Seed dispersal differences and determinants of ranges of narrow endemic and widespread eastern North American trilliums

A Dissertation Presented for the

Doctor of Philosophy

Degree

The University of Tennessee, Knoxville

Chelsea Nicole Miller

May 2020

ACKNOWLEDGEMENTS

I would like to thank my advisor, Dr. Charles Kwit, in addition to the members of my doctoral committee, Dr. James Fordyce, Dr. Edward Schilling, and Dr. Joseph Clark for their patience, advice, willingness to listen, and for their feedback on this document in all of its forms. I will always be grateful for their support. I would like to extend additional thanks to Dr. Fordyce for his invaluable assistance with statistical analyses and experimental design.

I would like to thank Drs. Susan Whitehead and Monica Papeş, without whom the second and third chapters of this dissertation would not have been completed, for their willingness to pass along their expertise and advice, and for their generosity in allowing me to make use of their lab spaces and limited time. I would like to thank Dr. Jessica Budke for her mentorship during my Research Assistantship at the University of Tennessee Herbarium, and for her excellent career advice. The intellectual and emotional guidance provided to me by these mentors were critical to my success.

I would like to thank Kane Lawhorn, Jaime Albert, and Joanna Huntoon for their assistance with field work, laboratory experiments, specimen identification, and maintenance of ant colonies in the Kwit Lab. I would like to thank the Georgia, North Carolina, and South Carolina Departments of Natural Resources for allowing me to collect data and specimens with their permission from state forests, natural areas, and wildlife management areas. I would like to thank Dr. Aaron Floden, Dr. Edward Schilling, Jayne Lampley, and the late Dr. Thomas Patrick for their assistance in locating study sites across the southern Appalachian region and for teaching me to identify my study species. I would like to thank the Garden Club of America, the Botanical Society of

America, the University of Tennessee Herbarium, the University of Tennessee Graduate School, and the University of Tennessee Ecology and Evolutionary Biology Department for funding. I would like to extend special thanks to my colleagues Chloe Lash, Dr. Todd Pierson, Tyson Paulson, Hailee Korotkin, Harmony Yomai, Alex Edwards, and Dr. Zach Marion for their encouragement, their feedback on manuscripts, grants, and experimental design, and for their friendship. Finally, I would like to thank my partner and my family for their ceaseless support and love.

ABSTRACT

Drivers of species rarity can be classified as abiotic (environmental) or biotic (species interactions), and the spatial scale at which factors impact rarity ranges from fine to coarse. This dissertation explores drivers of narrow endemism, a kind of rarity in which species are restricted to geographic regions or habitat types, in the genus *Trillium,* a group of perennial understory forest herbs in eastern North America. To identify factors characterizing narrow endemism in this system, I used a framework comparing ecological attributes of co-occurring endemic and widespread congeners. At a fine spatial scale, I explored whether seed dispersal effectiveness differed for endemic and widespread species. Trilliums rely on ants for seed dispersal; my work establishes that rates and probabilities of seed dispersal are lower in the field for endemic trilliums, and that ants display a preference for the seeds of widespread trilliums, suggesting that narrowly endemic species may be dispersal limited. To investigate underlying mechanisms, I quantified interspecific differences in the morphology and chemistry of seeds and elaiosomes (i.e. seed-coat appendages that are attractive to ants). I found that seed morphology, concentrations of fatty acids, and aspects of elaiosome phytochemistry were significantly related to dispersal probability for five species of *Trillium,* suggesting that endemic species produce less-attractive seeds. At a coarse spatial scale, I generated species distribution models (SDMs) using Maxent for 21 species of *Trillium*; SDMs produced areas of predicted suitability for each species based on abiotic conditions. Within this framework, I found that rarer species of greater conservation concern had lower proportional occupancy of the fundamental niche than common species of lower conservation concern; that SDMs performed best for species of higher conservation

concern with larger predicted suitable areas; and finally, that flower type and ovule number were significantly related to model performance. These discoveries provide valuable insight into the ecology and biogeography of species in the genus *Trillium*, thus advancing the conservation-relevant knowledgebase for this horticulturally-important group of plants, and contribute to our understanding of factors that maintain plant endemism at fine and coarse spatial scales, from both abiotic and biotic perspectives.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

The study of species rarity has been granted a great deal of attention by ecologists and evolutionary biologists over the past 150 years (von Mueller, 1855; Hooker, 1872; Rabinowitz, 1978; Leitao et al., 2016). In particular, the causes and persistence of plant rarity have been of great interest to the ecological community (Karron, 1991; Levin et al., 1996; Speed and Austrheim, 2017). Studies of plant rarity have frequently been undertaken in the spirit of protecting and conserving rare plant populations (Ellstrand and Elam, 1993; Schmidt and Jensen, 2000; Wu and Smeins, 2000; Honnay and Jacquemyn, 2007). To this end, ecologists have employed a myriad of empirical and theoretical methods to study of rare plants. These include censuses, inventories, and surveys (Palmer, 1987; Jalali and Jamzad, 1999), manipulative experiments (Fischer and Matthies, 1998; Klironomos, 2002), investigations of management techniques (Falk et al., 1996; Bevill et al., 1999), and mathematical models aimed at defining current distributions of rare plant species (Engler et al., 2004; Williams et al., 2009), identifying the ecological niche of rare plants (Aguirre-Gutiérrez et al., 2015), and projecting future range shifts of rare plants under conditions of climate change (Huntley et al., 1995; Van der Putten et al., 2010).

