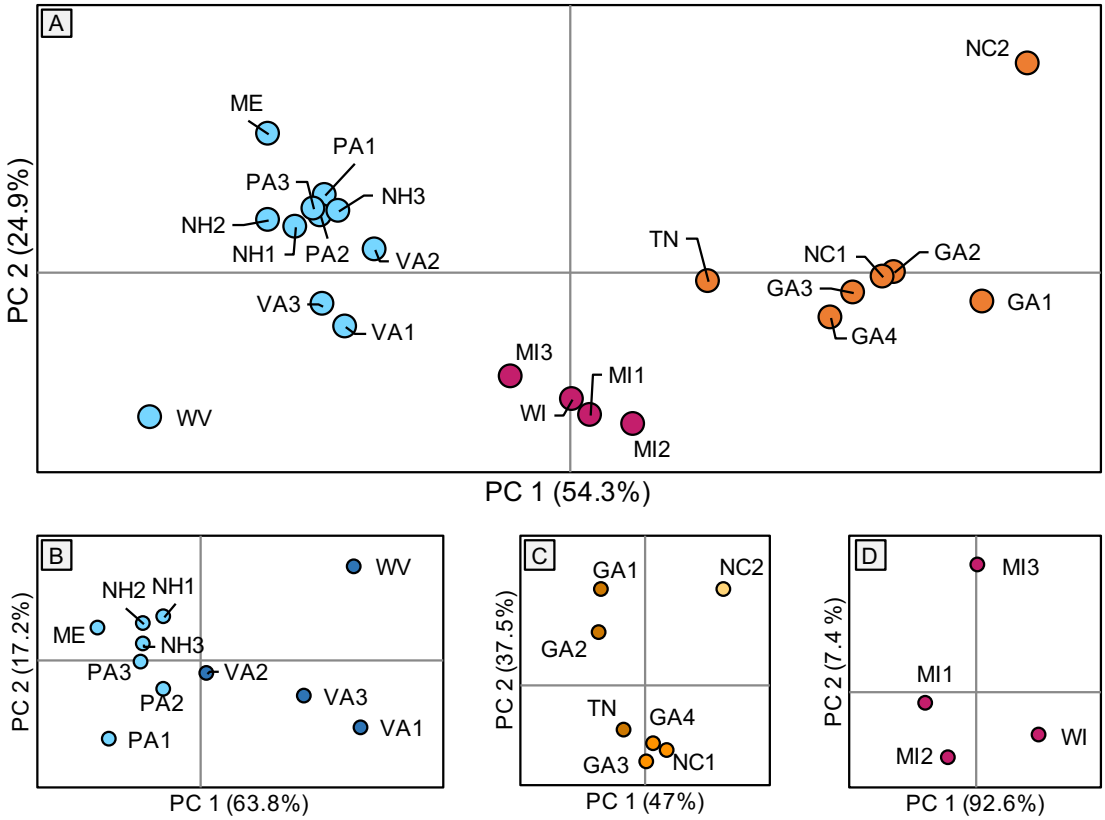

Figure 3.5. Principal coordinates analysis (PCoA) based on Nei's unbiased genetic distances of (A) all populations, (B) "*NEast*" populations, (C) "*SApps*" populations, and (D) "*GLakes*" populations.

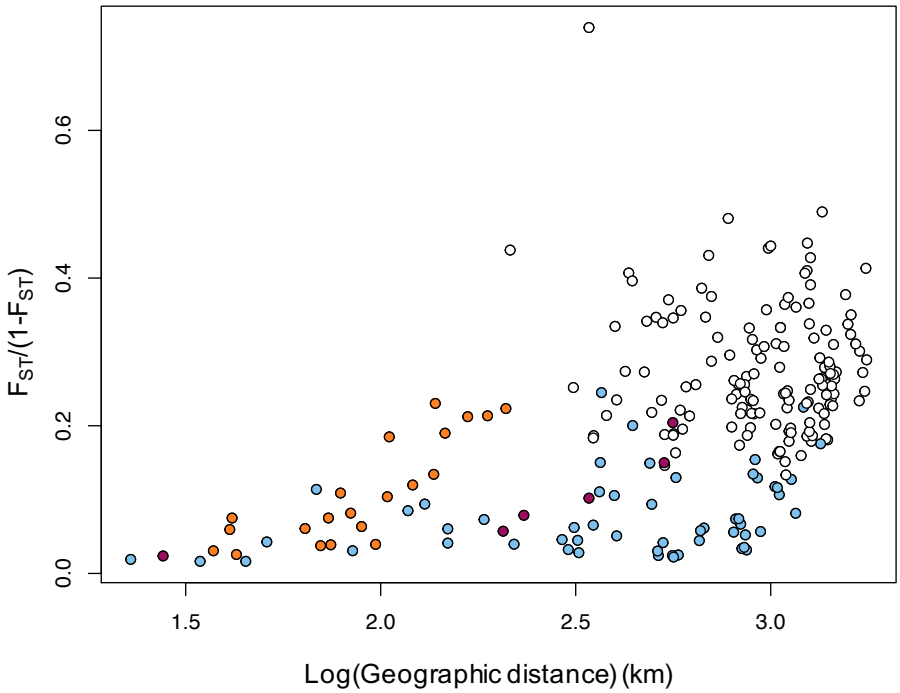

Figure 3.6. Association between geographic distance (log transformed) and Slatkin's linearized genetic distance among populations of *M. macrocicatrices*. Colored points indicate pairwise population comparisons within a cluster (blue = "*NEast*," orange = "*SApps*," and purple = "*GLakes*") and white points indicate pairwise population comparisons across clusters. There was significant isolation-by-distance within each cluster, but not over the entire range when controlling for cluster assignment (see Table 3.6).

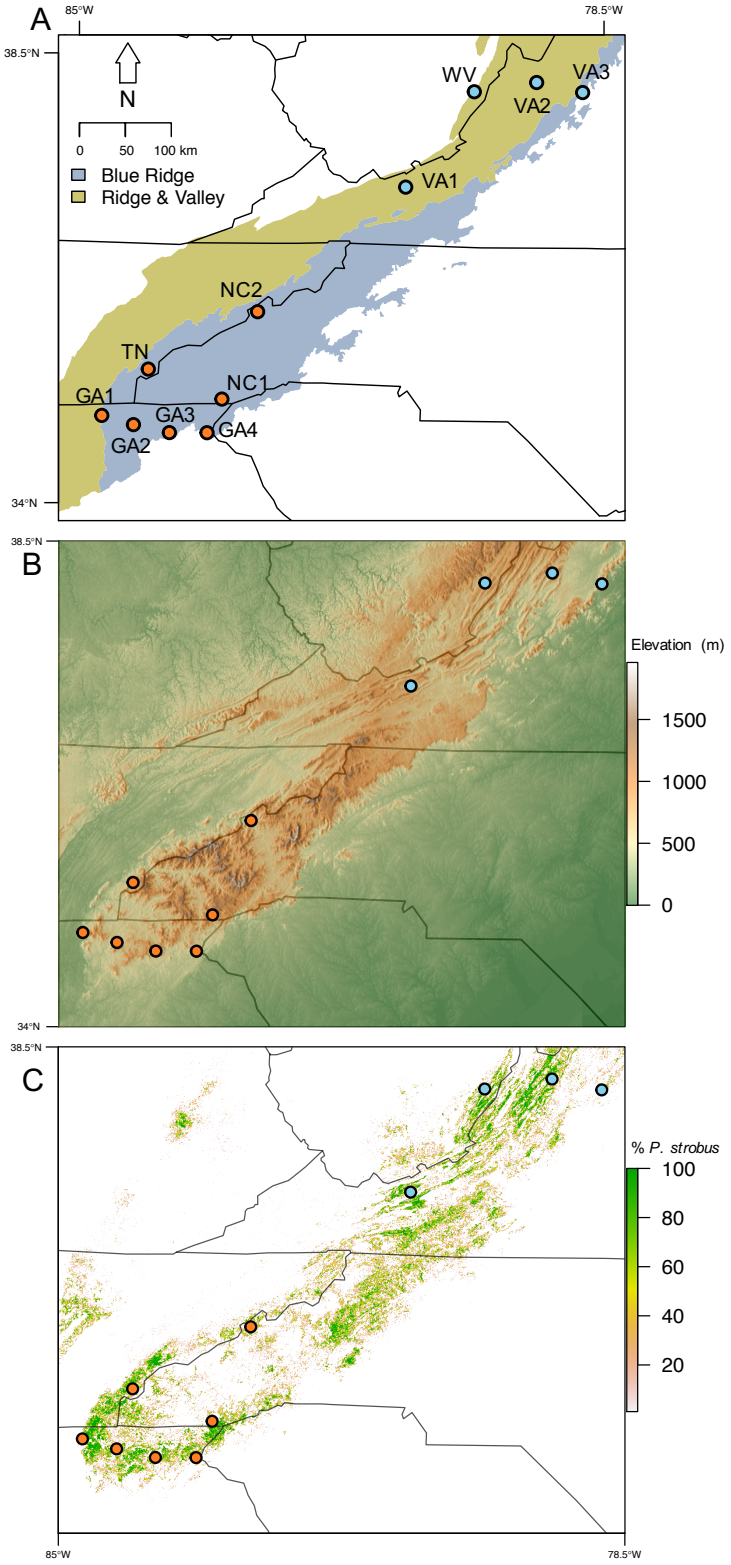

Figure 3.7. Abiotic variables that may influence the genetic barrier in the Blue Ridge mountains, including (A) Level III ecoregions, (B) elevation, and (C) host-tree density using FIA data, where each pixel indicates the percentage of *P. strobus* compared to total tree species for individuals ≥12.7 cm DBH, at 250-m resolution.

CHAPTER 4

EXTANT POPULATION GENETIC VARIATION AND STRUCTURE OF EASTERN WHITE PINE (*PINUS STROBUS* L.) IN THE SOUTHERN APPALACHIANS[5]

[5] Whitney T.D., Gandhi K.J.K., Hamrick, J. L., Lucardi R.D., 2018. In preparation for *Tree Genetics and Genomes.*

Abstract

Eastern white pine (*Pinus strobus* L.) is a widespread conifer species in eastern North America. A novel dieback phenomenon, as well as increasing global temperatures contributing to the contraction of suitable habitat, is threatening the long-term persistence of this species in the Southern Appalachian Mountains. Within this southern edge of its range is where *P. strobus* is hypothesized to have survived in refugial populations during the last glacial maximum. As a result, extant populations located here may hold high levels of ancestral genetic diversity, and by extension, adaptive potential for future modifications. We genotyped 432 individuals from 23 populations in the Southern Appalachians and 34 individuals from 2 populations in northern North America using ten established microsatellite markers. Levels of genetic diversity in the southern portion of its range were comparable but not higher than populations in the northern portion. There was an overall heterozygote deficiency and high inbreeding coefficient (F_{IS} = 0.173), but values were also similar to studies of northern populations. There was low overall genetic differentiation (F_{ST} = 0.055) in the Southern Appalachians and the little geographically defined population structure was best explained by ecoregional boundaries. Isolated populations west of the Appalachian Mountain chain harbored fewer private alleles but were not significantly depauperate genetically. These results show that *P. strobus* in the Southern Appalachians, and throughout its range, is a fairly heterogenous and admixed species with relatively high genetic diversity mostly partitioned within populations. The Southern Appalachians remains an important area for *P. strobus* conservation, but not necessarily because it is a unique genetic pool.

Keywords: conservation, genetic differentiation, genetic diversity, last glacial maximum, microsatellite, refugium

4.1 Introduction

The standing genetic variation of a species is essential for adaptation when faced with novel evolutionary pressures, because beneficial alleles that have been filtered through natural selection in the past are readily available at high frequencies (Barrett and Schluter 2008). Diverse gene pools can be especially important for the adaptive potential of long-lived species, such as forest trees, which often endure numerous disturbances over a lifespan (Schaberg et al. 2008). The rate of environmental change in the Anthropocene (e.g. due to climate change, and/or increased contact with exotic pests and pathogens) has become so rapid that it is likely to out-pace the rate that trees, with their long generation times, can modify and adapt (Aitken et al., 2008; Alfaro et al. 2014; Kuparinen et al. 2010, Savolainen et al. 2004, 2007). Hence, it is imperative to gather foundational data of the extant genetic variation for forest trees, especially for those species or populations that face imminent threats and/or may have particular conservation value (Potter et al. 2017).

Eastern white pine (*Pinus strobus* L.) is an important tree species in eastern North America, occurring from Minnesota, USA, to Nova Scotia, Canada, and south into northern Georgia, USA, along the Appalachian Mountain range. Its broad distribution is indicative of its versatility, thriving as both an early and late successional species, growing at both low (sea-level) and high elevations (up to 1220 m), and outcompeting hardwoods in both riparian valleys and nutrient-poor uplands (Abrams 2001, Wendel and

Smith 1990). Due to several favorable ecophysiological traits, *P. strobus* has been the most widely planted tree in eastern North America, often for use in erosion control and reclamation of previously mined or cultivated land (Hepp et al. 2015, Wendel and Smith 1990). Growing in pure or mixed stands as a scattered super-canopy tree, it provides resources and unique habitat qualities for wildlife (Abrams et al. 1995, Rogers and Lindquist 1992). With the continued decline of eastern hemlock (*Tsuga canadensis* L. Carriére) due to the invasive hemlock woolly adelgid (*Adelges tsugae* Annand), the ecological and economic importance of *P. strobus* – a frequently co-occurring canopy species – has become even more significant (Lovett et al. 2006).

In recent centuries, *P. strobus* has endured significant and recurring disturbances due to anthropogenic (e.g., intensive logging and fire exclusion) and biotic (e.g, white pine weevil, *Pissodes strobi* Peck, and white pine blister rust, *Cronartium ribicola* J.C. Fisch) factors (Costanza et al. 2018). Since the beginning of the 21st century, a novel and complex dieback phenomenon has also emerged in *P. strobus*, which has renewed concern for the immediate and long-term health and persistence of the species. Across its North American range, unique symptoms – including the presence of pathogenic fungal cankers, needle loss, branch dieback, and stem girdling – have arisen in conjunction with native biotic agents: a canker-forming pathogen (*Caliciopsis pinea* Peck) and a scale insect (*Matsucoccus macrocicatrices* Richards) (summarized in Costanza et al. 2018). Furthermore, climate prediction models suggest increasing temperatures are shifting the suitable climatic envelope of *P. strobus* northward at a rate faster than it can effectively disperse, which would result in significant range contraction at the expense of populations in the Southern Appalachians (Joyce and Rehfeldt 2013). At this southern

extent of its range in northwest Georgia, is where refugial populations are hypothesized to have survived during the last glacial maxima (LGM) of Pleistocene glaciations, based on evidence from the fossilized pollen record (Davis 1983, Jackson et al. 2000) and molecular phylogeography (Nadeau et al. 2015, Zinck and Rajora 2016). These populations served as source populations for post-glacial recolonization northward thousands of years ago, and thus, conventional wisdom suggests that the trees currently located in these areas would harbor the highest levels of extant genetic diversity (Hewitt 1996, but see Petit et al. 2003). South-to-north clines in genetic diversity, consistent with this "southern richness and northern purity" hypothesis (Hewitt 2000), have been reported in other plants and animals (Soltis et al. 1997, Talarico et al. 2019, Wielstra et al. 2013). If *P. strobus* follows a similar pattern, then their most vulnerable populations in the Southern Appalachians may hold exceptional adaptive potential and conservation value.

Prior to 2015, no study had investigated the population genetics of *P. strobus* south of Pennsylvania in the Southern Appalachians region (defined herein, as the USA states surrounding the area of the Appalachian Mountains: Georgia, Kentucky, North Carolina, South Carolina, Tennessee, Virginia, and West Virginia). As part of two recent phylogeographic studies, Nadeau et al. (2015) and Zinck and Rajora (2016) examined *P. strobus* populations across the entire North American range to reconstruct possible post-glacial recolonization routes. However, only 10 of 133 (Nadeau et al. 2015) and 2 of 33 (Zinck and Rajora 2016) total sampled populations were located in the Southern Appalachians. The majority of past research on the genetic diversity and structure of *P. strobus* predominately comes from the Canadian and the northern USA ranges (e.g.,

Beaulieu and Simon 1994, Rajora et al. 2000, Epperson and Chung 2001, Mehes et al. 2009). This geographic imbalance leaves a considerable gap in our understanding of the extant genetic diversity and adaptive potential in the area where *P. strobus* survived during and subsequently recolonized after the LGM.

The present study aimed to elucidate the genetic diversity, variability, and population structure of *P. strobus* in the Southern Appalachians. Its southern extent remains the most understudied, eminently at-risk, and potentially important part of its range for conservation. We used 10 highly-utilized, highly-polymorphic microsatellite markers (Echt et al. 1996) to genotype 432 individual *P. strobus* trees from 23 populations within what we have defined as its Southern Appalachian range. Eighteen of these populations were sampled along the main Appalachian Mountain chain where *P. strobus* is fairly contiguous from Virginia to Georgia. The remaining five populations were sampled west of the mountains in Tennessee, Kentucky and West Virginia in highly isolated, possibly relict, patches of *P. strobus*. For the sake of comparison and relationship-testing, we also genotyped 34 individual trees from 2 populations located in New Hampshire and Michigan. Additionally, we modelled the suitable climatic envelope of *P. strobus* at the LGM (~22,000 years ago), mid-Holocene (~6,000 years ago), and present day, to contextualize expected and emergent patterns from our genetic analyses. In bridging this foundational knowledge gap in *P. strobus* biology, we hoped to determine the similarity, or lack thereof, between the northern North American and Southern Appalachian range portions in: (1) the extant genetic diversity within and among populations, (2) the levels and patterns of genetic differentiation, and (3) the degree of genetic isolation in geographically isolated populations.

4.2 Methods

4.2.1 Sample collection

We sampled *P. strobus* needle tissue between 2014 to 2016 from 432 individual trees from 23 sites (populations, hereafter) throughout the Southern Appalachians in the USA, in the states of Georgia, Kentucky, North Carolina, South Carolina, Tennessee, Virginia, and West Virginia, as well as 14 individuals from one population in Michigan, and 20 individuals from a New Hampshire population (N = 466 individuals from 25 total populations) (Figure 4.1, Table 4.1). The mean size of sampled trees per site ranged from small saplings to poletimber: 2 cm to 15.9 cm diameter at breast height (DBH). Two of the populations were located in state forests, and the remaining populations were located within National Forests (USDA Forest Service). Sites were spaced by at least 10 km, a distance further than white pine pollen usually disperses (Epperson and Chung 2001). We sampled 14-20 individual trees per site, ensuring each tree was $\geq$10 m apart to mitigate the sampling of siblings. Within 48 hours of collection, needles were preserved at -80 °C.

4.2.2 Molecular analysis

To assess genetic diversity, variability, and population structure of *P. strobus*, we used twelve nuclear microsatellite markers – RPS1b, RPS2, RPS6, RPS12, RPS20, RPS25b, RPS34b, RPS39, RPS50, RPS60, RPS84, and RPS127 – previously designed by Echt (1996) (Table 4.2). These markers have been used extensively in past population genetics studies of *P. strobus* (Chhatre and Rajora 2014, Marquardt et al. 2007, Marquardt and Epperson 2004, Mehes et al. 2009, Myers et al. 2007, Mandak et al. 2013, Rajora et al. 2000, Walter and Epperson 2004, and Zinck and Rajora 2016).

All DNA extractions utilized the Qiagen DNeasy Plant Tissue Extraction Kit (Qiagen Inc., Chatsworth, California, USA) following the manufacturer's instructions. For each tree, five mature needles, weighed to approximately 0.1 g, were homogenized with a 3.2-mm diameter stainless steel bead within individual 2.0 µL microcentrifuge tubes placed within a Mini-Beadbeater-96 (BioSpec Products, Inc., Bartlesville, Oklahoma, USA). Genomic DNA quantity and quality was assessed with a BioPhotometer Plus (Eppendorf, Hamburg, Germany).