One established method for examining factors regulating rare plant populations is through demographic comparisons of related rare and common taxa (Karron, 1987; Walck et al., 1999). This method has been used to identify a variety of characteristics that differ for related rare and common taxa, including seed-affiliated life history transitions (Combs et al., 2013), differences in width of ecological amplitude for soil

biogeochemical parameters (Kleijn et al., 2008), and patterns of genetic variation (Soltis and Gitzendanner, 1999) that differ between co-occurring rare and common congeners. In 1999, Bevill and Louda comprehensively reviewed the literature comparing attributes of rare vascular plants with those of related, common relatives and found that the frequency of these studies was lower than expected. They concluded that the database underlying the understanding and management of rare compared to common plant species is scattered and conceptually weak. Although the number of studies comparing rare and common plants has increased since this review, their conclusions highlight the need for more attention to the comparative ecology of rare and common plants.

Research into the causes of plant rarity and of the phytogeography of rare plants have often focused on coarse-scale abiotic (i.e. environmental) variables. Climate, hydrology, topography, precipitation, temperature, and elevation have all been shown to regulate plant distributions (Baskin and Baskin, 1988; Harrison et al., 2006), and have been investigated in many systems with a variety of empirical and modelling strategies. Specifically, environmental and land-use changes (Lavergne et al., 2005), edaphic and light-level factors (Baskin and Baskin, 1988), and topographic factors (Miller, 1986) have all been shown to impact rare plant distributions and abundance. Developing an understanding of the ecological niche of rare plants using species distribution modeling (SDM) can reveal important focal areas for conservation planning and predict future distributions under conditions of anthropogenic change. SDM targets the current geographic distributions of species using presence (i.e. occurrence) data and environmental (i.e. climate) data layers, and produce maps of predicted suitable area (Soberon and Peterson, 2005). These models are valuable tools for understanding how

abiotic variables shape species distributions. In the context of a rare-common species framework, species distribution modeling may help identify coarse-scale abiotic factors that contribute to plant rarity at a landscape scale.

Fine-scale biotic factors also contribute to plant rarity in important ways. Biotic factors are critical components of the realized niche (Hutchinson, 1957). Negative interactions such as parasitism, predation, or disease can result in a reduced realized niche, which may translate to reduction in geographic distribution, restriction in habitat suitability, or decrease in the size of local populations. Any of these outcomes could increase rarity. Genetic interactions, such as interspecific hybridization, can also lead to rarity in plants, sometimes with the ultimate consequence of extinction due to demographic swamping and genetic assimilation by an abundant relative (Levin et al., 1996). Conversely, positive species interactions such as mutualism, symbioses, or facultative exchanges, may enable a greater match between the fundamental and realized niche, possibly translating to an increase in geographic distribution, a widening of habitat suitability, or an increase in local population size. These outcomes may result in an increase in the number of species occurrences, and thus a decrease in rarity.

The context in which a positive biotic interaction has the potential to decrease the rarity of a species is dependent on external environmental dynamics. Rapid changes to the conditions under which mutualisms evolved, such as landscape fragmentation, habitat alteration, and land-use conversion (Winfree et al., 2009), or introduction of non-native species (Traveset and Richardson, 2014), can jeopardize mutualistic interactions and endanger one or both partners. Local or global extirpation of one partner, for natural or anthropogenic reasons, can leave the other partner at risk of becoming rare or similarly

going extinct, especially if the relationship is obligate (Mawdsley et al., 1998). In addition, successful species interactions require that both partners share a similar cue (Warren and Bradford, 2014). Cues are context-dependent, but can include environmental changes such as seasonal warming or cooling, or changes in precipitation. Thus, shifts in seasonal patterns stimulated by climate change can place mutualisms at risk of breakdown (Tylianakis et al., 2008; Warren and Bradford, 2014), which has the potential to contribute to rarity in one or both partners. For plants that rely on mutualistic interactions for services such as nutrient uptake, pollination, protection, or seed dispersal, interspecific competition for mutualistic partners may contribute to rarity in cases where a plant is out-performed by co-occurring competitors. In these cases, a positive interaction in the form of a mutualism may be structured by a negative interaction in the form of competition (Warren et al., 2014). Furthermore, the degree to which mutualistic interactions are negatively impacted by external environmental conditions may impact plant rarity.

Identifying which abiotic and biotic factors, at both coarse and fine scales, play significant roles in regulating distributions of rare plant species is key to biological conservation. This is particularly true given the disproportionate roles rare species play in functional richness, specialization, and assemblage originality (Leitao et al., 2016). Plants in the genus *Trillium* constitute an exemplary system for investigating these questions. There are dozens of native *Trillium* species in eastern North America (Ohara, 1989), some of which are rare (range-restricted, narrowly endemic, or threatened), and some of which are geographically widespread. Often, rare (hereafter referred to as narrow endemics: those species inhabiting restricted geographic regions or specific habitat types)

and widespread species co-occur. This provides a framework for studying the effects of environmental and biotic variables on a number of endemic-widespread species pairs while controlling for environmental heterogeneity and phylogenetic relatedness. Trilliums are myrmecochores—plants that have their seeds dispersed by ants. This enables comparison of the dynamics of this facultative mutualism, and of interspecific competition for this service, in an endemic-widespread context. Ant-mediated seed dispersal has been studied from a number of ecological, demographic, and genetic perspectives (Beattie, 1985; Boyd, 1996; Kalisz et al., 1999; Lengyel et al., 2010; Warren et al., 2011; Zhou et al., 2007), but whether seed dispersal metrics differ for related, sympatric endemic and widespread myrmecochores has not been considered. Studying seed dispersal metrics in this context can reveal inequities in seed dispersal effectiveness (SDE), i.e. the number of new adults produced by the dispersal activities of a disperser (*sensu* Schupp et al., 2010), for endemic and widespread species.

This dissertation addresses questions surrounding the biotic and abiotic drivers of narrow endemism in the genus *Trillium*. In the first chapter, which was published in *The American Journal of Botany* (Miller and Kwit, 2018), I characterize quantitative and qualitative differences in SDE for endemic and widespread congeneric species pairs stemming from ant-seed dispersal in both the field and laboratory. In the second chapter, which was published in *Ecology and Evolution* (Miller et al., 2020), I investigate interspecific differences in the underlying mechanisms driving the observed differences in SDE: chemistry and morphology of seeds and elaiosomes (i.e. seed coat-derived appendages that are attractive to ants). In the third chapter, I quantify the fundamental niche and predict suitable habitat for 21 species in the genus *Trillium* in eastern North

America using SDE to target coarse-scale environmental factors shaping the distributions of endemic and widespread species.

CHAPTER 1: OVERALL SEED DISPERSAL EFFECTIVENESS IS LOWER IN ENDEMIC *TRILLIUM* SPECIES THAN IN THEIR WIDESPREAD CONGENERS

ABSTRACT

Comparing ecological attributes of endemic species with related, widespread species can reveal differences accounting for rarity. Forests of the southeastern U.S. are home to many range-restricted endemic and widespread species of *Trillium*, a genus of ant-dispersed herbs. Evidence suggests that aspects of seed-related life history stages are often correlated with plant rarity, but few studies have tested whether the process of seed dispersal differs for endemic and widespread species. To address this question, I compared aspects of seed dispersal effectiveness (SDE) for three sympatric widespread-endemic *Trillium* species pairs. I observed ant seed dispersal for *Trillium* species pairs at eight sites, recorded numbers of seeds dispersed and dispersal distances, and described disperser interactions. To test disperser preference, I presented seeds of each pair to captive colonies of *Aphaenogaster picea*, a keystone disperser. Seeds were assigned scores based on worker behavior, and I recorded proportions of seeds dispersed after 1 h and 24 h. Field observations yielded some significant within-pair differences. Ants dispersed more seeds of widespread species for all pairs, although dispersal distances did not differ. In laboratory experiments, after 24 h, ants dispersed more seeds of widespread species into nests. Endemic *Trillium* species exhibited lower overall SDE than their widespread congeners. These findings add to the list of ecological and demographic

challenges that face endemic plants when compared to common congeners. Lower SDE may negatively impact reproductive rates and the colonization of new habitats, which may contribute to patterns of endemism.

INTRODUCTION

Endemic species are key components of global biodiversity and have been used in conservation prioritization (Myers et al., 2000; Lamoreaux et al., 2006). Of the characteristics noted by Rabinowitz (1981), restricted geographic ranges (i.e. endemism) are of great interest in the field of rare species biology (Morrone, 1994). Rare species have been shown to disproportionately support the functional structure of species assemblages and are thus crucial in maintaining ecosystems (Leitao et al., 2016), so understanding the processes affecting rarity is critical for successful conservation efforts. Furthermore, determining why some species are rare while others are common is central to our understanding of the dynamics of species and communities (Brown et al., 2003), and is a fundamental question in ecology.

In flowering plants, ecological traits have been shown to correlate with species rarity. Some studies suggest that the reproductive and recruitment stages of a species' life history are most likely to explain rarity (Brown et al., 2003), while others have implicated specific factors such as limited fecundity, lack of seed-soil banks, and reduced seed dispersal (Osunkoya and Swanborough, 2001). Theoretical work suggests that the success of seed dispersal might be related to relative commonness when seed availability is in short supply, such as in species with few large seeds, in the context of a scarce seed bank,

or in species with little to no clonal abilities (Bond, 1994). Indeed, direct comparisons of narrow endemic and widespread congeners have shown that endemic species generally have lower seed availability (Fiedler, 1987; Byers and Meagher, 1997), and classic work in the field of rare species biology indicates that seed dispersal syndromes are likely correlated with aspects of relative rarity and commonness (Rabinowitz, 1978; Rabinowitz and Rapp, 1981; Pirie et al., 2000). This body of work establishes that seed-related life history stages play important roles in plant rarity, but that this hypothesis is in need of further investigation.

The use of demographic comparisons of related rare and common taxa is an established method for examining factors regulating rare plant populations (Karron, 1987; Walck et al. 1999). This method has been used to compare functional traits (Hand et al., 2017), seed production and germination (Brown et al., 2003), reproductive and ecophysiological attributes (Osunkoya and Swanborough, 2001), responses to herbivory and invasive species removal (Leege et al., 2010), predispersal seed predation (Combs et al., 2011; Combs et al., 2013), and the effects of invasive competitors (Combs et al., 2011) for rare and common taxa. In these studies, many of which compare endemic and widespread congeners, the noted demographic differences highlight the challenges that rare species face. Such a comparative framework can provide a better understanding of species characterized by narrow endemism, or those that occur in one or a few populations and thus are confined to a single domain or a few localities (Kruckeberg and Rabinowitz, 1985). Furthermore, the use of a framework that examines rare-common differences in sympatric congeneric populations may be especially insightful because confounding factors such as differences in phylogeny, ecology, and local environmental

factors can be controlled or reduced (Combs et al., 2013). Despite the applicability of this method, few studies have directly tested whether the process of seed dispersal differs for rare and common species (Pirie et al., 2000).