The majority of polymerase chain reactions (PCRs) to amplify microsatellite regions of interest were conducted in volumes of 12.5 µL, consisting of 2 µL 10x Takara buffer (Takara Bio Inc., Shiga, Japan), 1.5 µL 2.5 mM dNTPs, 0.25 µL both 10 µM forward and reverse primer, 0.1 µL 5 U/mL Takara *Ex Taq*™ polymerase (Takara), and 1-10 ng template DNA. For reactions amplifying RPS12 and RPS50, 0.1 µL 20 mg per mL bovine serum albumin (BSA) was added. The reactions that amplified the loci RPS1b and RPS2 were multiplexed in 25-µL volumes, consisting of 5 µL 10x Takara buffer, 4 µL 2.5 mM dNTPs, 0.15 µL both RPS1b primers and 0.2 µL both RPS2 primers, 0.25 µL 5 U/mL Takara *Ex Taq*™ polymerase, and 1-10 ng template DNA. Each forward primer for every reaction was tagged with a fluorophore: FAM, NED, PET, or VIC (Life Technologies, Carlsbad, California, USA). Each reverse primer had a GTTT pig-tail attached (Integrated DNA Technologies, Coralville, Iowa, USA) to reduce stutter during genotyping (Brownstein et al. 1996). Mastercycler Pro S Thermal Cyclers (Eppendorf) were used for PCR amplification under the following touchdown protocol: 2 cycles of 94 °C for 1 min, 60 °C for 1 min, and 70 °C for 35 s; 18 cycles of 93 °C for 1 min, 59-50.5 °C for 45 s decreasing 0.5 °C every cycle, and 70 °C for 35 s; 20 cycles of 92 °C for 30s,

50 °C for 30 s, and 72 °C for 60 s; and a final extension of 70 °C for 5 min. Amplicon sizes were determined on a 3730 capillary sequencer (Applied Biosystems, Foster City, California, USA) at the Arizona State University DNA Core Lab, with GeneScan LIZ 500 as the size standard (Life Technologies). Allele sizes were visualized and scored using the microsatellite plugin for GENEIOUS version 10.2.3 (Biomatters, Auckland, New Zealand).

4.2.3 Statistical analyses

4.2.3.1 Genetic diversity and variability

We tested all pairs of loci for linkage disequilibrium using the probability test, and we tested every locus-population combination for deviations from Hardy-Weinberg equilibrium with exact tests in GENEPOP version 4.2 (Raymond and Rousset 1995). We also calculated null allele frequencies of each locus for each population using the Brookfield (1996) method in GENEPOP and categorized the overall means as either negligible ($r < 0.05$), moderate ($0.05 \leq r < 0.2$), or large ($r \geq 0.2$) according to Chapuis and Estoup (2007).

Genetic diversity was estimated using effective number of alleles (A_E) and mean frequency of private alleles (A_P), and mean number of locally common alleles (≥5%) occurring in ≤50% of populations (A_{LC}) in GENALEX version 6.503 (Peakall and Smouse 2006, 2012). Rarefied allelic richness (A_R) was calculated in the R package 'hierfstat' (Goudet 2005). Observed heterozygosity (H_O), expected heterozygosity (H_E), and inbreeding coefficients (F_{IS}) were calculated using the Bayesian Individual Inbreeding Model (IIM) in INEST version 2.2 (Chybicki and Burczyk 2009), which takes into

account the presence of null alleles and the upward bias they can cause to F_{IS} estimates (Campagne et al. 2012). The Gibbs sampler in INEST utilized 20,000 burn-in steps, followed by 200,000 cycles, keeping every 1000th update.

We conducted generalized linear models (GLMs) in R version 3.5.1 (R Core Team 2019) to determine if the surrounding density of conspecifics, a measurement of the degree of isolation, influenced metrics of genetic diversity for our *P. strobus* populations in the Southern Appalachians only. We first used data from the USDA Forest Service Forest Inventory and Analysis (FIA) to create a density raster in R using the 'raster' package (Hijmans 2017), where each pixel (250 x 250 m) holds a value equal to the percentage of *P. strobus* comprising the total composition of trees ≥12.7 cm DBH. We then extracted the mean proportional density of *P. strobus* for the pixels surrounding each sample site at four different buffer sizes (radii): 1, 10, 50, and 100 km. In independent GLMs, we assessed the association of proportional conspecific density at these four spatial envelopes with each of the genetic diversity metrics calculated above. We also assessed if mean pairwise F_{ST} was associated with proportional density to determine if more isolated populations were more genetically distant. In a separate set of GLMs, we evaluated the association between latitude and longitude with these metrics of genetic diversity to determine if any geographic clines exist among the 23 sampled populations within the Southern Appalachians (N = 432). Latitude and longitude were included in these models simultaneously as covariates.

The program BOTTLENECK (Cornuet and Luikart 1996, Piry et al. 1999) was used to detect recent bottleneck events in our dataset. This program tests for deviations from mutation-drift equilibrium with the assumption that an excess in heterozygotes

results from a shrinking population and a deficiency in heterozygotes results from a population expansion. We tested all 25 populations (New Hampshire & Michigan included) separately each with 1000 permutations. We used the single-step mutation model (SMM) and the two-phase model (TPM) with 95% single-step mutations and 5% multi-step mutations (Piry et al. 1999). Significant excess and deficiency in heterozygosity was determined with one-tailed Wilcoxon signed-rank tests.

4.2.3.2 Population structure

We used STRUCTURE version 2.3.4 (Pritchard et al. 2000), a genetic clustering algorithm, to assign individuals to genetic groups. There were two sets of analyses, one assessing all individuals, including those from New Hampshire and Michigan (N = 466), and one assessing only the individuals from the Southern Appalachians focal region (N = 432). Each set consisted of two STRUCTURE runs: one with and one without the *LOCPRIOR* option. This parameter, which can improve detection of weak population structure without bias, informs STRUCTURE that individuals sampled at the same location are more likely to share ancestry (Hubisz et al. 2009). We assumed that individuals followed an admixture model and that allele frequencies were correlated among groups (Falush et al. 2003). Each STRUCTURE run varied the number of clusters (K) from 1 to 15 with 20 replicates for each K, using 25,000 burn-in steps, followed by 50,000 Markov Chain Monte Carlo (MCMC) iterations. The most optimal K in each analysis was determined using the $\mathrm{LnPr}(X|K)$ and ΔK methods described by Evanno et al. (2005) and implemented in the STRUCTURE HARVESTER pipeline (Earl and Vonholdt 2012).

To visualize pairwise genetic differentiation, we performed principal coordinates analyses (PCoA) using Nei's unbiased genetic distances (Nei 1978) in GENALEX. We also calculated F_{ST} across loci and between all pairs of populations using the *exclusion null alleles* method with 10,000 replicates in FREENA (Chapuis and Estoup 2007), which accounts for the presence of null alleles. The pairwise matrix of linearized F_{ST} values and a pairwise matrix of log-transformed geographic distances (km) of the 23 Southern Appalachian populations (N = 432) were used for a Mantel test (Mantel 1967) in GENALEX to detect if an isolation-by-distance (IBD) relationship existed within our focal study area.

We conducted analyses of molecular variance (AMOVA) using ARLEQUIN version 3.5 (Excoffier and Lischer 2010) to determine the hierarchical partitioning of genetic variance within individuals, within populations, among populations, and among pre-defined genetic groups in the Southern Appalachian focal region (23 populations, N = 432). We conducted three AMOVAs with 10,000 permutations to test for significance. The first test (A) utilized no priors and assumed no genetic structure. The second test (B) informed structure with two pre-defined groups: the five isolated populations west of the main Appalachian Mountain chain (Table 4.1: TN4, KY1, KY2, KY3, and WV1) and the remaining 18 populations within the main distribution of *P. strobus* along the Appalachian Mountain chain (Figure 4.2). The area within the Tennessee River Valley extending northeast along the Kentucky-Virginia border and into central West Virginia lacks *P. strobus* entirely, so this AMOVA sought to determine if this valley might act as a barrier to dispersal, influencing the partitioning of genetic variation. The third test (C) assigned each of the 23 populations based on Level III ecoregion (U.S. Environmental

Protection Agency 2013; see Table 4.1) to inform genetic structure: Blue Ridge (n = 15), Ridge and Valley (n = 5), Southwestern Appalachians (n = 1), and West Allegheny Plateau (n = 2) (Figure 4.3). This AMOVA examined if the ecoregions of the Southern Appalachians, which vary widely in elevation, aspect, geology, and soils, influenced *P. strobus* genetic differentiation.

4.2.4 Climate suitability modelling

To provide a historical context to the current *P. strobus* distribution and genetic structure, we estimated its suitable climatic envelope at three time points: the LGM (~22,000 years ago), the mid-Holocene (~6,000 years ago), and the "present day" (1970-2000). We utilized MAXENT version 3.4.0 (Phillips et al. 2006), which uses a maximum entropy algorithm to predict species distribution based on models of the relationship between occurrence data and in this case, abiotic environmental variables. *Pinus strobus* occurrence data for MAXENT was derived from the relative density raster created above (Figure 4.2b). All cells containing 100% *P. strobus* among trees ≥12.7 cm DBH were extracted as points (totaling 55,141) to be used for the model. We obtained 19 climatic variables from the WorldClim database (www.worldclim.org) (Hijmans 2005) as biolayers in our models for the climatic suitability of *P. strobus* (Table 4.3) at a resolution of 2.5 arc-minutes. The bioclimatic envelope was projected onto LGM and mid-Holocene climate data based on CCSM4. Biolayers BIO2, BIO3, BIO14, and BIO15 were removed from LGM simulations due to their inconsistent estimation among models (Varela et al. 2015). Beyond this, we opted not to use *a priori* variable selection methods; the regularization procedure in MAXENT addresses model selection by weighing both

informative and uninformative variables appropriately to balance model fit and complexity (Elith et al. 2011). We used a convergence threshold of 10^{-5} and 500 iterations. A randomly chosen set of 3,899 occurrence points was used for training the model, and 1,299 points (25%) were used for testing the model. The model's AUC score (area under the curve of the receiver operating characteristic) was evaluated for predictive accuracy, which varies from 0 (imperfect) to 1 (perfect), with a value of 0.5 indicating random predictive power. An AUC score above 0.7 is considered good model performance (Fielding & Bell 1997).

4.3 Results

4.3.1 Microsatellite loci

Two loci, RPS25b and RPS60, were removed from analyses due to inconsistent amplification and scoring. The ten remaining microsatellite loci amplified in this study yielded 128 total alleles across the 466 samples of *P. strobus* (Table 4.4), ranging from a minimum of 5 alleles (RPS127) to a maximum of 26 (RPS12). There was 100% polymorphism across loci for each population, and each individual had a unique multi-locus genotype. Estimates of null allele frequency were considered either moderate ($0.05 \leq r < 0.2$) or negligible ($r < 0.05$) for all loci. The Bayesian IIM procedure from INEST revealed that both inbreeding and null alleles were important factors; the full model, including inbreeding, null alleles, and genotyping errors (DIC = 19980.21), fit the data better than either the model assuming no null alleles (DIC = 20069.64) or the model assuming no inbreeding (DIC = 20056.59). The total adjusted inbreeding coefficient (F_{IS}) from INEST was 0.187. This is indicative of an overall deficiency of heterozygotes.

Observed heterozygosity (mean of 0.492 across all 10 loci) was much lower than expected heterozygosity (mean of 0.620). Of the 250 locus-population combinations, exact tests revealed a significant departure from Hardy-Weinberg Equilibrium in 120 combinations, but with no clear concentration in any particular locus or population. No pair of loci showed significant linkage disequilibrium except between RPS6 and RPS34b.

4.3.2 Genetic diversity and variability

Genetic diversity metrics were assessed per population and are summarized in Table 4.1. There was a mean of 4.86 alleles per locus (A_N) and 4.21 alleles per locus when standardizing for sample size (A_R). Private alleles, which are alleles that only occur in one population but no other, were present in 13 of the 25 total sampled populations. Private alleles were absent from all Kentucky populations, all Tennessee populations, 3 of 4 North Carolina populations, a Virginia population, and a Georgia population. The majority of private alleles within the Southern Appalachians were present in the lowest latitude populations in Georgia and South Carolina and also in higher latitude populations in Virginia and West Virginia. The New Hampshire and Michigan populations had similar levels of genetic diversity to the overall values of the Southern Appalachian study region according to most metrics, with the exception that they had comparatively lower inbreeding coefficients (F_{IS} = 0.046 and 0.073, respectively).

Generalized linear models (GLMs) provided little support that proportional density of surrounding conspecifics within the Southern Appalachians region influenced genetic diversity, as there was no association between *P. strobus* density and inbreeding coefficient, heterozygosity, or number of rare alleles at any radius. However, in

populations where the surrounding proportional density of *P. strobus* increased at radii of 50 km and 100 km, respectively, the number of private alleles significantly increased ($F_{1,21}$ = 5.38, t = 2.32, P = 0.03) and allelic richness increased ($F_{1,21}$ = 4.08, t = 2.021, P = 0.06). Additionally, as proportional density at a radius of 100 km increased, mean pairwise F_{ST} values decreased ($F_{1,21}$ = 3.88, t = -1.97, P = 0.06). When investigating if longitudinal or latitudinal patterns in genetic diversity exist, GLMs revealed no association with any of the tested indices.

No population we sampled had an excess of heterozygotes from the one-tailed Wilcoxon tests of the SMM and TPM (Table 4.5), which would be indicative of recent bottleneck event(s). We detected significant heterozygote deficiency, however, in 14 and 16 of the 23 Southern Appalachian populations according to the SMM and TPM, respectively. We also detected a heterozygote deficiency in the New Hampshire but not the Michigan population.

4.3.3 Population structure

The maximum ΔK value and the plateau of the LnPr(X|K) curve for runs in STRUCTURE revealed the best number of clusters at the highest level of hierarchical analysis (N = 466, 25 populations; Figure 4.4a) and within the Southern Appalachians only (N = 432, 23 populations; Figure 4.4b) to be K = 2 (Figure 4.5). The mean posterior probabilities of each population, including those from New Hampshire and Michigan, were mostly admixed across the landscape and no clear geographical pattern emerged. The STRUCTURE runs conducted with the *LOCPRIOR* option differed negligibly, also

resulting in a best $K = 2$, with little population structure emerging at higher K-values (Figure 4.5).

Nei's unbiased genetic distances, visualized by a Principle Coordinates Analysis (PCoA), accounted for 77.4% of genetic variation explained by the first two axes. Results showed little overall genetic differentiation (Figure 4.5a) among the majority of populations, including NH and MI. The only populations with substantial differentiation on one or both of the principal coordinates were from Georgia (population GA2), West Virginia (population WV1), and Kentucky (population KY3). These three populations also had the highest mean pairwise F_{ST}, with values from 0.105-0.149, whereas all other populations had F_{ST} values less than 0.071 (Table 4.6). Another PCoA, excluding populations GA2, WV1, and KY3, resulted in 56% of variation explained by the first two axes (Figure 4.5b). The Mantel test revealed a significant association between genetic distance and geographic distance ($R = 0.31$, $P = 0.01$) among the 23 Southern Appalachian populations, suggesting *P. strobus* has spatial genetic structure consistent with isolation-by-distance in this region (Figure 4.7).

The overall F_{ST} value of our entire dataset, calculated using the *exclusion null alleles* method in FREENA, was 0.060. Within the Southern Appalachians only, the F_{ST} of 0.055 from the AMOVA without any population structure was significant ($P < 0.001$), but 94.5% of genetic variation was partitioned within populations and only 5.5% of genetic variation partitioned among populations (Table 4.7a). The AMOVA with predefined structure, separating the five isolated populations west of the Appalachian Mountain chain from the 18 populations located within the mountains, explained slightly more genetic differentiation than the null model ($F_{ST} = 0.056$, $P < 0.001$), however, the

amount of variation among groups (F_{CT} = 0.002, P = 0.25) was not significant (Table 4.7b). The last AMOVA to utilize pre-defined population structure, which separated populations according to their respective Level III ecoregion (Figure 4.3), best explained genetic differentiation across the Southern Appalachians (F_{ST} = 0.062, P < 0.001) and accounted for a significant amount of variation among groups (F_{CT} = 0.019, P = 0.02) (Table 4.7c).

4.3.4 Climate suitability modelling

The maximum entropy model performed substantially better than random (AUC = 0.844) (Figure 4.8). The maximum temperature in the warmest month (BIO05) and the mean temperature in the warmest quarter (BIO10) were consistently among the highest contributing variables to model performance in all three projections. The suitable climatic envelope for *P. strobus* during the LGM (~22,000 years ago) was highest in what is now the North Carolina coastal plain and the Piedmont region of Alabama, Georgia, North Carolina and South Carolina (Figure 4.9a). There is also a narrow, long stretch of suitable habitat directly west of the Appalachian Mountains in northern Georgia and eastern Tennessee, an area of the Ridge and Valley ecoregion where no *P. strobus* exists today (Figure 4.2b). By the mid-Holocene warm period (~6,000 years ago), the climate suitability for *P. strobus* had shifted northward substantially (Figure 4.9b), consistent with the steady movement of pollen from the fossil record (Davis 1983). Isolated islands of suitable habitat characterize most of the landscape during this time, including within our focal Southern Appalachians study region. The present-day climatic envelope (Figure 4.9c) is consistent with the current *P. strobus* range and density (Figures 4.1 & 4.2), with

high and contiguous suitability values along the southern half of the Appalachian Mountain chain and moderate habitat suitability west of the mountains where samples were collected from populations TN4, KY1, KY2, KY3, and WV1.

4.4 Discussion

4.4.1 Genetic diversity

Levels of genetic diversity in populations of *P. strobus* in the Southern Appalachians were comparably high, but not higher than populations in the northern latitudes of North America. The mean observed heterozygosity across populations of 0.477 falls within the range (albeit, on the lower end) of values reported in recent microsatellite studies from northern populations of 0.432 to 0.740 (Table 4.8). Allozyme studies (e.g. Beaulieu and Simon 1994, Buchert et al. 1997, Rajora et al. 1998) report much lower levels of observed heterozygosity (0.121 – 0.265), but this is likely due to allozyme loci being less variable than microsatellite loci (Hedrick 1999). From our analysis, the New Hampshire and Michigan populations possessed higher allelic richness and heterozygosity than populations in the Southern Appalachians on average. Within the Southern Appalachian region, we found no latitudinal gradients in genetic diversity estimates (from Table 4.1). The two previous studies to have included any Southern populations yielded both contrasting and affirmatory results. Zinck and Rajora (2016) reported a south-to-north decrease in several metrics, including allelic richness and heterozygosity, but Nadeau et al. (2015) found no latitudinal clines in genetic diversity. In fact, Nadeau et al. (2015) showed there was slightly less genetic diversity in populations at the southern extent of their sampling. They attributed this pattern to

genetic drift acting on small and sparse northern Georgia refugial populations during the LGM (Jackson et al. 2000). Similarly, the highest levels of genetic diversity in red pine (*P. resinosa* Ait.) do not occur at the origin of their postglacial recolonization in the Southern Appalachians (Walter and Epperson 2001).