The framework of seed dispersal effectiveness (SDE; Schupp et al., 2010) can be used to compare dispersal-related outcomes of different plant species. This model specifies that poor seed dispersal may stem from ineffective quantitative and qualitative aspects of the seed dispersal process. For a given plant species, quantitative aspects of seed dispersal include numbers of seeds moved away from parent plants, while qualitative aspects pertain to survival through the dispersal process and facets of the deposition site that promote successful germination and establishment (Schupp, 1993). In the context of animal-mediated seed dispersal, the quantitative portion of SDE can be measured by recording numbers of seeds removed by partners during dispersal events, while qualitative aspects can be approached by identifying disperser preferences, the identity of disperser assemblages (knowing that some are better than others), the strength and resiliency of disperser assemblages (single vs multiple effective dispersal agents), and potentially through quantification of dispersal distances. When shown to be ineffective, any of the above can contribute to lower overall SDE (Schupp et al., 2010). Comparing quantitative and qualitative aspects of SDE for sympatric rare and common congeners may reveal meaningful differences in aspects of seed dispersal, especially in cases where seed dispersal for both species is mediated by the same animal partners.

In this study, I employed field observations and laboratory experiments to investigate whether four aspects of seed dispersal effectiveness—numbers of seeds dispersed, dispersal distance, disperser preference, and interactions with dispersers—

differed between sympatric populations of narrow endemic and widespread species of ant-dispersed (myrmecochorus) *Trillium* (Order Liliales, Family Melanthiaceae). My focus on seed dispersal metrics is relevant given that seed dispersal connects the reproductive phase of adult plants to the establishment of the next generation (Wang and Smith, 2002), and that the seed life history stage is important to population growth rates of myrmecochores (Horvitz and Schemske, 1995). I hypothesized that endemic trilliums would more consistently exhibit less effective metrics of seed dispersal reflecting the quantitative and qualitative aspects of SDE than their sympatric widespread congeners. Specifically, I predicted that widespread trilliums would (1) have more seeds dispersed, (2) be dispersed longer distances, (3) be preferred by the keystone seed-dispersing genus of ants in eastern North America (*Aphaenogaster*; Ness et al., 2009), and (4) have more observed interactions with effective seed dispersers than sympatric narrow endemic trilliums.

MATERIALS AND METHODS

Study species

Plants of the genus *Trillium*, a group of long-lived perennial understory herbs, encompassed my study system. Trilliums are perennial monocot rhizomatous herbs found in temperate deciduous forests of eastern Asia, western North America, and eastern North America (Freeman, 1975). Like other myrmecochores, *Trillium* seeds are characterized by the presence of elaiosomes: fleshy seed coat-borne appendages rich in lipids and other nutrients (see Turner and Frederickson, 2013). Trilliums flower from March to June, with

flowering lasting 2-4 weeks, followed by the production of a single fruit containing a few to dozens of diaspores (collective unit of a seed with its attached elaiosome). Ripe fruits dehisce and drop mature diaspores beneath the maternal plant. Members of *Trillium* are among dozens of plant species in the southern Appalachian region that share a mutualism with *Aphaenogaster* ants, a keystone seed-dispersing genus in the mesic forests of eastern North America (Ness et al., 2009; Schultz, 2014).

At least twenty-nine species of *Trillium* are native to eastern North America (Ohara, 1989), including 15 that are characterized as rare and are ranked as G3 or higher conservation priority (NatureServe, 2017). Some trilliums are also endemic in the sense that they inhabit restricted geographic regions or specific habitat types, and many of these co-occur with widespread congeners. For the purposes of this study, I used definitions of widespread species as those found throughout deciduous forests of the southeastern Appalachian region, and endemic species as those found only in restricted geographic localities or in specific habitats. Endemic/widespread status of the study species was determined using the NatureServe (2017) conservation status system coupled with maps of endemism generated by the Biota of North America Program (BONAP; Kartesz, 2015), and was confirmed by *Trillium* experts (E. Schilling, University of Tennessee, Knoxville; T. Patrick, Georgia Department of Natural Resources, personal communications).

The widespread species used in this study, *Trillium cuneatum* Raf. and *Trillium catesbaei* Elliot., are globally and locally secure (NatureServe, 2017) and considered common in the counties in which they are found (see Kartesz, 2015). The endemic species, *Trillium decumbens* Harbison., *Trillium lancifolium* Raf., and *Trillium discolor*

Wray ex Hook., are considered rare in the counties in which they are found, or, in the case of *T. lancifolium*, have patchy distributions throughout each state in their range (see Kartesz, 2015). The endemic species range from globally secure (*T. decumbens* and *T. discolor)* to globally vulnerable (*T. lancifolium*), and are vulnerable, imperiled, or critically imperiled in the states in which they occur (NatureServe, 2017).