With severe inbreeding depression (Johnson 1945, Fowler 1965) and a higher outcrossing rate than what is typical in other conifers (Beaulieu and Simon 1995), *P. strobus* has been reported with low levels of inbreeding (Rajora et al. 2000, Epperson and Chung 2001, Marquardt and Epperson 2004). More recent studies, genotyping individuals over a broader distribution and with the same microsatellite loci used in this study, report fairly high mean F_{IS} values ranging from 0.072 to 0.215 (Table 4.8); the mean inbreeding coefficient of only our Southern Appalachian populations (F_{IS} = 0.173) was comparable within this range. There was also an overall deficit in heterozygosity, similar to northern *P. strobus* populations. The fact that we sampled young trees (<16 cm DBH) may have also inflated our F_{IS} values, because as stands of trees age, selection generally occurs against inbred progeny (Hufford and Hamrick 2003, Strauss 1986). However, inbreeding may yet be slightly stronger in the Southern Appalachians, as the New Hampshire and Michigan populations observed in this study had lower F_{IS} values. Zinck and Rajora (2016) reported a similar pattern with higher inbreeding coefficients in a single Virginia (F_{IS} = 0.020) and a single North Carolina (F_{IS} = 0.021) population relative to the overall mean of mostly northern populations (F_{IS} = 0.10). The fact that populations in the Southern Appalachians are generally smaller and more scattered (Abrams 2001) may explain why signals of inbreeding may be stronger than those populations further north.

The degree of isolation for *P. strobus* populations in the Southern Appalachians influenced some but not all metrics of genetic diversity. Heterozygosity and F_{IS} remained mostly unchanged, but number of private alleles and mean allelic richness decreased as the proportional density of conspecifics varied at a large geographic scale (radii of 50 and 100 km). Chhatre and Rajora (2014) similarly found that disjunct populations of *P. strobus* in Ontario, Canada, had reduced allelic richness and number of private alleles, but had similar heterozygosity and F_{IS} values when compared to continuous populations. Although we found no evidence of genetic bottleneck in any population, the smaller size of more isolated populations in our dataset may account for the lack of private alleles in general. Due to the longevity of *P. strobus*, it is possible that mechanisms like genetic drift have not yet reduced heterozygosity or increased inbreeding levels in more isolated populations. Only at the largest geographic scales and the highest degrees of isolation were proportional *P. strobus* densities correlated with any differences in genetic diversity indices. This is probably indicative of the great dispersal ability of *P. strobus*. Wind-dispersed pollen and seed usually travel less than 1 km, but in open landscapes they have the potential to travel several kilometers (Wilson and McQuilkin 1963, Munzbergova et al. 2010, Williams 2010). This long-range dispersal ability likely maintains a high level of genetic diversity among distant and/or isolated *P. strobus* populations.

4.4.2 Population structure

Little overall genetic differentiation is typical of gymnosperms (Hamrick and Godt 1996) and also of *P. strobus* in northern North America. In the Southern Appalachians, we similarly found a low mean F_{ST}, comparable with northern studies

(Table 4.8). Although low, this value (F_{ST} = 0.55) was also significant according to an AMOVA with no *a priori* structure (Table 4.7a), suggesting there is at least a non-negligible amount of genetic differentiation. The Mantel test also suggested there is a spatial genetic structure consistent with isolation-by-distance, a finding that is to be expected of wind-pollenated plants (Hamrick and Nason 1996) and is consistent with *P. strobus* in northern populations (e.g., Epperson and Chung 2001, Marquardt et al. 2007, Nijensohn et al. 2005, but also see Mandak et al. 2013). Nevertheless, population structure appeared to be weak overall among the 23 Southern Appalachian populations, as we did not identify any geographic patterns of substructure. Although STRUCTURE revealed K = 2, the posterior probabilities of cluster assignments suggested nearly universal admixture across populations (Figure 4.4). This was likely an erroneous default result, because STRUCTURE cannot assess K = 1 as a potential scenario (Janes et al. 2017). The PCoA echoed this, as 22 of the 25 total populations, including those from New Hampshire and Michigan, were clustered closely together (Figure 4.6a). Overall, genetic diversity within the Southern Appalachians was overwhelmingly partitioned within populations, and not among them, a result also consistent with *P. strobus* in the northern region (Table 4.8).

This was the first study to explore the genetic diversity and structure of *P. strobus* in Kentucky. These three populations (KY1, KY2, and KY3), as well as the Tennessee population on the western side of the Tennessee River Valley (TN4) and the West Virginia population (WV1), are all isolated from the main distribution of *P. strobus* occurring along the main Appalachian Mountain chain. These five isolated populations are also quite small in terms of surrounding conspecific density (Figure 4.2b). No records

exist to indicate whether these are relict glacial populations or are remnants of a once contiguous *P. strobus* corridor that was present during postglacial recolonization. In many cases, isolated populations are of high conservation interest, because they may harbor unique alleles, potentially available for ongoing or future evolutionary change (Fady et al. 2016). However, none of these populations except for one (WV1) harbored a private allele. In fact, genetic diversity was fairly comparable to levels in the other eighteen populations. Two of the populations (KY3 and WV1) were highly differentiated, but the other three showed high relatedness with the rest of the Southern Appalachian populations (Table 4.6). Less than half of these geographically distinct populations were genetically very distinct, suggesting their current isolation may be a recent phenomenon (Hewitt 1996, Provan and Bennett 2008).

Categorization by Level III ecoregions (U.S. Environmental Protection Agency 2013) best explained genetic differentiation in the Southern Appalachians, accounting for the most variation among *P. strobus* populations and between groups (Table 4.7c, Figure 4.3). Successive uplift events formed the ancient Appalachian Mountain chain starting >1 billion years ago. This created geologically varied ecoregions in the Southern Appalachians, characterized by their distinct topography, soils, and climatic conditions. Despite weak overall *P. strobus* differentiation in the Southern Appalachians, this finding may best explain extant population structure and be an indication of possible local adaptation to specific ecological zones.

4.4.3 Climate suitability modeling

Pollen macrofossil evidence indicates *P. strobus* refugia existed south of the Appalachian Mountains in northwestern Georgia during the LGM (Jackson 2000). As the glaciers receded, *P. strobus* first recolonized cooler, higher elevations and continued range expansion northeastward along the Appalachian Mountain chain, arriving in Virginia's Shenandoah Valley (where populations VA3 and VA4 are located) ~13,000 years ago. According to our climatic suitability envelope during the LGM, suitable habitats existed in the area surrounding the southern extent of the Appalachian Mountain chain, including areas west in what is present-day Tennessee, USA. In addition to populating the cooler, higher elevations of the Appalachian Mountains after glacial retreat, migrants from these populations may also have moved west into higher elevation sites opposite the Ridge and Valley ecoregion along the Cumberland Plateau. This may explain the presence of isolated *P. strobus* patches in this area today (e.g. populations TN4, KY1, KY2, KY3, and WV1). By the mid-Holocene, *P. strobus* had recolonized an area of North America similar in extent to its current distribution (Davis 1983), but our climate suitability envelope during this time suggests it may have existed in patchy, sparse populations. This was especially true in the Southern Appalachians and the area west of the mountain chain, where climatic suitability was low to moderate ~6,000 years ago. Since then, the climatic suitability improved within this region, potentially facilitating population expansions. We propose that isolated stands of *P. strobus* currently in Central Tennessee and Kentucky are not relict populations. Their genetic similarity to our other Southern Appalachian populations, as well as the presence of suitable climatic

habitat in this area over time, suggests there has been high historic gametic connectivity through what may have been a relatively contiguous corridor at one time.

4.4.4 Summary and conclusions

Species that experienced glacial range contraction and subsequent interglacial range expansion are predicted to harbor greater levels of genetic diversity where their glacial refugia were located (Hewitt 1996, 2000). Genetic diversity is critical for adaptation, so populations in these areas may hold important conservation potential (Hampe and Petit 2005). *Pinus strobus* is thought to have survived the LGM in refugia located in the Southern Appalachians (Davis 1983, Jackson 2000), the same part of its range expected to contract within the next 100 years as global temperatures rapidly increase (Joyce and Rehfeldt 2013). Recently emerged pests and pathogens are also causing dieback and mortality, exacerbating the risk of population reductions in the southern extent (Asaro et al. 2018, Schulz et al. 2018, Whitney et al. 2018). Our snapshot of the standing genetic variability and population structure in the Southern Appalachians provides little evidence that this region is a hotspot of exceptionally high extant and/or ancestral genetic diversity. Instead, our genetic analyses suggest that *P. strobus* throughout North America are essentially a single heterogeneous, nearly admixed species.

Natural and anthropogenic phenomena, as well as our methodological choices, may help to contextualize our findings: (1) the mating system of *P. strobus* and the lack of barriers on the landscape facilitate long-distance gametic exchange. Although the Southern Appalachian Mountains served as a discontinuity driving vicariance in several

species during glaciation (Soltis et al. 2006), they have been considered a weak barrier to gene flow for wind-dispersed trees, like *P. strobus* (Nadeau et al. 2015). Consistent south-to-north pollen and seed movement over many generations since the LGM may explain the shallow population structure and consistently high genetic diversity throughout the its range. (2) Anthropogenic activities within the last century may have also affected our estimates of genetic diversity and population structure. As part of the "New Deal" in the 1930s through the 1950s, the Tennessee Valley Authority and the Civilian Conservation Corps cleared significant swaths of land throughout the Southern Appalachians and in many cases replanted millions of seedlings where reforestation was needed (e.g. Vimmerstedt 1962). Some records indicate these replanted seedlings were locally sourced, but it is impossible to know the origin and extent of all planted *P. strobus* stands in this region. Occurring only ~80 years ago (less than a single *P. strobus* generation), there is a possibility that such activities by the federal government may have dampened signal(s) of population structure and/or altered detectable levels of genetic diversity within this region. (3) Although the microsatellite markers we chose for genetic analyses have been used extensively, they may not have allowed for optimal temporal resolution or allelic breadth to detect substantial genetic differentiation among groups. Nadeau et al. (2015) reported the presence of substructure ($K = 4$) within their "southern group" (comprising of eight sample sites), based on *P. strobus* genotypes from ~150 single nucleotide polymorphisms (SNPs). Future investigations into the population structure of *P. strobus* in the Southern Appalachians should consider using SNPs or developing additional microsatellite markers for a finer resolution.

Estimating variation in selectively neutral markers is an important barometer for the adaptive potential in trees (Jump et al. 2009). With its great capacity for gene flow, *P. strobus* populations in the Southern Appalachians and across its entire range, including the highly-isolated populations tested herein, appear to be highly diverse and highly admixed. Despite the lack of clear population structure found in our study, provenance tests have shown *P. strobus* to have substantial clinal variation in certain phenotypic traits according to latitude and climate (e.g. Joyce and Rehfeldt 2013). This plasticity, as well as the high genetic variation and high genetic mobility of *P. strobus* bode well for its adaptability (Hamrick 2004). Although populations in the Southern Appalachians were not uniquely high in genetic diversity as we hypothesized, they remain an important part of the range deserving of conservation priority.

4.5 Acknowledgements

We are grateful to those who assisted with *P. strobus* sampling: Chris Asaro (USDA Forest Service), Brittany Barnes (University of Georgia), Megan Bartz (USDA Forest Service), Amanda Block (USDA Forest Service), William Caudill (USDA Forest Service), Lori Chamberlain (Virginia Department of Forestry), Beth Christensen (USDA Forest Service), Kara Costanza (University of New Brunswick), Don Grosman (Arborjet), TJ Hall (Pennsylvania DCNR), Bob Heyd (USDA Forest Service), Erica Roberts (USDA Forest Service), Derek Robertson (USDA Forest Service) Laura Vincent (USDA Forest Service), and Joe White (USDA Forest Service). This research was supported in part by the USDA Forest Service, Southern Research Station (13-CA-

11330129-056 and 16-CS-11330129-045), Southern Region-Forest Health Protection, and the University of Georgia (2015) Innovative and Interdisciplinary Research Grant.

4.6 References

Abrams, M.D., 2001. Eastern white pine versatility in the presettlement forest. Bioscience 51, 967-979.

Abrams, M.D., Orwig, D.A., Demeo, T.E., 1995. Dendroecological analysis of successional dynamics for a presettlement-origin white-pine mixed-oak forest in the Southern Appalachians, USA. Journal of Ecology 83, 123-133.

Aitken, S.N., Yeaman, S., Holliday, J.A., Wang, T.L., Curtis-McLane, S., 2008. Adaptation, migration or extirpation: climate change outcomes for tree populations. Evolutionary Applications 1, 95-111.

Alfaro, R.I., Fady, B., Vendramin, G.G., Dawson, I.K., Fleming, R.A., Saenz-Romero, C., Lindig-Cisneros, R.A., Murdock, T., Vinceti, B., Navarro, C.M., Skroppa, T., Baldinelli, G., El-Kassaby, Y.A., Loo, J., 2014. The role of forest genetic resources in responding to biotic and abiotic factors in the context of anthropogenic climate change. Forest Ecology and Management 333, 76-87.

Asaro, C., Chamberlin, L.A., Rose, J.A., Mooneyham, K., Rose, A.K., 2018. Mortality of eastern white pine (*Pinus strobus* L.) in association with a novel scale insect-pathogen complex in Virginia and West Virginia. Forest Ecology and Management 423, 37-48.

Barrett, R.D.H., Schluter, D., 2008. Adaptation from standing genetic variation. Trends in Ecology & Evolution 23, 38-44.

Beaulieu, J., Simon, J.P., 1994. Inheritance and linkage relationships of allozymes in *Pinus strobus* L. Silvae Genetica 43, 253-261.

Beaulieu, J., Simon, J.P., 1995. Mating system in natural populations of eastern white pine in Quebec. Canadian Journal of Forest Research-Revue Canadienne De Recherche Forestiere 25, 1697-1703.

Brownstein, M.J., Carpten, J.D., Smith, J.R., 1996. Modulation of non-templated nucleotide addition by taq DNA polymerase: Primer modifications that facilitate genotyping. Biotechniques 20, 1004-1010.

Buchert, G.P., Rajora, O.P., Hood, J.V., Dancik, B.P., 1997. Effects of harvesting on genetic diversity in old growth eastern white pine in Ontario, Canada. Conservation Biology 11, 747-758.

Campagne, P., Smouse, P.E., Varouchas, G., Silvain, J.F., Leru, B., 2012. Comparing the van Oosterhout and Chybicki-Burczyk methods of estimating null allele frequencies for inbred populations. Molecular Ecology Resources 12, 975-982.

Chapuis, M.P., Estoup, A., 2007. Microsatellite null alleles and estimation of population differentiation. Molecular Biology and Evolution 24, 621-631.

Chhatre, V.E., Rajora, O.P., 2014. Genetic divergence and signatures of natural selection in marginal populations of a keystone, long-lived conifer, eastern white pine (*Pinus strobus*) from northern Ontario. Plos One 9.

Chybicki, I.J., Burczyk, J., 2009. Simultaneous estimation of null alleles and inbreeding coefficients. Journal of Heredity 100, 106-113.

Cornuet, J.M., Luikart, G., 1996. Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. Genetics 144, 2001-2014.

Costanza, K.K.L., Whitney, T.D., McIntire, C.D., Livingston, W.H., Gandhi, K.J.K., 2018. A synthesis of emerging health issues of eastern white pine (*Pinus strobus*) in eastern North America. Forest Ecology and Management 423, 3-17.

Davis, M.B., 1983. Quaternary history of deciduous forests of eastern North America and Europe. Annals of the Missouri Botanical Garden 70, 550-563.

Earl, D.A., Vonholdt, B.M., 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources 4, 359-361.

Echt, C.S., MayMarquardt, P., Hseih, M., Zahorchak, R., 1996. Characterization of microsatellite markers in eastern white pine. Genome 39, 1102-1108.

Elith, J., Phillips, S.J., Hastie, T., Dudik, M., Chee, Y.E., Yates, C.J., 2011. A statistical explanation of MaxEnt for ecologists. Diversity and Distributions 17, 43-57.

Epperson, B.K., Chung, M.G., 2001. Spatial genetic structure of allozyme polymorphisms within populations of *Pinus strobu*s (Pinaceae). American Journal of Botany 88, 1006-1010.

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

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

Fady, B., Aravanopoulos, F.A., Alizoti, P., Matyas, C., von Wuhlisch, G., Westergren, M., Belletti, P., Cvjetkovic, B., Ducci, F., Huber, G., Kelleher, C.T., Khaldi, A., Kharrat, M.B.D., Kraigher, H., Kramer, K., Muhlethaler, U., Peric, S., Perry, A., Rousi, M., Sbay, H., Stojnic, S., Tijardovic, M., Tsvetkov, I., Varela, M.C., Vendramin, G.G., Zlatanov, T., 2016. Evolution-based approach needed for the conservation and silviculture of peripheral forest tree populations. Forest Ecology and Management 375, 66-75.

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

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

Fowler, D.P., 1965. Natural self-fertilization in three jack pines and its implications in seed orchard management. Forest Science 11, 55-58.

Goudet, J., 2005. HIERFSTAT, a package for R to compute and test hierarchical F-statistics. Molecular Ecology Notes 5, 184-186.

Hampe, A., Petit, R.J., 2005. Conserving biodiversity under climate change: the rear edge matters. Ecology Letters 8, 461-467.

Hamrick, J.L., 2004. Response of forest trees to global environmental changes. Forest Ecology and Management 197, 323-335.

Hamrick, J.L., Nason, J.D., 1996. Consequences of dispersal in plants. Population dynamics in ecological space and time, 203-236.

Hepp, T.E., Vimmerstedt, J.P., Smalley, G.W., McNab, W.H., 2015. Estimating yields of unthinned eastern white pine plantations from current stocking in the Southern Appalachians. Forest Science 61, 114-122.

Hewitt, G., 2000. The genetic legacy of the Quaternary ice ages. Nature 405, 907-913.

Hewitt, G.M., 1999. Post-glacial re-colonization of European biota. Biological Journal of the Linnean Society 68, 87-112.

Hewitt, G.M., 2004. Genetic consequences of climatic oscillations in the Quaternary. Philosophical Transactions of the Royal Society of London Series B-Biological Sciences 359, 183-195.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G., Jarvis, A., 2005. Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology 25, 1965-1978.

Hijmans, R.J., van Etten, J., Cheng, J., Mattiuzzi, M., Sumner, M., Greenberg, J.A., Lamigueiro, O.P., Bevan, A., Racine, E.B., Shortridge, A., Hijmans, M.R.J., 2017. Package 'raster'. R package.

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

Hufford, K.M., Hamrick, J.L., 2003. Viability selection at three early life stages of the tropical tree, *Platypodium elegans* (Fabaceae, Papilionoideae). Evolution 57, 518-526.

Jackson, S.T., Webb, R.S., Anderson, K.H., Overpeck, J.T., Webb, T., Williams, J.W., Hansen, B.C.S., 2000. Vegetation and environment in Eastern North America during the Last Glacial Maximum. Quaternary Science Reviews 19, 489-508.

Janes, J.K., Miller, J.M., Dupuis, J.R., Malenfant, R.M., Gorrell, J.C., Cullingham, C.I., Andrew, R.L., 2017. The K=2 conundrum. Molecular Ecology 26, 3594-3602.

Johnson, L.P.V., 1945. Reduced vigour, chlorophyll deficiency, and other effects of self-fertilization in *Pinus*. Canadian Journal of Research 23, 145-149.

Joyce, D.G., Rehfeldt, G.E., 2013. Climatic niche, ecological genetics, and impact of climate change on eastern white pine (*Pinus strobus* L.): Guidelines for land managers. Forest Ecology and Management 295, 173-192.

Jump, A.S., Marchant, R., Penuelas, J., 2009. Environmental change and the option value of genetic diversity. Trends in Plant Science 14, 51-58.

Kopelman, N.M., Mayzel, J., Jakobsson, M., Rosenberg, N.A., Mayrose, I., 2015. Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. Molecular Ecology Resources 15, 1179-1191.

Kuparinen, A., Savolainen, O., Schurr, F.M., 2010. Increased mortality can promote evolutionary adaptation of forest trees to climate change. Forest Ecology and Management 259, 1003-1008.

Lovett, G.M., Canham, C.D., Arthur, M.A., Weathers, K.C., Fitzhugh, R.D., 2006. Forest ecosystem responses to exotic pests and pathogens in eastern North America. Bioscience 56, 395-405.

Mandak, B., Hadincova, V., Mahelka, V., Wildova, R., 2013. European invasion of North American *Pinus strobus* at large and fine scales: High genetic diversity and fine-scale genetic clustering over time in the adventive range. Plos One 8.

Marquardt, P.E., Echt, C.S., Epperson, B.K., Pubanz, D.M., 2007. Genetic structure, diversity, and inbreeding of eastern white pine under different management

conditions. Canadian Journal of Forest Research-Revue Canadienne De Recherche Forestiere 37, 2652-2662.

Marquardt, P.E., Epperson, B.K., 2004. Spatial and population genetic structure of microsatellites in white pine. Molecular Ecology 13, 3305-3315.

Mehes, M., Nkongolo, K.K., Michael, P., 2009. Assessing genetic diversity and structure of fragmented populations of eastern white pine (*Pinus strobus*) and western white pine (*P. monticola*) for conservation management. Journal of Plant Ecology 2, 143-151.

Munzbergova, Z., Hadincova, V., Wild, J., Herben, T., Maresova, J., 2010. Spatial and temporal variation in dispersal pattern of an invasive pine. Biological Invasions 12, 2471-2486.

Nadeau, S., Godbout, J., Lamothe, M., Gros-Louis, M.-C., Isabel, N., Ritland, K., 2015. Contrasting patterns of genetic diversity across the ranges of *Pinus monticola* and *P. strobus*: A comparison between eastern and western North American postglacial colonization histories. American Journal of Botany 102, 1342-1355.

Nei, M., 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 89, 583-590.

Nijensohn, S.E., Schaberg, P.G., Hawley, G.J., DeHayes, D.H., 2005. Genetic subpopulation structuring and its implications in a mature eastern white pine stand. Canadian Journal of Forest Research-Revue Canadienne De Recherche Forestiere 35, 1041-1052.

Peakall, R., Smouse, P.E., 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Molecular Ecology Notes 6, 288-295.

Peakall, R., Smouse, P.E., 2012. GENALEX 6.5: genetic analysis in Excel. Population genetic software for teaching and research: an update. Bioinformatics 28, 2537-2539.

Petit, R.J., Aguinagalde, I., de Beaulieu, J.L., Bittkau, C., Brewer, S., Cheddadi, R., Ennos, R., Fineschi, S., Grivet, D., Lascoux, M., Mohanty, A., Muller-Starck, G.M., Demesure-Musch, B., Palme, A., Martin, J.P., Rendell, S., Vendramin, G.G., 2003. Glacial refugia: Hotspots but not melting pots of genetic diversity. Science 300, 1563-1565.

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

Piry, S., Luikart, G., Cornuet, J.M., 1999. BOTTLENECK: A computer program for detecting recent reductions in the effective population size using allele frequency data. Journal of Heredity 90, 502-503.

Potter, K.M., Jetton, R.M., Bower, A., Jacobs, D.F., Man, G., Hipkins, V.D., Westwood, M., 2017. Banking on the future: progress, challenges and opportunities for the genetic conservation of forest trees. New Forests 48, 153-180.

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

Provan, J., Bennett, K.D., 2008. Phylogeographic insights into cryptic glacial refugia. Trends in Ecology & Evolution 23, 564-571.

R Core Team, 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rajora, O.P., DeVerno, L., Mosseler, A., Innes, D.J., 1998. Genetic diversity and population structure of disjunct Newfoundland and central Ontario populations of eastern white pine (*Pinus strobus*). Canadian Journal of Botany-Revue Canadienne De Botanique 76, 500-508.

Rajora, O.P., Rahman, M.H., Buchert, G.P., Dancik, B.P., 2000. Microsatellite DNA analysis of genetic effects of harvesting in old-growth eastern white pine (*Pinus strobus*) in Ontario, Canada. Molecular Ecology 9, 339-348.

Rogers, L.L., Lindquist, E.L., 1992. Supercanopy white pine and wildlife. In: Stine, R.A. (Ed.), White Pine Symposium, Duluth, Minnesota, USA, pp. 39-43.

Rosenberg, N.A., 2004. DISTRUCT: a program for the graphical display of population structure. Molecular Ecology Notes 4, 137-138.

Savolainen, O., Bokma, F., Garcia-Gil, R., Komulainen, P., Repo, T., 2004. Genetic variation in cessation of growth and frost hardiness and consequences for adaptation of *Pinus sylvestris* to climatic changes. Forest Ecology and Management 197, 79-89.

Savolainen, O., Pyhajarvi, T., Knurr, T., 2007. Gene flow and local adaptation in trees. Annual Review of Ecology Evolution and Systematics 38, 595-619.

Schaberg, P.G., DeHayes, D.H., Hawley, G.J., Nijensohn, S.E., 2008. Anthropogenic alterations of genetic diversity within tree populations: Implications for forest ecosystem resilience. Forest Ecology and Management 256, 855-862.

Schulz, A.N., Mech, A.M., Cram, M.M., Asaro, C., Coyle, D.R., Lucardi, R.D., Lucas, S., Gandhi, K.J.K., 2018. Association of *Caliciopsis pinea* Peck and *Matsucoccus macrocicatrices* Richards with eastern white pine (*Pinus strobus* L.) seedling dieback. Forest Ecology and Management 423, 70-83.

Soltis, D.E., Gitzendanner, M.A., Strenge, D.D., Soltis, P.S., 1997. Chloroplast DNA intraspecific phylogeography of plants from the Pacific Northwest of North America. Plant Systematics and Evolution 206, 353-373.

Soltis, D.E., Morris, A.B., McLachlan, J.S., Manos, P.S., Soltis, P.S., 2006. Comparative phylogeography of unglaciated eastern North America. Molecular Ecology 15, 4261-4293.

Strauss, S.H., 1986. Heterosis at allozyme loci under inbreeding and crossbreeding in *Pinus attenuata*. Genetics 113, 115-134.

Talarico, L., Babik, W., Marta, S., Mattoccia, M., 2019. Genetic drift shaped MHC IIB diversity of an endangered anuran species within the Italian glacial refugium. Journal of Zoology 307, 61-70.

U.S. Environmental Protection Agency, 2013. Level III ecoregions of the continental United States. U.S. EPA - National Health and Environmental Effects Research Laboratory, Corvallis, Oregon, USA.

Varela, S., Lima-Ribeiro, M.S., Terribile, L.C., 2015. A short guide to the climatic variables of the last glacial maximum for biogeographers. Plos One 10.

Vimmerstedt, J.P., 1962. Southern Appalachian white pine plantations site, volume, and yield. USDA Forest Service, Southeastern Forest Experimental Station, p. 15.

Walter, R., Epperson, B.K., 2001. Geographic pattern of genetic variation in *Pinus resinosa*: area of greatest diversity is not the origin of postglacial populations. Molecular Ecology 10, 103-111.

Walter, R., Epperson, B.K., 2004. Microsatellite analysis of spatial structure among seedlings in populations of *Pinus strobus* (Pinaceae). American Journal of Botany 91, 549-557.

Wendel, G.W., Smith, H.C., 1990. *Pinus strobus* L. eastern white pine. In: Burns, R.M., Honkala, B.H. (Eds.), Silvics of North America. USDA Forest Service, Washington DC, pp. 476-488.

Whitney, T.D., Cram, M.M., Barnes, B.F., Yao, J.M., Lucardi, R.D., Gandhi, K.J.K., 2018. Tree-level distribution of a novel insect-pathogen complex and its potential contribution to eastern white pine dieback. Forest Ecology and Management 423, 49-58.

Wielstra, B., Crnobrnja-Isailovic, J., Litvinchuk, S.N., Reijnen, B.T., Skidmore, A.K., Sotiropoulos, K., Toxopeus, A.G., Tzankov, N., Vukov, T., Arntzen, J.W., 2013. Tracing glacial refugia of *Triturus* newts based on mitochondrial DNA phylogeography and species distribution modeling. Frontiers in Zoology 10.

Wilson, R.W., McQuilkin, W.E., 1963. Silvical characteristics of eastern white pine (*Pinus strobus*). USDA Forest Service Northeasten Forest Experimental Station Research Paper 13, 1-28.

Zinck, J.W.R., Rajora, O.P., 2016. Post-glacial phylogeography and evolution of a wide-ranging highly-exploited keystone forest tree, eastern white pine (*Pinus strobus*) in North America: single refugium, multiple routes. Bmc Evolutionary Biology 16.

Table 4.1. Population information for microsatellite analysis of *M. macrocicatrices*, including location data and estimates of genetic diversity indices.

Site ID	Site Name	County, state	Ecoregion[a]	Latitude	Longitude	Elev. (m)	N	A_N	A_E	A_R	A_P	A_{LC}	F_{IS}	H_O	H_E
GA1	Mill Creek	Murray, GA	BR	34.8723	-84.7234	291	20	4.6	2.33	3.99	0.1	1.1	0.096	0.504	0.573
GA2	Rock Creek	Gilmer, GA	BR	34.7821	-84.3334	596	20	4.6	2.13	3.95	0.2	0.9	0.070	0.482	0.543
GA3	Canada Creek	Union, GA	BR	34.6803	-84.0425	845	15	4.4	2.10	3.94	0.2	1.3	0.117	0.424	0.512
GA4	Boggs Creek	Lumpkin, GA	BR	34.6940	-83.8914	554	15	4.7	2.12	4.24	0.2	1.4	0.139	0.481	0.591
GA5	Raper Creek	Habersham, GA	BR	34.7442	-83.5721	510	15	4.2	2.11	3.89	0.0	1.0	0.194	0.457	0.578
GA6	Panther Creek	Habersham, GA	BR	34.6995	-83.4199	486	18	5.9	2.57	5.09	0.4	2.1	0.226	0.478	0.645
SC1	Chattooga River	Oconee, SC	BR	34.9154	-83.1630	518	20	5.0	2.53	4.24	0.1	1.3	0.208	0.501	0.623
NC1	Muskrat Valley	Macon, NC	BR	35.1380	-83.5038	808	19	4.9	2.70	4.19	0.0	1.6	0.176	0.471	0.595
NC2	Long Branch Trail	Transylvania, NC	BR	35.2729	-82.8393	926	20	5.9	2.49	4.65	0.2	1.6	0.270	0.436	0.620
NC3	Little Buck Creek	McDowell, NC	BR	35.7461	-82.0975	551	20	4.8	2.33	3.94	0.0	1.6	0.099	0.469	0.548
NC4	Wilson Creek	Avery, NC	BR	36.0696	-81.7990	890	19	4.8	2.67	4.23	0.0	1.1	0.160	0.477	0.588
TN1	Tellico Plains	Monroe, TN	BR	35.3375	-84.1488	625	20	4.9	2.87	4.40	0.0	1.7	0.109	0.532	0.605
TN2	Green Corner	Cocke, TN	BR	35.7986	-83.0899	688	19	5.0	2.47	4.31	0.0	1.7	0.204	0.439	0.581
TN3	Nolichucky River	Unicoi, TN	BR	36.1005	-82.4328	543	20	5.2	2.86	4.34	0.0	1.7	0.112	0.518	0.613
TN4	Norris Dam State Park	Campbell, TN	RV	36.2409	-84.1077	411	20	4.9	2.71	4.12	0.0	1.6	0.199	0.446	0.594
VA1	Pugh Mountian	Smyth, VA	BR	36.7871	-81.4679	953	19	5.6	2.82	4.71	0.2	1.5	0.114	0.578	0.668
VA2	Price Ridge	Bland, VA	RV	37.1551	-80.9561	792	18	4.5	2.94	4.12	0.0	1.3	0.189	0.479	0.584
VA3	Falling Springs	Alleghany, VA	RV	37.8771	-79.9884	451	19	5.7	3.04	4.76	0.6	1.6	0.296	0.421	0.632
VA4	Deerfield	Augusta, VA	RV	38.2113	-79.3297	634	20	4.7	2.66	4.11	0.1	1.3	0.353	0.325	0.543
WV1	Horseshoe Run	Tucker, WV	RV	39.1878	-79.6211	676	19	3.8	2.04	3.47	0.1	0.8	0.124	0.417	0.496
KY1	Laurel Road	Laurel, KY	SWA	37.0085	-84.2697	381	19	4.9	2.23	4.21	0.0	1.5	0.254	0.407	0.586
KY2	Sheltowee Trace Connector	Menifee, KY	WAP	37.8256	-83.6284	228	20	4.6	2.22	3.88	0.0	1.6	0.194	0.420	0.544
KY3	Pretty Ridge	Menifee, KY	WAP	38.0242	-83.4430	340	18	3.6	2.31	3.36	0.0	0.6	0.079	0.555	0.589
Mean: Southern Appalachians							*18.8*	*4.8*	*2.49*	*4.18*	*0.10*	*1.4*	*0.173*	*0.466*	*0.585*
NH	Mast Yard State Forest	Merrimack, NH	NECZ	43.2398	-71.6529	12	20	5.7	2.86	4.63	0.3	1.3	0.046	0.614	0.664
MI	Poverty Perch	Crawford, MI	NLF	44.7569	-84.8366	360	14	4.6	2.67	4.39	0.1	0.7	0.073	0.597	0.640
Mean: All Populations							*18.6*	*4.9*	*2.51*	*4.21*	*0.11*	*1.4*	*0.164*	*0.477*	*0.590*

[a] U.S. Environmental Protection Agency, 2013, Level III ecoregions: BR = Blue Ridge, NECZ = Northeastern Coastal Zone, NLF = Northern Lakes and Forests, RV = Ridge and Valley, SWA = Southwestern Appalachians, WAP = Western Allegheny Plateau.

A_N: mean number of alleles; A_E: effective number of alleles; A_R: rarefied allelic richness; A_P: mean frequency of private alleles; A_{LC}: mean number of locally common alleles ($\geq$5%) occurring in $\leq$50% of populations; F_{IS}: inbreeding coefficients; H_O: observed heterozygosity; H_E: expected heterozygosity.

Table 4.2. Microsatellite primer sequences (Echt et al. 1996) and assigned fluorophore used in this study.

Locus[a]	Sequence (5' – 3')	Repeat motif	Allele size range (bp)
RPS1b^{V}	F: GCCCACTATTCAAGATGTCA	$(AC)_{10}$	193-213
	R: GATGTTAGCAGAAACATGAGG		
RPS2^{P}	F: CATGGTGTTGGTCATTGTTCCA	$(AC)_{15}$	155-177
	R: TGGAGGCTATCACGTATGCACC		
RPS6^{N}	F: TTTTCTAATCAGTGTGCGCTACA	$(AC)_{14}$	159-187
	R: CACCGCTGCCCTATTTTACA		
RPS12^{F}	F: TCAATGTGGAGATGGTGATT	$(AC)_{17}$	145-199
	R: ACTTCTGACCTAACCAGAAACC		
RPS20^{N}	F: ACTTCCCCACAGGTTAACACA	$(AC)_{18}(AT)_{6}$	120-177
	R: AACAAGATAGGCGGGATTCA		
RPS25b*	F: CACATATGGCAGAACACACA	$(AC)_{17}AG\ (AT)_{9}$	120-177
	R: GATCGTCGCACTATCGAAC		
RPS34b^{F}	F: CAGTGTTCTCTTATCACAGCG	$(AC)_{14}$	140-170
	R: GCACTATAATGAAATAGCGCA		
RPS39^{F}	F: GCCAGCTCCAACCAGAATC	$(AC)_{17}$	154-184
	R: GGCTCGCTGACCCAATAA		
RPS50^{P}	F: CCCAGAAATCTGTTTTAGAGC	$(AC)_{17}$	157-187
	R: ACACATGAAATGTCAGAATGC		
RPS60*	F: ACGATAATGGCGGTGAGAACAA	$(AC)_{19}(AT)_{7}$	199-280
	R: CCACCTGTCCTTCGTACATCCA		
RPS84^{P}	F: CCTTTGGTCATTGTATTTTTGGAC	$(CT)_{10}(AC)_{11}$	145-169
	R: CTTCCTTTTCCTTCTTGCTCCAC		
RPS127^{V}	F: ACTTCCTCCAAGTTACTATTGTCA	$(AC)_{10}(AT)_{5}$	191-199
	R: CCTTGTCTTCTAAAAAACACTTTT		

[a] Fluorophore used to label primers: F = FAM, V = VIC, N = NED, P = PET
[*]Removed from analyses.

Table 4.3. List of biolayers (worldclim.org) used in MAXENT models.

	Name	Variable definition
BIO1	Annual mean temperature	
BIO2	Mean monthly temperature range	
BIO3	Isothermality	(BIO2/BIO7)*(100)
BIO4	Temperature seasonality	(Standard deviation*100)
BIO5	Max temperature warmest month	
BIO6	Min temperature coldest month	
BIO7	Temperature annual range	(BIO5-BIO6)
BIO8	Mean temperature wettest quarter	
BIO9	Mean temperature driest quarter	
BIO10	Mean temperature warmest quarter	
BIO11	Mean temperature coldest quarter	
BIO12	Annual precipitation	
BIO13	Precipitation wettest month	
BIO14	Precipitation driest month	
BIO15	Precipitation seasonality	Coefficient of variation
BIO16	Precipitation wettest quarter	
BIO17	Precipitation driest quarter	
BIO18	Precipitation warmest quarter	
BIO19	Precipitation coldest quarter	

Table 4.4. Summary statistics of genetic diversity, *F*-statistics, and null allele frequencies for each of the 10 microsatellite loci used in the study (developed by Echt et al. 1996).

Locus	A	A_R	H_O	H_E	F_{ST}	F_{IS}	Null
RPS1b	8	3.33	0.296	0.359	0.031	0.112	0.033
RPS2	11	3.82	0.489	0.599	0.030	0.111	0.055
RPS6	12	4.58	0.529	0.678	0.064	0.277	0.114
RPS12	26	9.31	0.686	0.874	0.050	0.399	0.162
RPS20	19	5.94	0.587	0.735	0.060	0.144	0.056
RPS34b	12	3.86	0.464	0.618	0.040	0.429	0.129
RPS39	12	3.02	0.395	0.507	0.067	0.197	0.064
RPS50	16	7.29	0.681	0.815	0.044	0.085	0.046
RPS84	7	2.72	0.331	0.396	0.045	0.116	0.042
RPS127	5	3.02	0.459	0.614	0.160	0.478	0.102
Mean	*12.8*	*4.69*	*0.492*	*0.620*	*0.060*	*0.187*	*0.080*

A: total number of alleles; A_R: rarefied allelic richness; H_O: observed heterozygosity; H_E: expected heterozygosity; F_{ST}: fixation index; F_{IS}: inbreeding coefficient.

Table 4.5. Results from BOTTLENECK analysis (Luikart and Cornuet 1996). Asterisks indicate a significant result ($P < 0.05$).

Population	N	A_N	H_E	SMM		TPM	
				P (1t)	P (2t)	P (1t)	P (2t)
GA1	39.6	4.6	0.51	0.99	0.02*	1.00	0.01*
GA2	39.2	4.6	0.42	1.00	0.00*	1.00	0.00*
GA3	29.6	4.4	0.42	1.00	0.01*	1.00	0.00*
GA4	28.8	4.7	0.48	1.00	0.00*	1.00	0.00*
GA5	28.4	4.2	0.49	1.00	0.01*	1.00	0.01*
GA6	34.6	5.9	0.54	1.00	0.01*	1.00	0.01*
SC1	38.4	5	0.53	0.98	0.08	1.00	0.01*
NC1	37.4	4.9	0.53	0.90	0.23	0.95	0.13
NC2	39	5.9	0.53	1.00	0.00*	1.00	0.00*
NC3	39.6	4.8	0.48	0.99	0.03*	1.00	0.01*
NC4	37.6	4.8	0.54	0.96	0.11	0.98	0.08
TN1	38.6	4.9	0.56	0.90	0.23	0.96	0.11
TN2	38	5	0.55	0.98	0.08	0.99	0.02*
TN3	38.8	5.2	0.56	0.88	0.28	0.93	0.16
TN4	40	4.9	0.57	0.78	0.49	0.90	0.23
VA1	35.4	5.6	0.58	0.99	0.03*	1.00	0.01*
VA2	35	4.5	0.53	0.54	1.00	0.58	0.92
VA3	37.8	5.7	0.60	1.00	0.01*	1.00	0.01*
VA4	40	4.7	0.52	0.88	0.28	0.93	0.16
WV1	37	3.8	0.41	0.99	0.02*	1.00	0.01*
KY1	36.8	4.9	0.52	1.00	0.01*	1.00	0.01*
KY2	39.8	4.6	0.49	0.92	0.19	0.99	0.02*
KY3	35	3.6	0.53	0.38	0.77	0.58	0.92
NH	39	5.7	0.57	1.00	0.01*	1.00	0.01*
MI	26.4	4.6	0.55	0.86	0.32	0.92	0.19

Table 4.6. Population pairwise F_{ST} values (25 populations, 466 individuals) calculated with the "exclusion null alleles" method from FREENA (Chapuis and Estoup 2007). All significant values ($P < 0.05$) are in bold.