A number of ant species are known to disperse seeds of myrmecochores in eastern North America, including the trilliums that comprise my study species (Gaddy, 1986; Zelikova and Sanders, 2011). Ants of the genus *Aphaenogaster* are the primary seed dispersal vector for most temperate ant-dispersed flora in eastern North America, approximately responsible for 74% of myrmecochore seed collection in eastern forests where encounters with *Aphaenogaster* have been reported (Ness et al., 2009). In the summer months, *Aphaenogaster* nests can be found in soil, decaying wood, under rocks, and in leaf litter. Nests are generally compact and shallow, and colony numbers range from 100 to 1,000 workers, with a single queen (Lubertazzi, 2011). While *Aphaenogaster* are general scavengers, primarily subsisting on insects and other invertebrates, members of this genus collect and move the seeds of dozens of myrmecochorous plants in eastern North America (Lubertazzi, 2011). This is especially likely in the early spring when other food items are less abundant, as has been noted for many temperate myrmecochorous systems (Thompson, 1981; Handel and Beattie, 1990; Ohkawara and Higashi, 1997; Oberrath and Bohning-Gaese, 2002; Gorb and Gorb, 2003), including those in eastern North America involving *Aphaenogaster* (Warren et al., 2014). Workers collect diaspores by grasping the elaiosome in their mandibles, and return diaspores to nests. Within nests, workers remove the elaiosome from the seed and feed it to the developing brood. Seeds

that have had the elaisome removed are then moved outside the nest, either to seemingly random locations (Canner et al., 2012), or to a refuse pile (Rico-Gray and Oliveira, 2007; C. Kwit, personal observation). *Aphaenogaster* species are considered effective seed dispersers due to their high abundance; subordinate nature and ability to rapidly discover and retrieve seeds before superior competitors interfere; predictable foraging schedules, corresponding with myrmecochore seed release; and utilization of the elaiosome without harming seeds (Warren and Giladi, 2014).

Study sites and pairwise framework

I compared seed dispersal effectiveness of endemic and widespread trilliums for three species pairs: Species Pair 1, *T. cuneatum* (widespread) and *T. decumbens* (endemic); Species Pair 2, *T. cuneatum* (widespread) and *T. lancifolium* (endemic); and Species Pair 3, *T. catesbaei* (widespread) and *T. discolor* (endemic) (Fig. 1). *T. cuneatum* was used as the widespread congener in Pairs 1 and 2 because it is ubiquitous throughout the southern Appalachian region and was sympatric with endemic species at multiple study sites. I located eight study sites in the southern Appalachian region containing sympatric populations of the three species pairs. Sites were unique patches roughly 50 m^2 and were separated from each other by at least 3 km. Seed dispersal metrics for Species Pair 1 was compared at three sites ranging in elevation from 210 - 282 m, all located in northeast Georgia (GA). Species Pair 2 was compared at three sites ranging in elevation from 218 – 252 m, all of which were located in northeast GA. Species Pair 3 was compared at two sites with elevations of 204 and 803 m, which were located on the western border between North Carolina and South Carolina (Fig. 1).

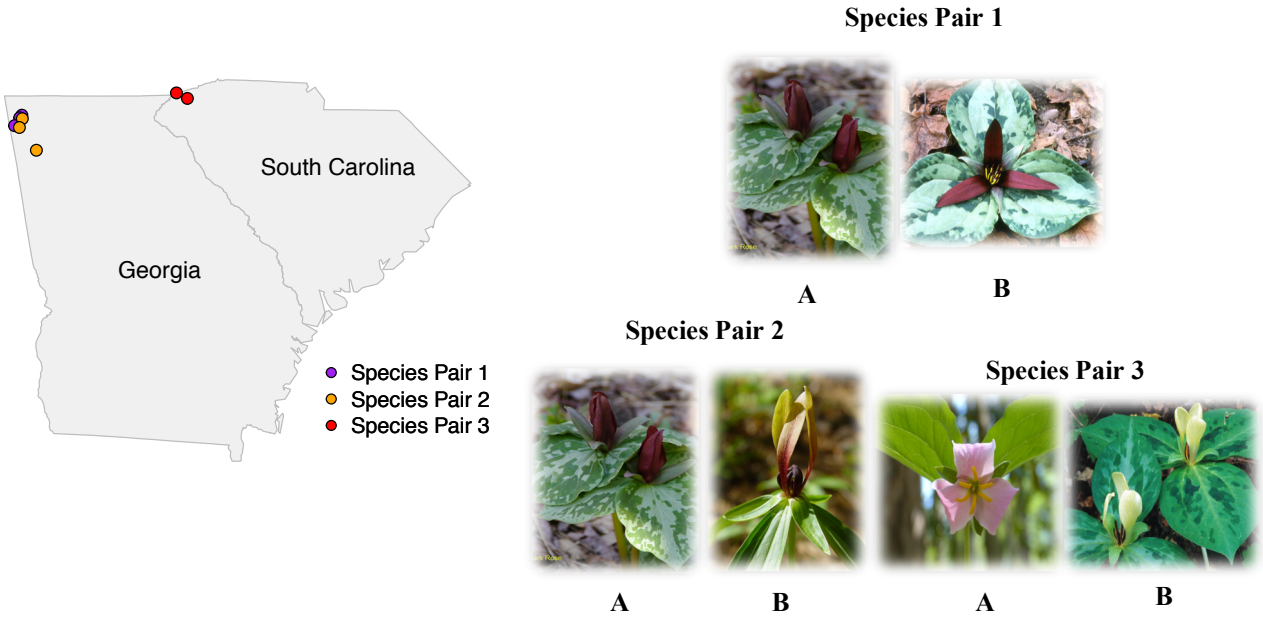


Figure 1: (Left) Site map for the three species pairs. Pair 1: 3 sites; Old Mine [34.746400, −85.346250]; Pocket Branch [34.714230, −85.378670]; Cave [34.613990, −85.456680]. Pair 2: 3 sites; Butler [34.30227, -85.12740]; Blue Hole [34.702330, −85.339170]; Dallas [34.59097, −85.38359]. Pair 3: 2 sites; Jocassee Gorges [34.95887, −82.85210]; Whitewater Falls [35.02855, −83.01711]. (Right) Photographs of the members of each species pair. A = widespread; B = endemic. Pair 1: A. *T. cuneatum;* B. *T. decumbens.* Pair 2: A. *T. cuneatum;* B. *T. lancifolium.* Pair 3: A. *T. catesbaei;* B. *T. discolor.*