	GA1	GA2	GA3	GA4	GA5	GA6	SC1	NC1	NC2	NC3	NC4	TN1	TN2	TN3	TN4	VA1	VA2	VA3	VA4	WV1	KY1	KY2	KY3	NH	MI
GA1																									
GA2	**0.113**																								
GA3	**0.077**	0.076																							
GA4	**0.042**	**0.105**	**0.030**																						
GA5	0.004	**0.107**	**0.032**	0.002																					
GA6	0.031	**0.089**	**0.041**	**0.029**	0.017																				
SC1	0.036	**0.113**	**0.059**	**0.060**	0.026	0.015																			
NC1	**0.058**	**0.086**	**0.041**	0.033	0.038	0.023	**0.046**																		
NC2	**0.043**	**0.091**	**0.032**	0.019	0.014	0.011	**0.028**	0.022																	
NC3	**0.027**	**0.111**	**0.044**	0.021	0.010	**0.045**	**0.059**	**0.052**	**0.040**																
NC4	**0.043**	**0.079**	**0.042**	**0.035**	**0.035**	**0.016**	**0.031**	0.001	0.004	**0.039**															
TN1	-0.009	**0.094**	**0.052**	**0.029**	-0.001	0.008	0.023	**0.030**	0.018	**0.018**	**0.017**														
TN2	**0.044**	**0.072**	0.024	0.024	0.007	0.011	**0.030**	**0.030**	0.013	**0.044**	0.021	0.010													
TN3	**0.040**	**0.124**	**0.116**	**0.073**	**0.051**	**0.048**	**0.043**	**0.055**	**0.061**	**0.077**	0.037	**0.040**	**0.061**												
TN4	**0.042**	**0.119**	**0.105**	**0.075**	**0.046**	**0.036**	**0.030**	**0.051**	**0.045**	**0.061**	**0.031**	**0.026**	**0.046**	0.015											
VA1	**0.085**	**0.128**	**0.121**	**0.096**	**0.086**	**0.050**	**0.051**	**0.038**	**0.048**	**0.098**	0.017	**0.058**	**0.062**	**0.026**	0.002										
VA2	**0.044**	**0.106**	**0.087**	**0.059**	**0.053**	**0.017**	**0.030**	**0.042**	0.015	**0.077**	0.014	**0.028**	**0.048**	**0.034**	**0.032**	**0.027**									
VA3	0.041	**0.117**	**0.096**	**0.069**	**0.055**	**0.037**	**0.042**	**0.043**	**0.023**	**0.064**	**0.018**	**0.028**	**0.043**	**0.022**	**0.034**	**0.032**	0.019								
VA4	**0.048**	**0.100**	**0.057**	**0.039**	**0.037**	**0.021**	0.024	**0.027**	0.000	**0.052**	0.004	0.025	**0.031**	**0.048**	**0.037**	**0.034**	0.011	0.012							
WV1	**0.161**	**0.026**	**0.138**	**0.174**	**0.160**	**0.142**	**0.147**	**0.131**	**0.137**	**0.162**	0.114	**0.141**	**0.132**	**0.176**	**0.157**	**0.167**	**0.161**	**0.157**	**0.139**						
KY1	**0.056**	**0.081**	0.035	**0.035**	**0.032**	0.008	**0.035**	0.011	0.011	**0.059**	0.002	**0.023**	0.006	**0.048**	0.032	**0.029**	**0.029**	**0.039**	**0.018**	0.123					
KY2	0.010	**0.129**	**0.066**	**0.045**	0.002	**0.046**	**0.038**	**0.070**	**0.044**	**0.048**	**0.066**	0.018	**0.046**	**0.065**	**0.058**	**0.115**	**0.067**	**0.068**	**0.064**	**0.169**	**0.062**				
KY3	**0.106**	**0.262**	**0.214**	**0.148**	**0.107**	**0.152**	**0.115**	**0.157**	**0.114**	**0.152**	**0.136**	**0.095**	**0.134**	**0.122**	**0.119**	**0.152**	**0.139**	**0.094**	**0.126**	**0.282**	**0.163**	**0.104**			
NH	**0.036**	**0.081**	**0.056**	**0.054**	**0.038**	0.008	0.017	**0.031**	0.010	**0.070**	0.006	0.015	0.012	**0.026**	**0.032**	**0.028**	0.004	0.011	0.018	**0.128**	0.009	**0.052**	**0.117**		
MI	**0.067**	**0.113**	**0.047**	**0.042**	**0.039**	**0.036**	**0.064**	**0.029**	0.018	**0.052**	0.018	**0.039**	0.022	**0.054**	**0.059**	**0.058**	**0.048**	**0.035**	**0.036**	**0.165**	0.030	**0.062**	**0.127**	**0.018**	
Avg	*0.052*	*0.105*	*0.070*	*0.056*	*0.042*	*0.039*	*0.048*	*0.048*	*0.036*	*0.062*	*0.034*	*0.034*	*0.040*	*0.061*	*0.054*	*0.067*	*0.050*	*0.050*	*0.042*	*0.149*	*0.041*	*0.063*	*0.143*	*0.037*	*0.053*

Table 4.7. Analyses of Molecular Variance (AMOVA) of only individuals within the Southern Appalachians (23 populations, N = 432). Significant *F*-statistics (P < 0.05) are bold.

Test	Group Structure	Source of variation	Degrees of freedom	Sum of squares	Variance components	Percent variation	Fixation indices
A	None	Among populations	22	173	0.14	5.5	F_{ST} = **0.055**
		Within populations	841	2082	2.48	95.5	
		Total	863	2255	2.62		
B	West of vs. Within the Appalachian Mountain chain (2 groups[a])	Between groups	1	10	0.01	0.2	F_{CT} = 0.002
		Among populations	21	163	0.14	5.4	F_{SC} = **0.054**
		Within populations	841	2082	2.48	94.4	F_{ST} = **0.056**
		Total	863	2255	2.62		
C	By Ecoregion (4 groups[b])	Between groups	3	44	0.05	1.9	F_{CT} = **0.019**
		Among populations	19	129	0.11	4.3	F_{SC} = **0.044**
		Within populations	841	2082	2.48	93.8	F_{ST} = **0.062**
		Total	863	2255	2.64		

[a] West group: TN4, KY1, KY2, KY3 & WV1; Within group: the 18 remaining populations.
[b] Four U.S. Environmental Protection Agency, 2013, Level III ecoregions (see Table 4.1 for population groupings): Blue Ridge, Ridge and Valley, Southwestern Appalachians & Western Allegheny Plateau.

Table 4.8. Comparison of the genetic diversity estimates and *F*-statistics reported for *P. strobus* in studies utilizing microsatellite markers developed by Echt et al. (1996). Other smaller-scale studies using these markers, as well as studies using allozymes, were omitted.

Publication	Microsatellites[a]	Range	Populations	N	A_R	H_O	H_E	F_{IS}	F_{ST}
This study	**1b, 2, 6, 12, 34b, 39, 50, 84, 127**	USA[b]: GA, KY, MI, NC, NH, SC, TN, VA, WV	25	466	4.21	0.477	0.590	0.164	0.060
Mehes et al. 2009	**1b**, **2**, **12**, **20**, 25b, **50,** 118b	Canada[c]: NB, NL, NS, ON, PE, QC	10	300-400	*NA*	0.740	0.802	0.072	0.084
Mandak et al. 2013	**1b, 2, 12**, 25b, **34b, 39, 50, 84**, 118b, **127**	USA[b]: CT, MA, ME, MI, NH, NY, PA, VT	30	592	4.23	0.432	0.528	0.215	0.025
Chhatre & Rajora 2014	**1b, 2, 12, 20**, 25b, **34b, 39, 50**, 118b, 119, **127**	Canada[c]: ON	6	614	10.58	0.525	0.608	0.139	0.083
Zinck & Rajora 2016	**1b, 2, 6, 12, 20**, 25b, **34b, 39, 50**, 118b, 119, **127**	Canada[c]: NB, NL, ON, QC USA[b]: ME, MN, NC, NH, NY, PA, VA	33	1650	10.36	0.680	0.740	0.100	0.104

[a] "RPS" microsatellite markers from Echt et al. 1996. **Bold** denotes a marker also used in this study.
[b] GA: Georgia, KY: Kentucky, ME: Maine, MI: Michigan, MN: Minnesota, NC: North Carolina, NH: New Hampshire, NY: New York, PA: Pennsylvania, SC: South Carolina, TN: Tennessee, VA: Virginia, WV: West Virginia.
[c] NB: New Brunswick, NL: Newfoundland and Labrador, NS: Nova Scotia, ON: Ontario, PE: Prince Edward Island, QC: Quebec.

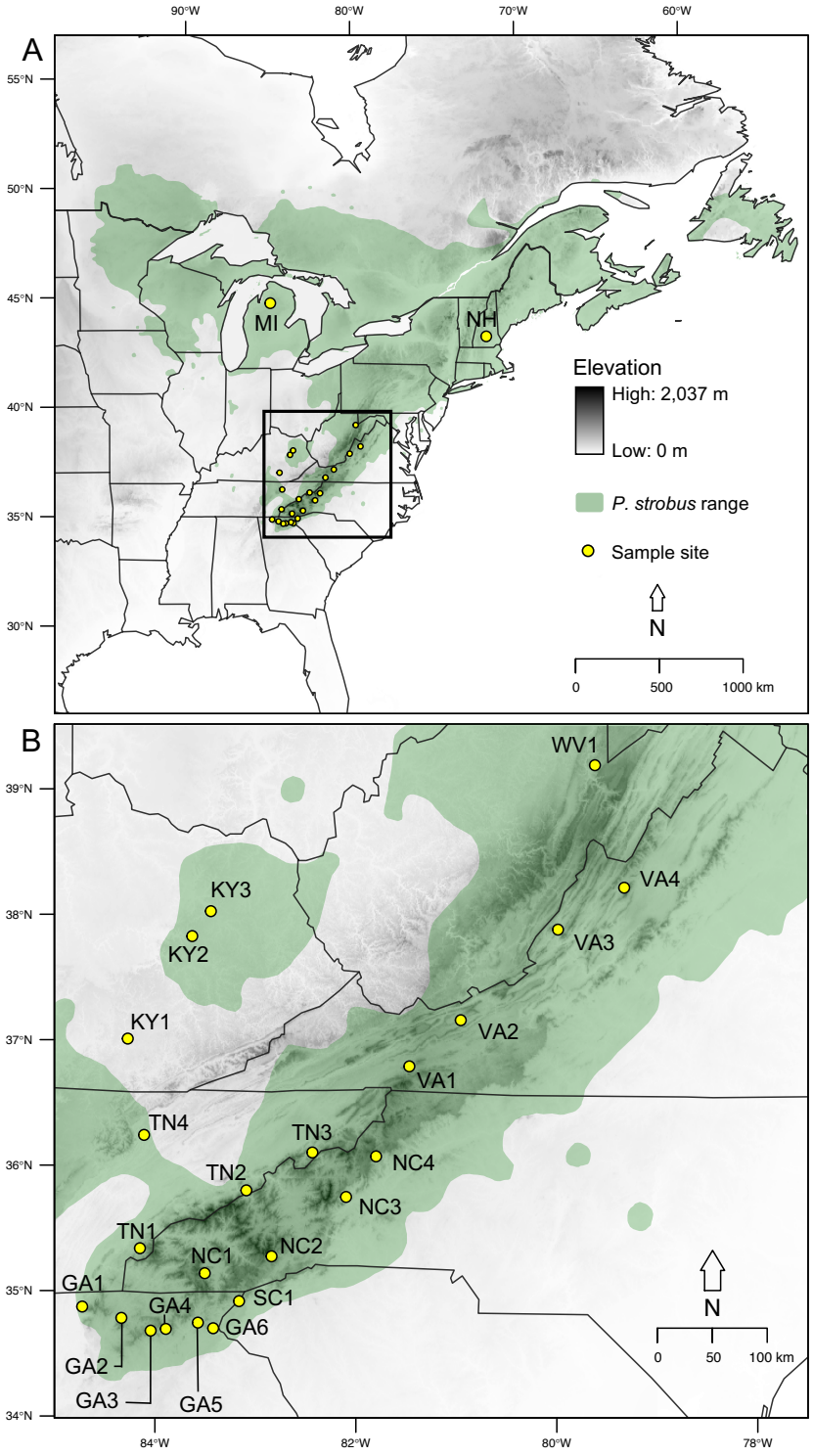

Figure 4.1. (A) Sampled population sites of *P. strobus* for molecular analyses. Shaded green area represents the range extent (Little, 1971). Inset (B) shows view of sampled populations within the Southern Appalachian region.

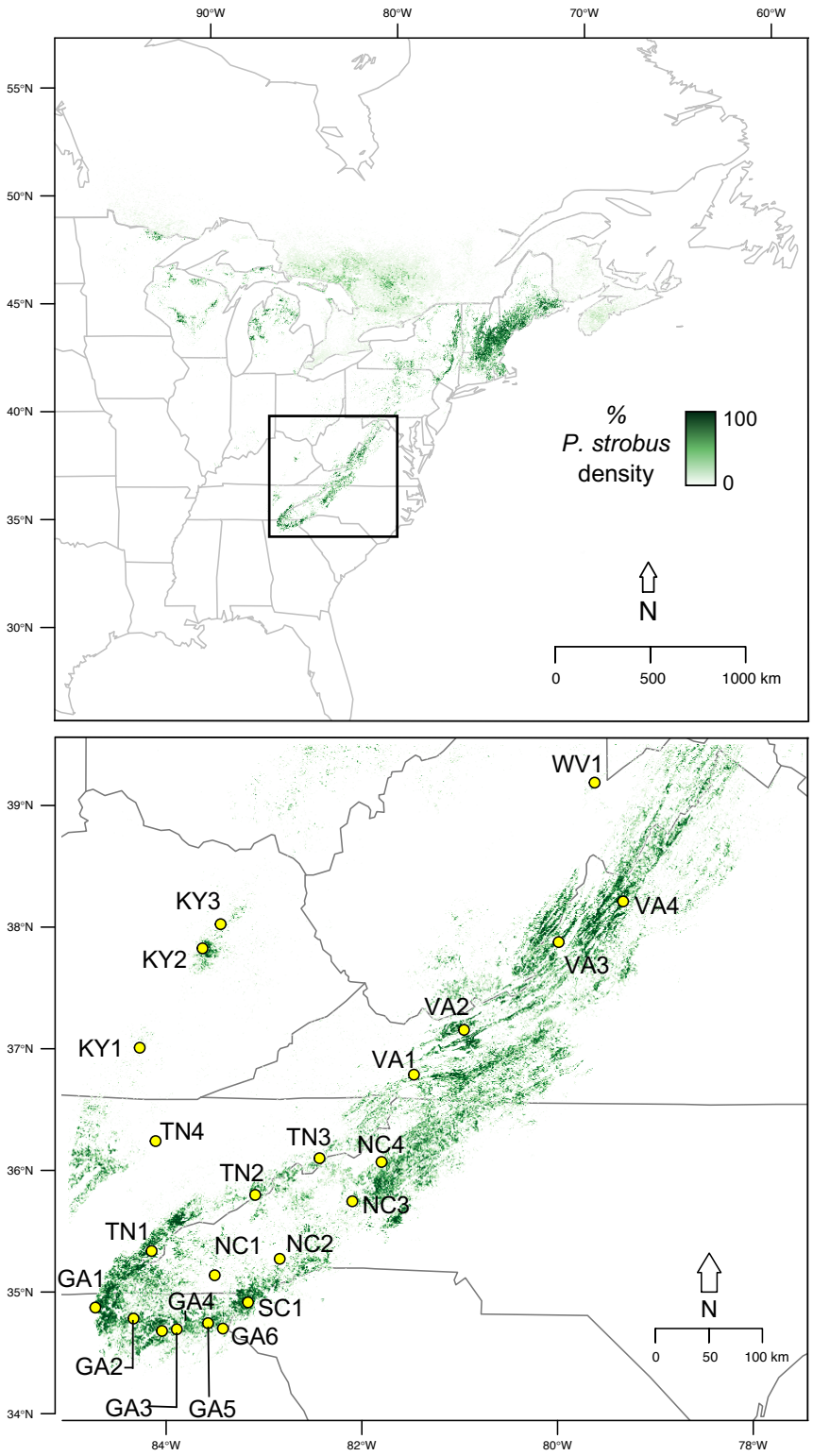

Figure 4.2. Relative density of *P. strobus* (A) in North America and (B) within the Southern Appalachian region. The value of each pixel (250 x 250 m) corresponds to the proportion of total tree area (≥12.7 cm DBH) that *P. strobus* encompasses.

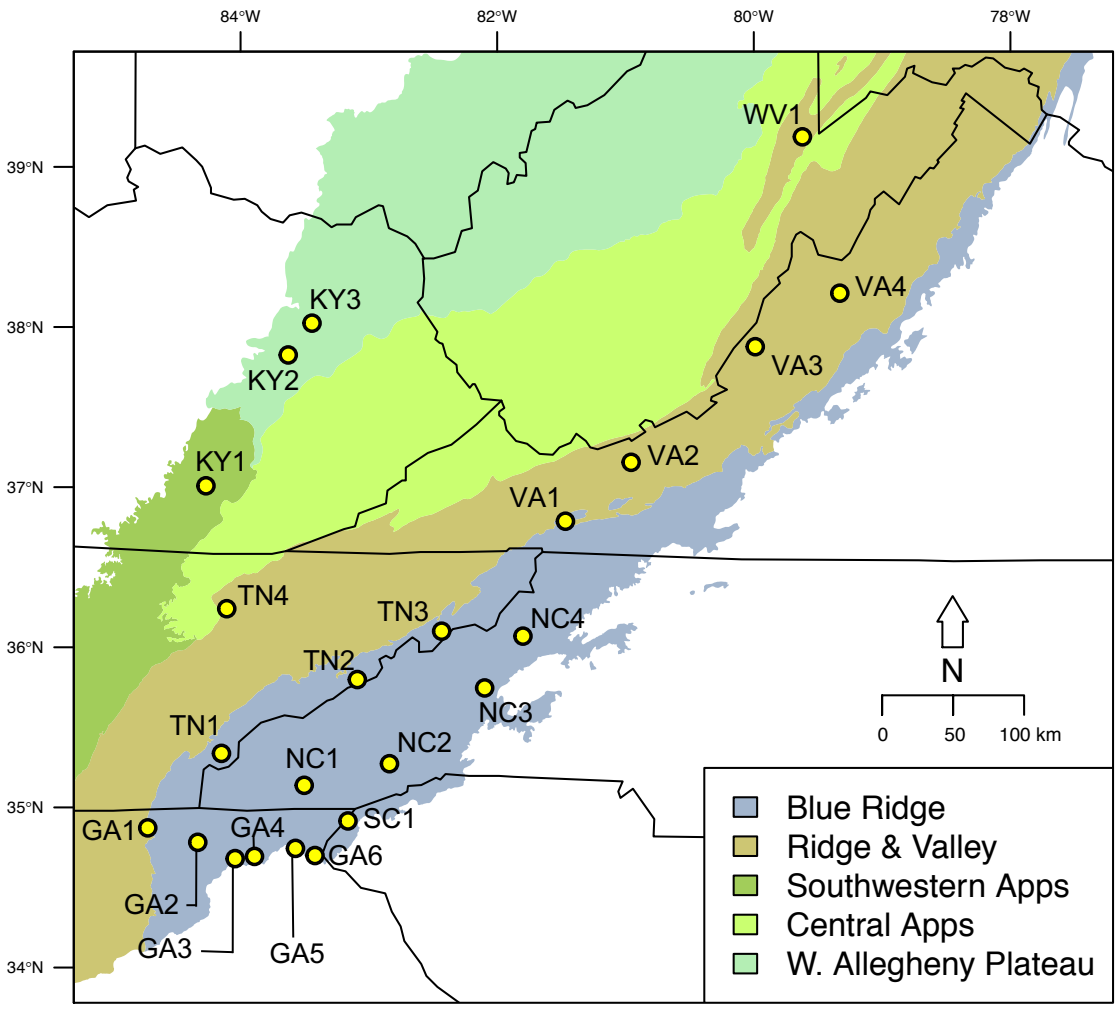

Figure 4.3. Sampled populations sites of *P. strobus* from the Southern Appalachians on a Level III ecoregion map (U.S. Environmental Protection Agency, 2013).

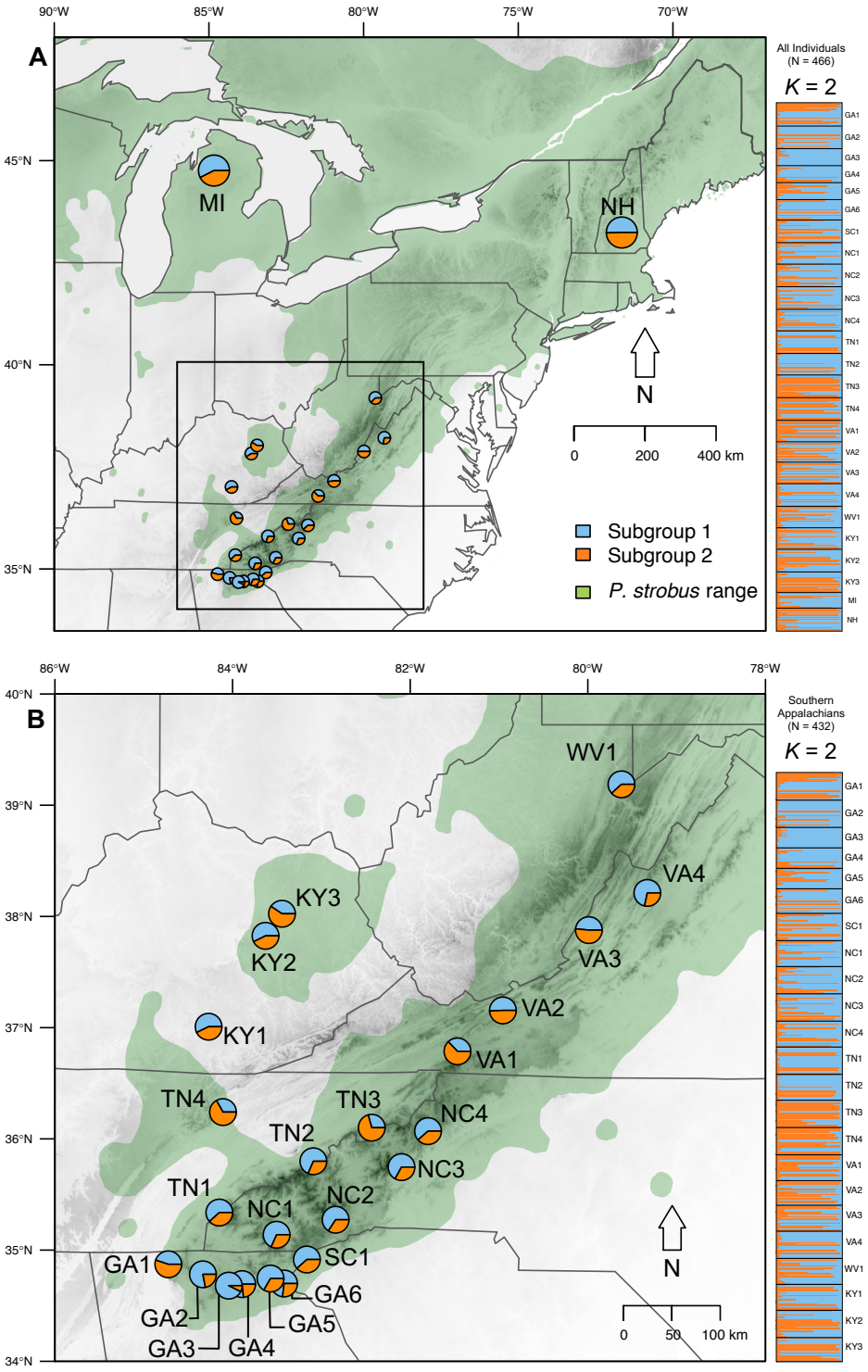

Figure 4.4. Population and individual cluster assignments from STRUCTURE for *P. strobus* (A) including all individuals (25 populations, N = 466), and (B) including only individuals from the Southern Appalachians (23 populations, N = 432).

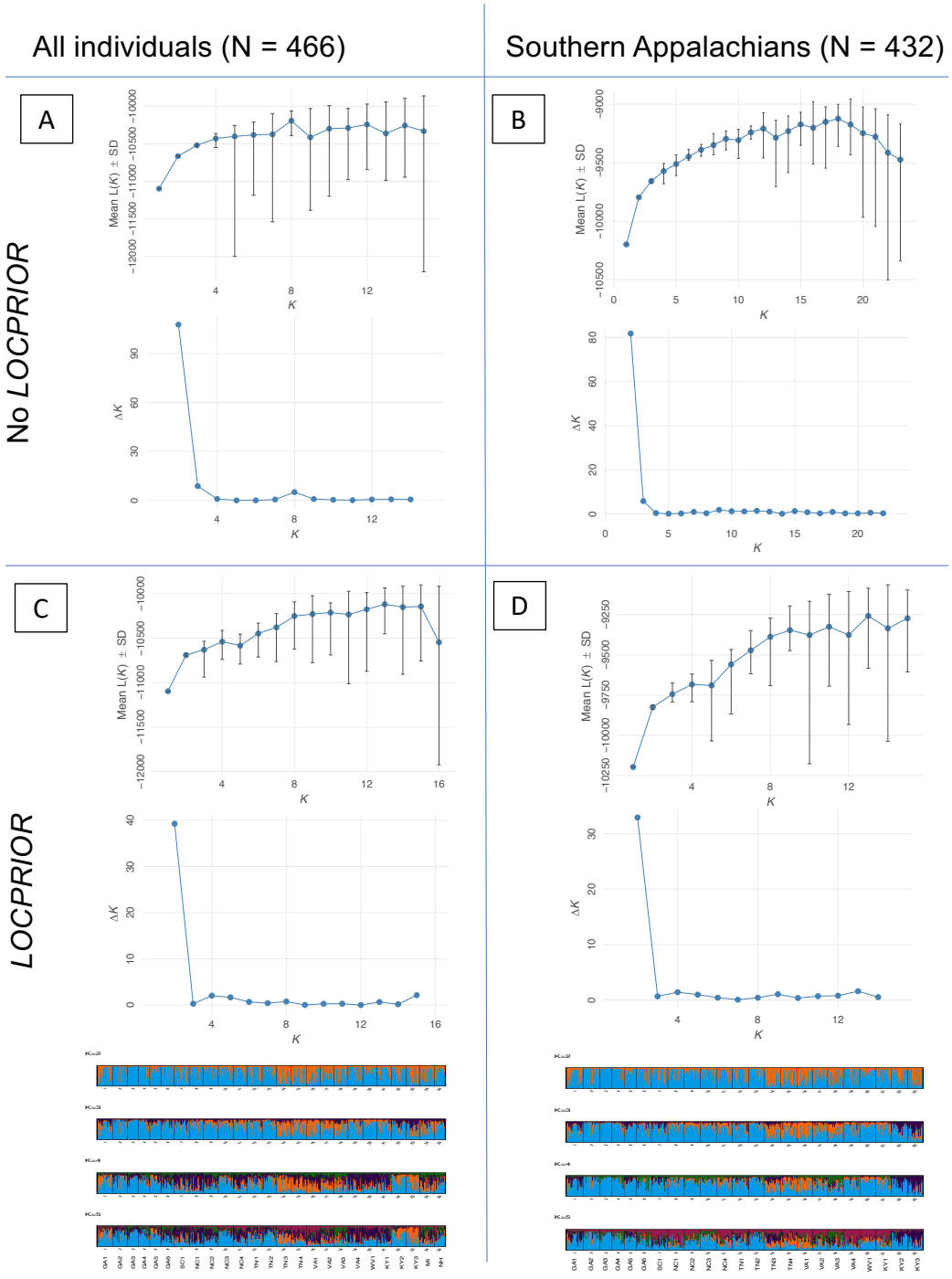

Figure 4.5. Post hoc tests LnPr(X|*K*) and Δ*K* (Evanno et al. 2005) for each run conducted in STRUCTURE and evaluated by STRUCTURE HARVESTER (Earl and Vonholdt 2012), categorized by the number of individuals tested and if the *LOCPRIOR* option was used (A-D). *Distruct* plots (Rosenberg 2004) of *K* = 2 to 5, generated from CLUMPAK (Kopelman et al. 2015), are displayed for STRUCTURE runs using the *LOCPRIOR* option (C, D).

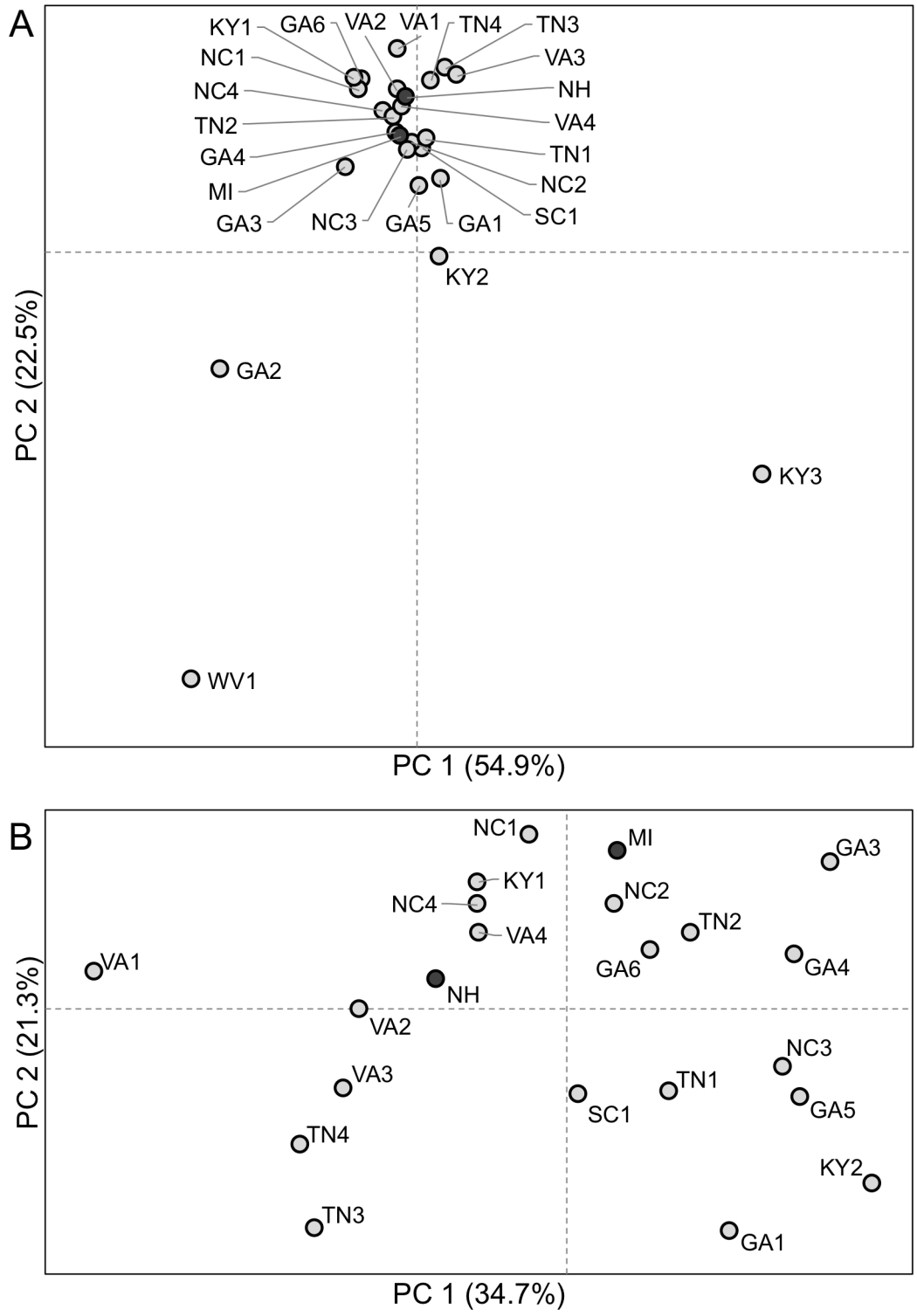

Figure 4.6. Principal coordinates analyses (PCoA) based on Nei's unbiased genetic distances of (A) all *P. strobus* populations (25 populations), and (B) all remaining populations after excluding GA2, KY3, and WV1. Dark gray points indicate the two northern populations from Michigan and New Hampshire.

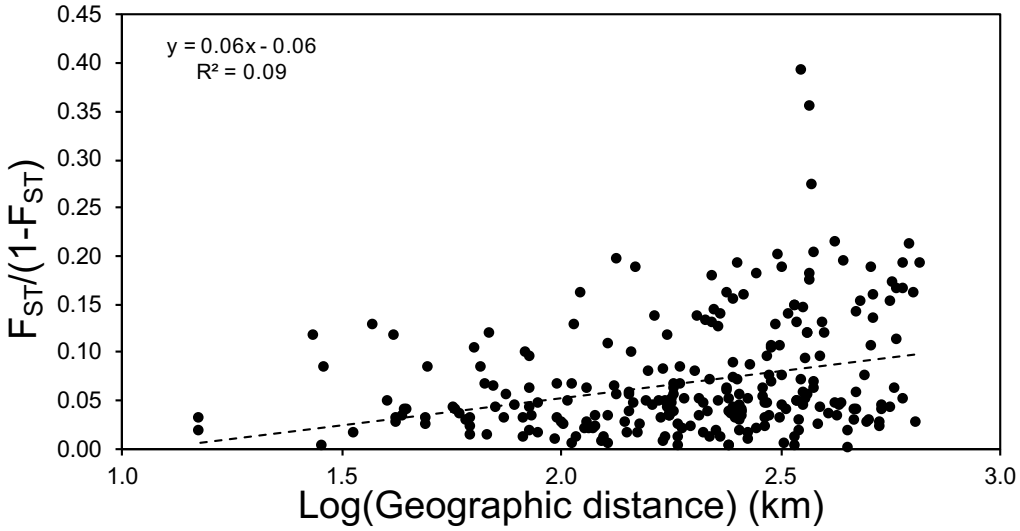

Figure 4.7. Mantel test comparing *P. strobus* linearized F_{ST} and log-transformed genetic distance.

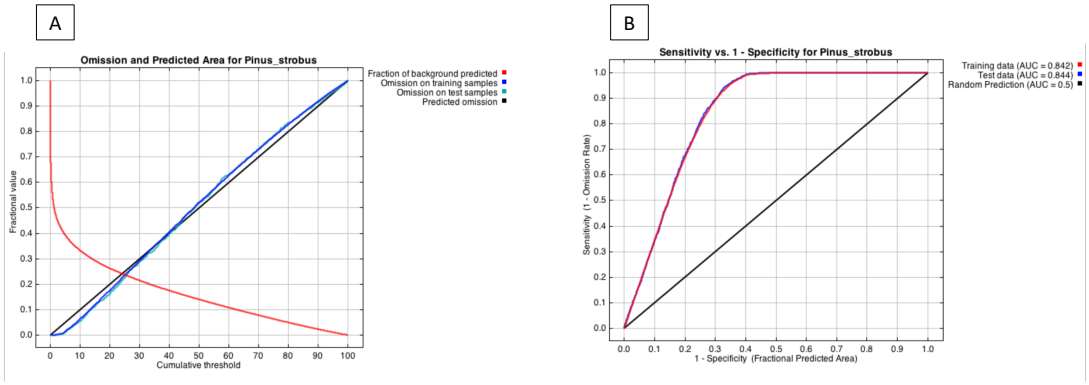

Figure 4.8. (A) Omission rate and (B) receiver operating characteristic (ROC) with AUC values for the MAXENT projections (Figure 4.9A-C).

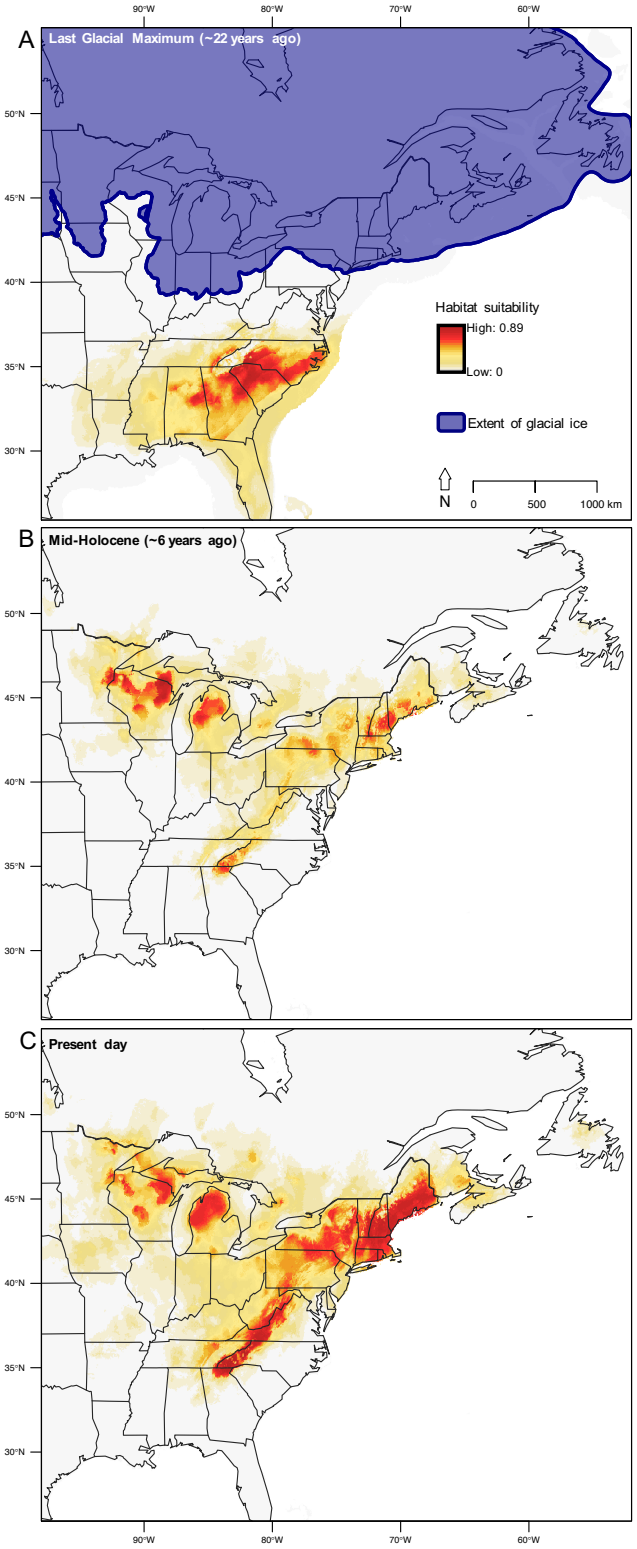

Figure 4.9. Modeled climate suitability envelopes of *P. strobus* during (A) the last glacial maximum (~22,000 years ago), (B) the mid-Holocene (~6,000 years ago), and (C) the present day (1970-2000).