All study sites were cove hardwood habitats, located in low-lying, mesic, deciduous forests with moderate to thick canopy cover. Study sites for Species Pairs 1 and 2 were located in the Limestone Valley soil province of GA, and sites for Species Pair 3 were located in the Blue Ridge soil province. Both soil provinces are characterized by mildly acidic soils high in phosphate. Sites for Species Pairs 1 and 2 were dominated by tulip poplar (*Liriodendron tulipifera*), American beech (*Fagus grandifolia*), and oak species (*Quercus* spp.). Sites for Species Pair 3 were characterized by tulip poplar, American holly (*Ilex opaca*), and eastern hemlock (*Tsuga canadensis*). All sites had

understory layers characterized by red buckeye (*Aesculus pavia*), goldenseal (*Hydrastis canadensis*), and jack-in-the-pulpit (*Arisaeama triphyllum*), as are as other myrmecochorous herbs such as trout lily (*Erythronium americanum*), violets (*Viola* spp.), bloodroot (*Sanguinaria canadensis*), wood anemone (*Anemone quinquefolia*), and sharp-lobed hepatica (*Anemone acutiloba*).

Field observations and laboratory experiments

For the three species pairs, I recorded observational metrics of seed dispersal in the field, including numbers of diaspores (hereafter referred to simply as "seeds") dispersed and dispersal distances. I collected specimens of ants observed dispersing seeds from pairs at all sites, which I identified and used to qualitatively characterize interactions with disperser assemblages for each trillium species. Following field observations, I conducted "cafeteria" experiments in the laboratory with captive *Aphaenogaster picea* Wheeler colonies to determine disperser preference for seeds of endemic and widespread species.

Number of seeds dispersed and dispersal distances

In early May 2016, I opportunistically located and flagged 30-50 flowering individuals of each trillium species at respective sites. In late May-July 2016, I randomly selected fifteen flagged individuals of each species for use in my observational study. At individual plants, I removed and manually dehisced the single ripe fruit. Fruits of species pairs at each site were generally at the same stage of ripeness at the time of these experiments, although the phenological window for peak ripeness may differ marginally for different species and individual plants even at the same environmental locations. For

logistical purposes, I removed fruits near or at peak ripeness in an effort to control when seeds were first exposed to the external environment and to identify definitive start-times for my observations. Despite the possible marginal differences in ripeness, I conducted this work at a time of year when *Trillium* are known to start dropping or dehiscing their fruits in the southeastern U. S. (early summer), such that within two weeks of my observations, fruits of maternal plants of the same species at the same sites had naturally dehisced and dropped diaspores (C. Miller, personal observation). Before starting observations, I counted and recorded the number of seeds produced per fruit (i.e. per plant, as trilliums produce one fruit per individual). I did not use this as a response variable to test the hypothesis that endemic trilliums experience lower SDE than widespread congeners, but I included it as a model covariate for assessing other response variables to account for variability in the sizes of seed piles available to ants. I then placed seeds under the maternal plant in the leaf litter and initiated an observational period of 1 h. This period was chosen because seed-dispersing ants, especially *Aphaenogaster,* exhibit the ability to remove myrmecochore seeds rapidly in nature and are adept at detecting such seeds (Ness et al., 2009), so it represented both a logistically feasible observational period and a sufficient timeframe to observe ant-seed dispersal.

During the observational period, I recorded the total number of seeds removed by ants, and the distances that ants carried seeds (for up to five seeds). Distances were then averaged per individual plant. I capped distance measurements at five seeds per individual because when ants removed more than one seed from a pile, they often did so simultaneously, which made tracking logistically challenging. Seeds were also almost always taken to the same location by the same ant colony (C. Miller, personal

observation), making additional measurements redundant. When ants were observed carrying seeds into confirmed nests, or when ants abandoned seeds in the leaf litter, the linear distance was measured in centimeters from the seed depot to the final location. If ants with seeds disappeared into the leaf litter, distance measurements were not recorded. During observations, I collected ant specimens observed interacting with seeds for use in recording trillium interactions with the ant disperser community (see below in *Disperser communities*). At the conclusion of each observational period, I left any remaining seeds at the base of the parent plant and removed all flagging.

To test whether there were significant differences in numbers of seeds dispersed by ants and in average distances of dispersal, I ran quasi-poisson generalized linear mixed effects models (PQL-GLMMs) for each species pair. PQL-GLMMs were used to deal with over-dispersion of the data (Ver Hoef and Boveng, 2007). Data from all study sites associated with a species pair were pooled, yielding sample sizes of 45 plants per species for Species Pairs 1 and 2 across three sites each, and 30 individuals per species for Species Pair 3 across two sites (Fig. 1). A power analysis required at least 30 experimental replicates, but based on permit restrictions I was unable to sample more than 15 individuals of each species at each site, so I observed plants at multiple sites to increase overall sample size. Species (endemic versus widespread) was the fixed effect in each model. I included site as a random effect in each model to account for environmental variability in the data, although it is important to note that because there is some association between sites and species pairs, this factor also includes the identity of a species within its respective category (endemic/widespread). PQL-GLMMs were conducted using the package mass (Venables and Ripley, 2002) in R (v. 3.4.0; R Core

Team, 2017). Variable significance was assessed using Type II Wald Chi-Square Tests in the package car (Fox and Weisberg, 2011) in R, which were run on full models. Variables were considered significant predictors if they yielded P-values less than $a = 0.05$.