CHAPTER 5

DISSERTATION CONCLUSIONS

5.1 Summary

The novel dieback phenomenon occurring in eastern white pine (*Pinus strobus* L.) has renewed concern for this important tree species. Previously inconspicuous biotic agents, now associated with each other and novel impacts, are being attributed as potential drivers in this dieback phenomenon. With little known about the damaging agents and the host tree, there is an urgency to establish foundational knowledge so that effective management strategies can be developed and deployed for the purposes of resource and ecological conservation. The research presented herein has elucidated valuable aspects about the biology, ecology and evolution of the eastern white pine bast scale (*Matsucoccus macrocicatrices*), its fungal pathogen affiliate, *Caliciopsis pinea*, and its host tree, *P. strobus*.

Past studies discovered correlative evidence to suggest *M. macrocicatrices* and *C. pinea* are involved with eastern white pine dieback. Not only is the presence of the insect-pathogen complex associated with trees experiencing dieback (Mech et al. 2013, Schulz et al. 2018a), but the incidence of each organism is also associated with the proportion of dead branches on saplings (Schulz et al. 2018b). Chapter two adds to this body of evidence, finding that the colonization hotspots within individual trees and across size classes are consistent with the characteristic symptoms. The pattern of eastern white pine branch dieback almost always occurs from the bottom up, and common observations by

tree size class have included: (1) poletimber experiencing the most bottom-up branch dieback, (2) saplings experiencing the most stem dieback, and (3) sawtimber experiencing the least amount/severity of dieback, in general. With the assumption, based on recent research, that *M. macrocicatrices* and *C. pinea* play a significant role in the symptomology, we hypothesized that the incidence of both scale insects and Caliciopsis cankers would be greatest on poletimber branches, sapling boles, and lower branches of every size class. Results showed that, as expected, patterns of colonization mirrored patterns of symptomology. Moreover, both *M. macrocicatrices* and *C. pinea* were most prevalent on bark tissue that was mature (i.e. not new growth), but still thin (i.e. <1 cm in thickness).

Matsucoccus macrocicatrices was first discovered in the Southern Appalachians in 2006 on trees infected with Caliciopsis canker (Mech et al. 2013), however the only known historical specimens were collected either in southeastern Canada or the New England states. A central question since then has been: "is it new or native to the Southern Appalachians?" This may provide clarity as to the role of *M. macrocicatrices* in the recent dieback phenomenon, especially if results showed it was a recent invader to the area. Chapter three sought to determine the historical biogeography of *M. macrocicatrices* since the LGM. We tested two competing hypotheses, that the presence of the insect outside of its purported native range (1) was due to recent introduction(s), or (2) was due to long co-occurring with its host, *P. strobus.* Results suggested *M. macrocicatrices* is native throughout its extant distribution. The fact that there was no documentation of this insect across most of its range, prior to revelations of its association with tree mortality, underscores the suddenness of its emergence as a pest.

Pinus strobus in the Southern Appalachians do not appear to be evolutionarily naïve to *M. macrocicatrices*, and yet trees are now showing unprecedented susceptibility to this insect and the Caliciopsis canker disease. The new threat from the insect-pathogen complex, compounded by warming climate trends that are predicted to cause significant *P. strobus* extirpation in its southern extent (Joyce and Rehfeldt 2013), calls for urgency to gather information pertinent for conservation of this forest tree in natural and managed systems. Chapter four investigated the genetic diversity and population structure of *P. strobus* populations in the Southern Appalachians to better understand its adaptive potential. We hypothesized there would be higher levels of genetic diversity (and by extension, adaptive potential) in this region compared to northern reference populations, consistent with the "southern richness, northern purity hypothesis" (Hewitt 2000). Instead, results indicated that *P. strobus* in the Southern Appalachians possesses high, but not higher levels of genetic diversity than populations in the northern range, and signatures of population structure were weak overall. Although not particularly distinct according to the neutral markers we used, *P. strobus* populations in the Southern Appalachians still appear to hold important conservation value.

The phylogeographic relationship between a host and a host-specialist organism should be fairly congruent if they co-evolved together. This has been documented extensively host-parasite systems (e.g. Johnson et al. 2001, Light and Hafner 2008, Nieberding et al. 2006) and to a lesser degree in tree-herbivore systems (e.g. Drag et al. 2018). Chapter three revealed evidence that *M. macrocicatrices*, like its host, may have survived the LGM at the southern extent of the Appalachian Mountains. However, there was an enormous difference between the population structure of the insect and the tree.

Populations of *M. macrocicatrices* were highly differentiated across the range (F_{ST} = 0.324) compared to *P. strobus* (F_{ST} = 0.060). Within the *SApps* cluster, which only included populations in Georgia, North Carolina, and Tennessee, *M. macrocicatrices* had much greater genetic differentiation (F_{ST} = 0.158) than *P. strobus* populations across the entire Southern Appalachians region (F_{ST} = 0.055). There was also significant substructure in *M. macrocicatrices*, whereas the *P. strobus* populations we tested appeared fairly admixed. This lack of congruence in population structure can likely be attributed to the differences in mating system and life history traits. Gametes from each organism can disperse long distances by wind, but *P. strobus* pollen travels further, likely because it is an order of magnitude smaller (0.075 mm; Science & Plants for Schools 2011) than the 1st instar crawlers of *M. macrocicatrices* (0.34-0.6 mm; Richards 1960). Wind may not carry seeds as far in some instances, but they can benefit from animal-facilitated dispersal. My results in chapters three and four show that barriers on the landscape affect each organism differently; the "Blue Ridge barrier" we found to restrict scale insect gene flow had no such effect on its host tree. Moreover, *M. macrocicatrices* cannot found new populations without its host already established, so this insect likely experienced a recolonization lag northward behind the more mobile *P. strobus* after the LGM. Additionally, the scale insect's generation time is only 1 or 2 years, whereas the tree's generation time can last decades. Therefore, less time is required for isolated *M. macrocicatrices* populations to differentiate than *P. strobus* populations. Faster generation time, a delay in recolonization after the LGM, and less subsequent inter-regional gene flow than its prolific host, may explain the incongruence in population structure between *M. macrocicatrices* and *P. strobus* we see today.

5.2 Implications

Assumed to be just one of over 250 innocuous herbivores of *P. strobus* (Wendel and Smith 1990), it remains a mystery why *M. macrocicatrices*, a native species, has recently been associated with severe tree injury and mortality. Anthropogenic disturbance is most often the driving force when natives become pests (Simberloff et al. 2012). Some have suggested *P. strobus* may simply be more susceptible to insect infestation and pathogen attack than before due to various contemporary abiotic norms (Costanza 2017, Munck et al. 2016, Mech et al. 2013, Schulz et al. 2018b). Costanza et al. (2018) reviewed several factors contributing to observed symptoms, which *M. macrocicatrices* may be exploiting, including climate change, land use, site conditions, and stand density. On the other hand, one documented biotic interaction involving the scale insect may deserve further attention. There are no known predators or parasitoids, but *M. macrocicatrices* may have historically been subject to population attenuation by the symbiont fungal species, *Septobasidium pinicola* Snell. This basidiomycete forms epiphytic mats on bark and parasitizes a proportion of the bast scales that settle underneath (Couch 1938). Watson et al. (1960) identified this relationship and hypothesized the two organisms actually engage in an obligate symbiosis, whereby 1st instar crawlers seek out and molt into 2nd instar cysts beneath these fungal mats for a protective feeding space, ultimately sacrificing the lives of a few individual in exchange. Today, *S. pinicola* is widespread, but is not common (Mech et al. 2013, Thomas Whitney, per obs.). This fungal symbiont does not currently appear to offer much in the way of biological control of the *M. macrocicatrices*, but perhaps it once did. Watson et al. (1960)

may too have overstated the obligate nature of this fungal-insect relationship, as their conclusions were based on a limited number of observations from a limited number of locales in Canada. Nonetheless, it may be worth exploring whether a biological control dynamic may have once been in place and has since been relaxed for some unknown reason(s).

The pathogen associated with eastern white pine dieback, *C. pinea*, is also presumed native to North America (Peck 1880, Ray 1936). Further investigation into what appears to be an insect-pathogen complex driving the Caliciopsis canker disease is a necessary next step in evaluating the actual mechanistic drivers of eastern white pine dieback. This system is reminiscent of the beech bark disease (BBD) complex, comprising of two beech (*Fagus* spp.) pathogens, both exploiting the feeding wounds of a scale insect, *Cryptococcus fagi* Baer., to cause tree mortality (Houston et al. 1994). The difference is that *C. fagi* was introduced from abroad, which triggered the disease when it established successfully on naïve beech hosts in North America. The organisms that appear to drive Caliciopsis canker, however, are both seemingly native. This presents a unique opportunity to understand how an indigenous species interaction, perhaps dormant or engaging infrequently in the past, can become more prominent over time. As with BBD, the exact mechanism for *C. pinea* infection is unknown. Bast scale feeding wounds are speculated to weaken localized bark tissue, which form ideal infection courts for *C. pinea* spores to penetrate, establish, and form cankers. A better grasp of their contemporary relationship may provide insights into their historical relationship.

5.3 Future directions

Pinus strobus remains an important tree species in eastern North American forests, and its value will only increase as one of its associate canopy species, eastern hemlock, continues to decline. The emergence of native organisms, interacting in novel ways, are affecting the density, distribution, and longevity of *P. strobus*, and thus management and conservation strategies will continue to adjust as we learn more about this forest health phenomenon. As part of a USDA multi-state project, work is already underway to demystify eastern white pine dieback and develop effective management and conservation strategies. As *M. macrocicatrices* has been solidified as an important component of this system, future research may consider the following objectives to further elucidate its role and recent emergence:

1. Determine if *M. macrocicatrices* feeding facilitates *C. pinea* infection of *P. strobus* trees. This may be achieved in a 2 x 3 factorial design on uninfected *P. strobus* trees grown in a greenhouse. Factor 1 would be "*C. pinea* inoculum," with the two levels being 'not inoculated' and 'inoculated.' Factor 2 would be "damage type," with the three levels being 'none' (control), 'mechanical damage,' and '*M. macrocicatrices* damage.' Mechanical damage would be achieved by puncturing bark with an insect pin, which simulates a scale insect feeding wound (this method is currently being implemented in experiments by Michelle Cram, USDA Forest Service). Actual damage caused by *M. macrocicatrices* would require rearing 1st instar crawlers to settle on experimental trees. There are several logistical challenges to *M. macrocicatrices* laboratory rearing, but if a factorial experiment such as this were successful, we may be able to disentangle the

specific effect of bast scale feeding damage on the infection of *C. pinea* and development of Caliciopsis canker.

2. Elucidate the mechanism. If scale feeding is found to have a synergistic effect on *C. pinea* infection, the next step would be to ascertain the mechanism at play. The broad hypothesis is that *M. macrocicatrices* feeding wounds weaken localized bark tissue to provide an infection court more suitable than a mechanical wound or a feeding wound from another similar insect. A thorough investigation into the scale insect's feeding biology may produce promising results. Whether it digests extra-orally, has toxic digestive enzymes in its saliva, and/or stymies defensive responses in its host, may all be pertinent to understanding the underlying mechanism causing Caliciopsis canker infection.
3. Develop a pheromone lure. A critical component in any pest prevention and mitigation program is monitoring, but we currently lack a practical and accurate method to survey this animal. Pheromone lures have proved highly effective to monitor congeners, and thus they should be developed for *M. macrocicatrices*. The ability to trap and count winged males every year will allow us (a) to conduct basic year-to-year density surveys, (b) understand the broad-scale and fine-scale distribution across the range of *P. strobus*, and even (c) identify where, latitudinally, the insect undergoes an annual versus a biennial life cycle (assuming generations do not overlap).
4. Assess the relationship between *M. macrocicatrices* and the fungus, *S. pinicola*. This fungus is presumably a specialist pathogen of *M. macrocicatrices* and is also the only possible biocontrol agent discovered as of now. Hence, this biotic

interaction warrants attention. Investigating if there are differences to the patterns and process of their interaction between sites in northeastern North America, where Watson et al. (1960) did their work, and in the Southern Appalachians, would be a possible angle to assess their ecological association.

5. Compare *M. macrocicatrices* with *M. matsumurae*, the red pine scale. Since 1946, *M. matsumurae* has invaded and become established as a pest of red pine (*P. resinosa* Sol. ex Aiton) in northeastern North America (Bean and Godwin 1955). The two now have overlapping ranges and, as congeners, have a similar functional biology. However, *M. macrocicatrices* is native and *M. matsumurae* is a non-native from Asia. This is a rare opportunity study side-by-side the demography, population dynamics, and host-use of two sister species with dissimilar historical biogeography, and therefore, possibly dissimilar evolutionary pressures and adaptations. There may be potential to not only determine how each pest independently proliferated, but more broadly, to also parse out the importance of a pest's biological/evolutionary characteristics, from their place of origin, as contributing factors to "invasiveness."
6. Explore the possibility of a cryptic invasion by *C. pinea*. This fungus is presumed to be native to North America but is also known to occur in Europe utilizing other pine species as hosts (Costanza et al. 2018). Using a population/phylo-genetic approach, it may be worthwhile to rule out that a morphologically similar, but genetically novel, non-native lineage of *C. pinea* did not recently and cryptically invade, resulting in multiple lineages, hybridization events, and/or species complexes.

7. Investigate resistant provenances of *P. strobus* to the insect-pathogen complex. Despite the overall low population structure, there is ample evidence that varying provenances exhibit varying responses to certain climatic conditions (Joyce and Rehfeldt 2013). Broadening future provenance testing to include susceptibility to *M. macrocicatrices* colonization and Caliciopsis canker development may be a worthy conservation avenue.

5.4 References

Bean, J.L., Godwin, P.A., 1955. Description and bionomics of a new red pine scale, *Matsucoccus resinosae*. Forest Sci 1, 164-176.

Costanza, K.K.L., 2017. Biological, ecological, and economic impact of Caliciopsis canker of *Pinus strobus* forests of northeastern North America. (PhD Dissertation) University of Maine, Orono, Maine, USA, p. 168.

Costanza, K.K.L., Whitney, T.D., McIntire, C.D., Livingston, W.H., Gandhi, K.J.K., 2018. A synthesis of emerging health issues of eastern white pine (*Pinus strobus*) in eastern North America. Forest Ecology and Management 423, 3-17.

Couch, J.N., 1938. The genus *Septobasidium*. University of North Carolina Press, Chapel Hill, North Carolina, USA.

Drag, L., Hauck, D., Rican, O., Schmitt, T., Shovkoon, D.F., Godunko, R.J., Curletti, G., Cizek, L., 2018. Phylogeography of the endangered saproxylic beetle Rosalia longicorn, *Rosalia alpina* (Coleoptera, Cerambycidae), corresponds with its main host, the European beech (*Fagus sylvatica*, Fagaceae). Journal of Biogeography 45, 2631-2644.

Hewitt, G., 2000. The genetic legacy of the Quaternary ice ages. Nature 405, 907-913.

Houston, D.R., 1994. Major new tree disease epidemics: beech bark disease. Annual Review of Phytopathology 32, 75-87.

Johnson, K.P., Drown, D.M., Clayton, D.H., 2001. A data based parsimony method of cophylogenetic analysis. Zoologica Scripta 30, 79-87.

Joyce, D.G., Rehfeldt, G.E., 2013. Climatic niche, ecological genetics, and impact of climate change on eastern white pine (*Pinus strobus* L.): Guidelines for land managers. Forest Ecology and Management 295, 173-192.

Light, J.E., Hafner, M.S., 2008. Codivergence in heteromyid rodents (Rodentia: Heteromyidae) and their sucking lice of the genus *Fahrenholzia* (Phthiraptera: Anoplura). Systematic Biology 57, 449-465.

Mech, A.M., Asaro, C., Cram, M.M., Coyle, D.R., Gullan, P.J., Cook, L.G., Gandhi, K.J.K., 2013. *Matsucoccus macrocicatrices* (Hemiptera: Matsucoccidae): First report, distribution, and association with symptomatic eastern white pine in the southeastern United States. Journal of Economic Entomology 106, 2391-2398.

Munck, I.A., Luther, T., Wyka, S., Keirstead, D., McCracken, K., Ostrofsky, W., Searles, W., Lombard, K., Weimer, J., Allen, B., 2016. Soil and stocking effects on Caliciopsis canker of *Pinus strobus* L. Forests 7.

Nieberding, C.M., 2006. Nieberding, C. M., & Morand, S. (2006). Comparative phylogeography: The use of parasites for insights into host history. In: Morand, S., Krasnov, B.R., Poulin, R. (Eds.), Micromammals and Macroparasites. Springer, New York, New York, USA.

Peck, C.H., 1880. Report of the Botanist. In, Annual Report N.Y. State Museum of Natural History, p. 32.

Ray, W.W., 1936. Pathogenicity and cultural experiments with *Caliciopsis pinea*. Mycologia 28, 201-208.

Richards, W.R., 1960. A new species of *Matsucoccus* Cockerell (Homoptera: Coccoidea). Canadian Entomologist 92, 179-181.

Schulz, A.N., Mech, A.M., Asaro, C., Coyle, D.R., Cram, M.M., Lucardi, R.D., Gandhi, K.J.K., 2018a. Assessment of abiotic and biotic factors associated with eastern white pine (*Pinus strobus* L.) dieback in the Southern Appalachian Mountains. Forest Ecology and Management 423, 59-69.

Schulz, A.N., Mech, A.M., Cram, M.M., Asaro, C., Coyle, D.R., Lucardi, R.D., Lucas, S., Gandhi, K.J.K., 2018b. Association of *Caliciopsis pinea* Peck and *Matsucoccus macrocicatrices* Richards with eastern white pine (*Pinus strobus* L.) seedling dieback. Forest Ecology and Management 423, 70-83.

Science & Plants for Schools, 2011. Pollen Image Library. http://www-saps.plantsci.cam.ac.uk/pollen/index2.htm (Accessed on September 8, 2018).

Simberloff, D., Souza, L., Nunez, M.A., Barrios-Garcia, M.N., Bunn, W., 2012. The natives are restless, but not often and mostly when disturbed. Ecology 93, 598-607.

Watson, W.Y., Underwood, G.R., Reid, J., 1960. Notes on *Matsucoccus macrocicatrices* Richards (Homoptera: Margarodidae) and its association with *Septobasidium pinicola* Snell in eastern Canada. Canadian Entomologist 92, 662-667.

Wendel, G.W., Smith, H.C., 1990. *Pinus strobus* L. eastern white pine. In: Burns, R.M., Honkala, B.H. (Eds.), Silvics of North America. USDA Forest Service, Washington DC, pp. 476-488.