I tested whether the number of seeds produced differed for endemic and widespread paired species using linear mixed effects models (LMEs). I included site as a random effect in each model as above. LMEs were conducted using the R package lme4 (Bates et al., 2015). Variable significance was assessed using Type II Wald Chi-Square Tests as above. All subsequent statistical analyses were conducted using these analyses and R packages unless specified otherwise.

Disperser preference

I conducted cafeteria experiments in the laboratory to determine whether disperser preference for widespread or endemic seed species was apparent by presenting seeds of species pairs simultaneously to seed-dispersing ants. Cafeteria experiments enabled me to control for spatial heterogeneity, size of seed pile, and distance to ant nests. I collected fruits from five maternal plants of each species pair at peak dehiscence times from field sites, which I transported to the University of Tennessee, Knoxville (UTK), and froze at -20 °C for 1-1.5 months. *Aphaenogaster picea* colonies were collected from the University of Virginia's Mountain Lake Biological Station (MLBS) grounds atop Salt Pond Mountain in Virginia in July 2016. This area did not contain populations of any of the trillium species used in this study, so *A. picea* colonies were naïve to my study pairs; however, at least one trillium, *T. undulatum* Willd., is found at MLBS. I collected colonies using weathered nest boxes consisting of two pine boards loosely screwed together with grooves carved into the bottom board to allow for colonization. Nest box

dimensions were 20.32 x 10.16 cm, with a width of 8.89 cm. I buried nest boxes in the leaf litter and checked them after two weeks. I then moved the colonies from the nest boxes to 25.40 x 15.24 cm plastic containers, with fine-mesh fabric glued into lids to provide airflow. I provided the colonies with artificial nests consisting of test tubes and cotton wrapped in aluminum foil for darkness. A petri dish was placed in each container, holding moistened cotton for a water source and ~0.5g of standardized ant diet, modified from the Bhatkar-Whitcomb diet (Dussutour and Simpson, 2008). I kept colonies in artificial nests indoors from July-September, under ambient light and temperature conditions. Twice a week, I replaced dried or moldy nests with fresh test tubes and cotton, and resupplied colonies with fresh water and food (Turner and Frederickson, 2013).

I conducted cafeteria experiments in September 2016 at MLBS. All colonies were starved for four days prior to the experiment to ensure equal hunger levels. To control for colony size, which I anticipated might impact dispersal behaviors of individual ant colonies, I pruned each colony four days prior to the experiment by haphazardly selecting thirty workers and one queen (eliminating additional gynes and larvae), which were transferred into new plastic containers (as in Turner and Frederickson, 2013). Out of thirty-five captive colonies, I presented seeds from Species Pair 1 to fifteen colonies; seeds from Species Pair 2 to twelve colonies; and seeds from Species Pair 3 to eight colonies. To increase statistical power, I pooled all trials and categorized the endemic species in each pair as "endemic" and the widespread species in each pair as "widespread" so that I assessed the effect of seed identity, either endemic or widespread, on ant preference.

Experiments began by setting up seed grids that were presented to colonies. One seed of either an endemic species or its widespread congener was placed at each grid point in an alternating pattern for a total of 10 seeds each (Fig. 2) to eliminate spatial clumping and to control for numbers of seeds available. Grids were drawn on the top surface of petri dish lids using a permanent marker and ants were able to access grids by climbing on top of the lids. To differentiate endemic from widespread seeds, which are similar in appearance, I marked endemic seeds with a dot of yellow paint and widespread seeds with a dot of silver paint (Elmer's Painters Yellow and Silver Paint Markers, Medium Tip). Paint marks were preliminarily tested for effects on ant seed dispersal and were shown to have no effect (unpublished data).

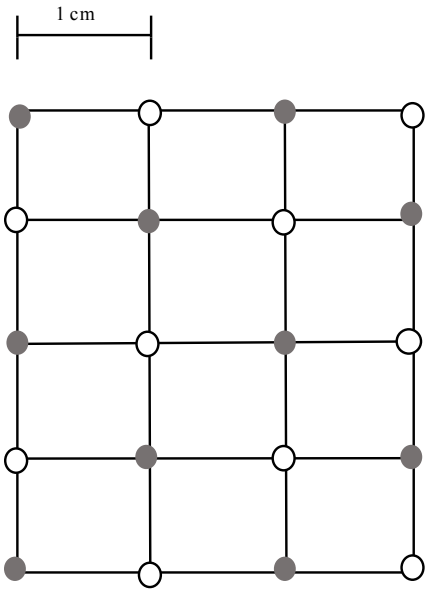


Figure 2: Design for placement of endemic and widespread seeds on grids for cafeteria experiments. Gray dots and odd numbers indicate widespread seed species; white dots and even numbers indicate endemic seed species.