APPENDICES

Appendix 1: Life cycle

In early 2016, I successfully raised *M. macrocicatrices* cysts through the adult, egg, and crawler stages in laboratory conditions. On January 28th, 2016, I collected branches infested with 2nd instar cysts from the Chattahoochee National Forest, Georgia, USA (Canada Creek, 34.6803, -84.0425), and put them in the refrigerator upon returning to the laboratory in Athens, Georgia. We have found that maintaining scales at ~4 °C keeps them in diapause, and they can be triggered to emerge into adulthood almost immediately when brought into a room temperature environment for an indefinite period of time. This artificially induced diapause in the refrigerator can be maintained for upwards of 3 months. On January 29th (day 0), I removed the collected branches from the refrigerator and put them in medium-sized plastic containers with a paper towel floor taped to the bottom. A moist paper towel ball was included for humidity regulation. In the following days, pre-adult males started emerging from their cysts (see Figure A1.1 for associated pictures of male development). Groups of them were removed and isolated in sealed petri dishes, with a taped towel floor and a wetted towel ball. By day 7 most males had entered the "pupal" stage in silken "cocoons." Wing pads were starting to develop, the compound eyes were larger, antennae were elongating, and a "skin-like" exuviae was starting to peel off the anterior end. By day 9 some of the male bodies were darkening in color, nearly done with wing development, and even twitching a bit. By day 12 they appeared to be nearly finished in their development. On day 14, many of the males were

fully formed and were walking around the petri dishes. At the same time, a single female was emerging from her cyst on one of the branches still in one of the plastic containers (see Figure A1.2 for associated pictures of female emergence, mating, egg-laying, and 1st instar hatching). After some time, she freed herself and I isolated her within a petri dish environment as above. I then facilitated a mating event with one of the males, by introducing him into the female's petri dish. The male immediately started to try and mate with her. At first, he probed her head repeatedly for a short time, reoriented and then successfully achieved copulation. The event lasted ~1 min, after which the male departed. On day 15, the female started producing silk from her abdomen. On day 16, she continued to produce silk, as well as some powdery white secretion. The male was also found dead on this day. On day 17, the female was producing eggs and laying them in a way that left them suspended within the silken mass. The dozen or so eggs produced were yellow. On day 18, the female had finished laying her eggs in the silk bundle and lay as a corpse, never having moved. Eggs incubated, eventually growing larger with little black eyes visible through the translucent exterior. On day 34, 1st instar crawlers hatched from the eggs, crawling with a vigorous pace within the silk, but not straying. Their appearance resembles a smaller, yellowish, more elongate version of the adults. On this day I attempted to inoculate a *P. strobus* twig, suspended in a test tube of water, by placing the silk mass carefully on the twig. In the days following, each crawler eventually dispersed from the silk, but unfortunately, none accepted the host or attempted feeding. This was a good first step toward establishing a future rearing program for *M. macrocicatrices*. Live trees in a laboratory greenhouse, preferably large enough to have a DBH, would probably yield a higher success rate for 1st instar colonization. Nonetheless,

the method demonstrated here, meticulously isolating emerged adults into petri dish environments, was successful at producing 1st instar progeny starting from 2nd instar individuals collected from the wild.

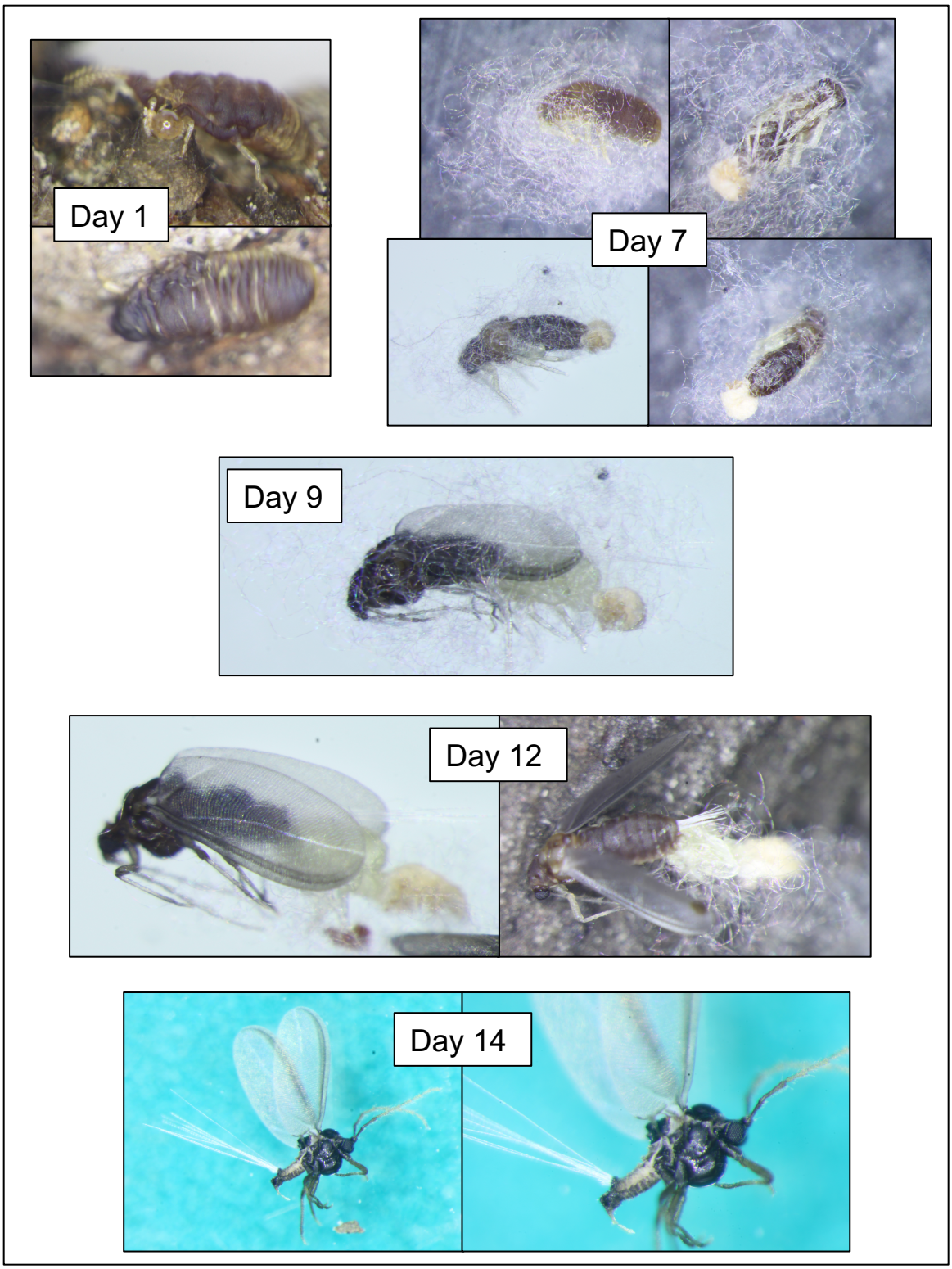

Figure A1.1. Development of male *M. macrocicatrices* from pre-adult (3rd instar) to final winged form.

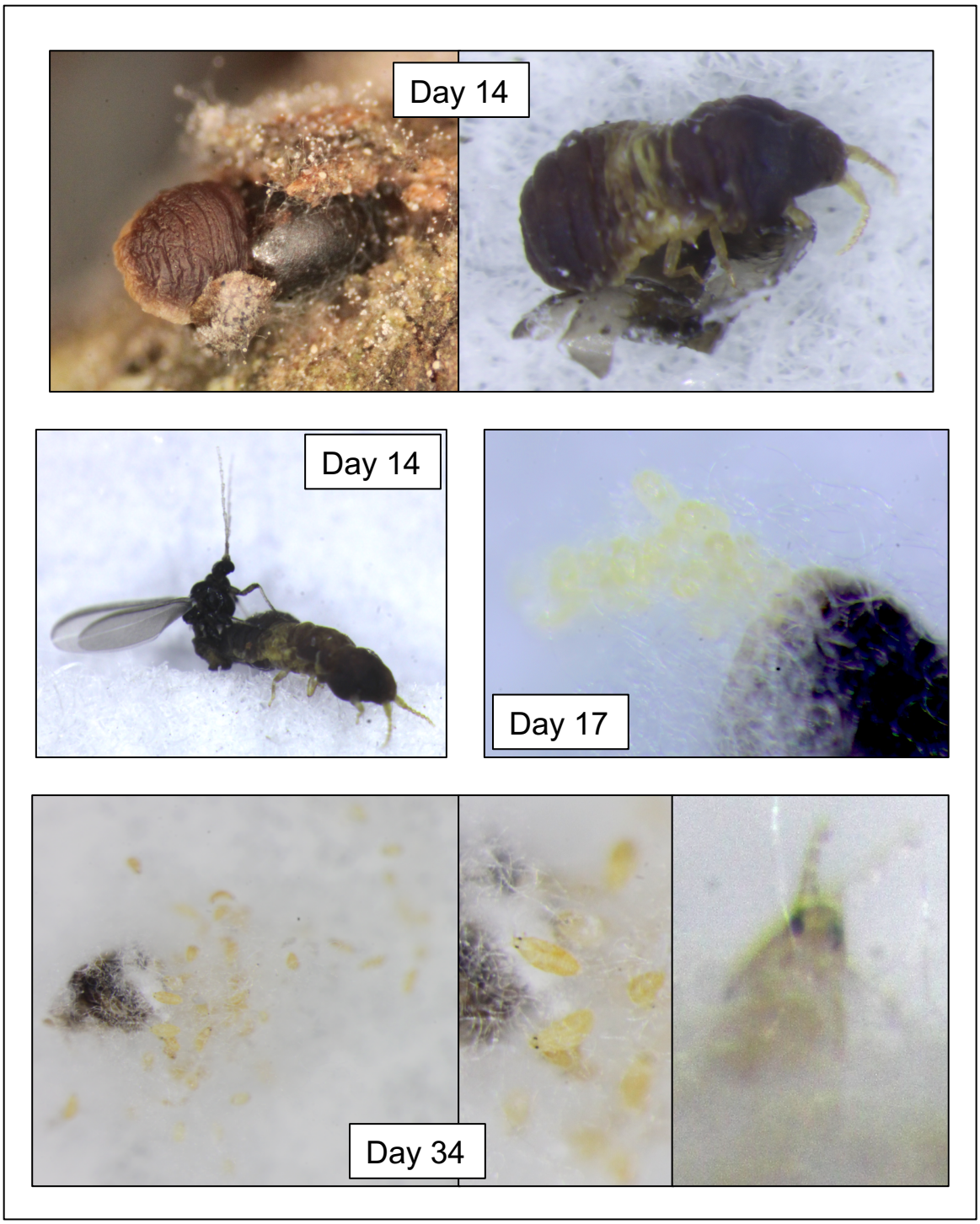

Figure A1.2. Emergence of female *M. macrocicatrices*, mating event, egg laying, and hatching of 1st instar crawlers.

Appendix 2: Cold tolerance

Insects are ectotherms, and thus must employ one of two strategies to avoid mortality while overwintering. “Freeze-tolerant” insects can tolerate internal freezing to overcome exposure to sub-zero temperatures (Doucet et al. 2009), but most insects are instead considered ‘freeze-avoiding’ (Sinclair et al., 2003), meaning they can maintain the liquidity of their hemolymph below its melting point, a process termed supercooling (Salt, 1961). The temperature at which an insect cannot supercool any further, and ice crystallization of its body fluids occurs, is deemed the supercooling point (SCP). The SCP is a commonly used metric for assessing the level of cold hardiness in insects. In the only published study to investigate cold tolerance in a scale insect, Wang et al. (2019), found that the crapemyrtle bark scale, *Acanthococcus lagerstroemiae* (Kuwana) (Hemiptera: Eriococcidae), employs various cold hardening strategies, including the reduction of internal water content and the production of a suite of cryoprotectant chemicals. It is likely *M. macrocicatrices* employs similar strategies but perhaps to varying degrees in the warmer southern range and the colder northern range. Here, we assessed the level of cold tolerance of *M. macrocicatrices* by quantifying the SCP of individuals from Michigan and Georgia, USA.

Twenty-eight live *M. macrocicatrices* cysts (see Table A2.1 for collection details) were sent overnight to Ben Philip (Rivier University, Nashua, New Hampshire, USA) for SCP analysis. Insects were kept at 4 °C until tests began. Supercooling points of an animal can be determined by steadily cooling the live individual and observing the release of latent heat that occurs during an exothermic reaction (e.g. liquid to solid).

Individual scales were put into 2.0 mL microcentrifuge tubes with a copper-constantan thermocouple. Each thermocouple was connected to a USB Thermocouple Data Acquisition Module (TC-08; OMEGA Engineering, Stamford, Connecticut, USA). Given the miniscule size of the scales, a dab of petroleum jelly was used as an adhesive to ensure scales were in contact with the tip of the thermocouple. The samples were placed into 50-mL conical tubes and suspended in an ethylene glycol cold bath (NESLAB RTE-740, Thermo Scientific, Waltham, Massachusetts, USA) set to 5 °C (Figure 3.1b). After all the samples standardized to 5 °C, we cooled the bath from 5 °C to -25 °C at a rate of 0.5 °C min^{-1}. We used OMEGA data acquisition software (OMEGA Engineering) to determine scale supercooling points from the latent heat of crystallization. The supercooling point for each individual *M. macrocicatrices* was determined from the latent heat of crystallization (after Lee 1989).

Raw results and summary statistics are presented in Table A2.1. The supercooling points of 39% of the tested scale insects exceeded -25 °C, which is beyond the cooling capacity of machinery we used. In these cases, we assigned a SCP value of -26 °C to enable rough, conservative estimates of mean SCP. The mean SCP of all 28 individuals was -21.6 °C. Individuals from Georgia (N = 21) had a mean SCP of -20.9 °C and individuals from Michigan (N = 7) had a mean SCP of -24.5 °C. Sample size was low and biased among groups, but these data indicate there is likely a difference in cold tolerance between *M. macrocicatrices* based on latitude. This may suggest *M. macrocicatrices* produces differing amounts or types of cryoprotectants in anticipation of winter depending on locale, which has implications for its phenology and physiology across its range.

Appendix 2 References

Doucet, D., Walker, V.K., Qin, W., 2009. The bugs that came in from the cold: molecular adaptations to low temperatures in insects. Cellular and Molecular Life Sciences 66, 1404-1418.

Lee, R.E., 1989. Insect cold-hardiness: to freeze or not to freeze - how insects survive low-temperatures. Bioscience 39, 308-313.

Salt, R.W., 1961. Principles of insect cold-hardiness. Annual Review of Entomology 6, 55-73.

Sinclair, B.J., Vernon, P., Klok, C.J., Chown, S.L., 2003. Insects at low temperatures: an ecological perspective. Trends in Ecology & Evolution 18, 257-262.

Wang, Z.N., Chen, Y., Diaz, R., Laine, R.A., 2019. Physiology of crapemyrtle bark scale, *Acanthococcus lagerstroemiae* (Kuwana), associated with seasonally altered cold tolerance. Journal of Insect Physiology 112, 1-8.

Table A2.1. Collection dates, test dates, collection locales, and supercooling points for SCP analysis conducted on 28 *M. macrocicatrices* cysts from Georgia and Michigan, USA.

#	Date collected	Date tested	State	Site name	SCP (°C)
1	2/18/16	2/19/16	Georgia	Boggs Creek	-12.5
2	2/18/16	2/19/16	Georgia	Boggs Creek	-13.5
3	2/18/16	2/19/16	Georgia	Boggs Creek	-26.0*
4	2/18/16	2/29/16	Georgia	Boggs Creek	-10.9
5	2/18/16	2/29/16	Georgia	Boggs Creek	-11.7
6	2/18/16	2/29/16	Georgia	Boggs Creek	-13.4
7	2/18/16	2/29/16	Georgia	Boggs Creek	-13.4
8	2/18/16	2/29/16	Georgia	Boggs Creek	-18.6
9	2/18/16	2/29/16	Georgia	Boggs Creek	-22.4
10	2/18/16	2/29/16	Georgia	Boggs Creek	-26.0*
11	2/18/16	2/29/16	Georgia	Boggs Creek	-26.0*
12	2/18/16	2/29/16	Georgia	Boggs Creek	-26.0*
13	2/18/16	2/29/16	Georgia	Boggs Creek	-26.0*
14	2/18/16	2/29/16	Georgia	Boggs Creek	-26.0*
15	2/18/16	2/29/16	Georgia	Boggs Creek	-26.0*
16	3/17/16	4/8/16	Georgia	Canada Creek	-23.9
17	3/17/16	4/8/16	Georgia	Canada Creek	-20.6
18	3/17/16	4/8/16	Georgia	Canada Creek	-22.7
19	3/17/16	4/8/16	Georgia	Canada Creek	-22.9
20	3/17/16	4/8/16	Georgia	Canada Creek	-22.4
21	3/17/16	4/8/16	Georgia	Canada Creek	-22.2
22	3/22/16	4/8/16	Michigan	Crawford Co.	-25.9
23	3/22/16	4/8/16	Michigan	Crawford Co.	-26.0*
24	3/22/16	4/8/16	Michigan	Crawford Co.	-26.0*
25	3/22/16	4/8/16	Michigan	Crawford Co.	-20.6
26	3/22/16	4/8/16	Michigan	Crawford Co.	-22.2
27	3/22/16	4/8/16	Michigan	Crawford Co.	-26.0*
28	3/22/16	4/8/16	Michigan	Crawford Co.	-26.0*
				Mean (Georgia)	**-20.9**
				Mean (Michigan)	**-24.7**
				Mean (All)	**-21.6**

* Indicates a supercooling point lower than -25 °C, but unable to register an accurate reading. These individuals were given a value of -26 °C for approximate summary statistics

Appendix 3: *Wolbachia*

Wolbachia is a genus of bacteria widely known to form endosymbioses with arthropods and nematodes (Werren et al. 2008). These interactions are sometimes facultative and sometimes obligatory. Infection occurs within reproductive organs and can have varied effects on fitness. The effects of *Wolbachia* infection include: (1) cytoplasmic incompatibility, preventing fertilization of eggs and sperm, (2) parthenogenesis, resulting in unfertilized eggs, (3) feminization of genetic males, and (4) male killing (Werren et al. 2008). Never before has *Wolbachia* infection been investigated in *Matsucoccus macrocicatrices.*

The wsp81F and wsp691R primers (Braig et al. 1998) were used to amplify strains of *Wolbachia* within *M. macrocicatrices* cysts. We tested the same 44 individuals used for 28S from section 3.2.2.2: two individuals from all 22 populations from chapter three. Polymerase chain reactions (PCRs) were carried out in 25-µL volumes and consisted of 2.5 µL 10x Takara buffer (Takara Bio Inc.), 0.125 mM dNTPs, 0.4 µM forward primer (wsp81F), 0.4 µM reverse primer (wsp691R) (Integrated DNA Technologies, Coralville, Iowa, USA), 1 U Takara *Ex Taq*™, and 1-10 ng template DNA. All reactions were conducted in Mastercycler Pro S Thermal Cyclers (Eppendorf), under the following PCR protocol: 94 °C for 4 min; 30 cycles of 94 °C for 40s, 55°C for 40 s, and 72 °C for 60 s; and a final extension of 72 °C for 10 min. Products were ran on gels and compared to positive controls (Linyphiidae spiders infected with *Wolbachia*), provided by Jennifer White (University of Kentucky).

Results show that *Wolbachia* is fairly common and widespread in *M. macrocicatrices*, as 29.5% tested positive for the bacterium. The 13 positive samples out

of 44 total were from the following locales: Maine (ME, N = 2 of 2), Michigan (MI1, N = 2 of 2), New Hampshire (NH3, N = 1 of 2), North Carolina (NC2, N = 2 of 2), Pennsylvania (PA1, N = 2 of 2), Virginia (VA2, N = 2 of 2), and West Virginia (WV, N = 2 of 2). All other samples were negative, including all "*SApps*" populations south of NC2 (i.e. all Georgia populations, TN, and NC1). I speculate *Wolbachia* is universal in *M. macrocicatrices* and would probably be found in all populations if a greater number of samples were tested. The role of this endosymbiont, if any, for this particular insect species, is still unknown.

Appendix 3 References

Braig, H.R., Zhou, W.G., Dobson, S.L., O'Neill, S.L., 1998. Cloning and characterization of a gene encoding the major surface protein of the bacterial endosymbiont *Wolbachia pipientis*. Journal of Bacteriology 180, 2373-2378.

Werren, J.H., Baldo, L., Clark, M.E., 2008. *Wolbachia*: master manipulators of invertebrate biology. Nature Reviews Microbiology 6, 741-751.