Seed grids were placed into colony containers 20 cm from the nest opening. In an observational period of 1 h (chosen to resemble my field observations), I recorded numbers of endemic and widespread seeds taken into nests, and numbers of endemic and widespread seeds removed from grids but not taken into nests. Following the observational period, I calculated the proportions of endemic and widespread seeds (out of 10) that fell into each of these categories. Colonies were revisited after 24 h, which enabled colonies to fully explore and discover seed piles, and the above measurements were repeated. Binomial generalized linear models (GLMs) were performed for data collected after 1 h and 24 h to determine if seed identity was a significant predictor of either the proportion of seeds dispersed into nests, or the proportion of seeds removed from the grid but not taken into nests. The effects of seed identity were assessed using Type II Wald Chi-Square Tests run on GLMs and effects were considered significant if tests yielded a P-value less than $a = 0.05$.

Following the initial 1 h observational period, individual seeds were assigned a behavioral score, 0-3, according to the level of ant interaction the seed experienced (Culver and Beattie, 1978). A score of 0 meant the seed was never discovered by ants; a score of 1 indicated discovery and handling, but no movement; a score of 2 indicated removal from the grid, but not dispersal into the nest; and a score of 3 indicated dispersal into the nest. I counted the number of observed behaviors expressed by ants in each behavioral category for each seed, and compiled these counts into a contingency table with n = 700 seeds (350 endemic, 350 widespread). Observed values were compared to expected values using a Likelihood Ratio Test. Significance was considered at the $a =$

0.05 level. The contingency table was created and analyzed using SAS Enterprise Guide (v.7.1).

During the initial observational period, numbers of *A. picea* workers handling endemic and widespread seeds was also recorded. At the end of 1 h, I calculated the average number of workers handling widespread and endemic seeds for each colony, and averages were rounded to the nearest whole number. I rounded averages because seeds cannot be handled by a fraction of a worker. This resulted in each colony receiving a single average worker integer for widespread and endemic seeds. Because of the relatively low number of ant-seed interactions observed during the initial 1 h periods (see *Disperser preference* below), the data were heavily skewed, and transformations did not yield normality. I therefore used a Poisson-distributed GLM to determine if seed identity was a significant predictor of average number of workers handling seeds. Effects were assessed for significance using a Type II Wald Chi-Square Test as above.

Disperser communities

To qualitatively describe and assess whether the endemic and widespread *Trillium* species pairs interacted differently with disperser assemblages, I haphazardly collected specimens of ants witnessed dispersing seeds of each trillium species at each study site during field observations. Specimens were collected based on the criteria that they were amongst several workers observed removing diaspores from a given seed pile. I collected at least one specimen at each maternal plant observed in this study, given that the plant experienced ant activity during the observational period. I attempted to collect at least one specimen of each distinguishable ant species visiting each plant. Mouth aspirators were used to collect specimens, which were then transferred to microcentrifuge vials and

stored in 95% EtOH. Specimens (n = 54) were stored at room temperature for five months and were then identified to species using the Mississippi State University Entomology Museum dichotomous key (MacGown, 2014). Identifications were independently confirmed by a second observer.

RESULTS

Number of seeds dispersed and dispersal distances

My observations of numbers of seeds dispersed by ants in the field yielded results consistent with lower quantitative SDE for endemic trillium species. I found that the number of seeds dispersed differed significantly for Species Pair 1 ($X^2(1) = 18.780$, $P < 0.001$), with the widespread *T. cuneatum* having greater numbers of seeds dispersed than the endemic *T. decumbens* (Fig. 3a.). The number of seeds available to ants (i.e. the number of seeds produced by an individual parent plant) was included as a covariate in each model, but was not a significant predictor of number of seeds dispersed for Species Pair 1, nor was the interaction between number of seeds available and species ($P > 0.05$, respectively). Species Pair 2 had significantly greater numbers of seeds dispersed for the widespread *T. cuneatum* than the endemic *T. lancifolium* (Fig. 3a.), with significant effects of both the number of seeds available and the interaction between number of seeds available and species ($X^2(1) = 8.344$, $P < 0.001$; $X^2(1) = 14.157$, $P < 0.001$, respectively). Species identity alone was not a significant predictor ($X^2(1) = 1.790$, $P = 0.181$), so I attribute the difference in the number of seeds dispersed for the two species primarily to the effects of the number of seeds available, in conjunction with an additive species

effect. For Species Pair 3, I found significantly greater numbers of seeds dispersed for the widespread *T. catesbaei* than the endemic *T. discolor* (Fig. 3a.). Species identity and the number of seeds available were significant predictors ($X^2(1) = 8.694$, $P < 0.001$; $X^2(1) = 5.0312$, $P = 0.025$, respectively), but their interaction was not ($P > 0.05$). My measurements of seed dispersal distances by ants in the field were not consistent with qualitative differences in SDE for widespread and endemic *Trillium* species pairs. None of the models indicated significant differences in linear seed dispersal distance from maternal plants for endemic and widespread species in any of the pairs ($P > 0.05$ in all cases; Fig. 3b.).

Disperser preference

I found no significant effect of seed identity (endemic vs widespread) on the proportion of seeds taken into nests after 1 h ($X^2(1) = 0.891$, $P = 0.345$; Fig. 4a.). There was also no significant difference in the proportion of endemic and widespread seeds removed from the grid, but not taken into the nest after 1 h ($X^2(1) = 0.005$, $P = 0.941$; Fig. 4b.). It should be noted that I observed very little ant activity of any kind during initial 1 h observational periods in the laboratory. I attribute the discrepancy between ant activity during 1 h observations in the field and laboratory to several possible factors. First, I conducted field observations during peak diurnal and annual foraging times in a natural setting, so ants were highly active in the field; the same was not necessarily true in the lab. Second, ants from many different colonies were active in the field, whereas in the lab, I was observing workers from one pruned colony per trial.

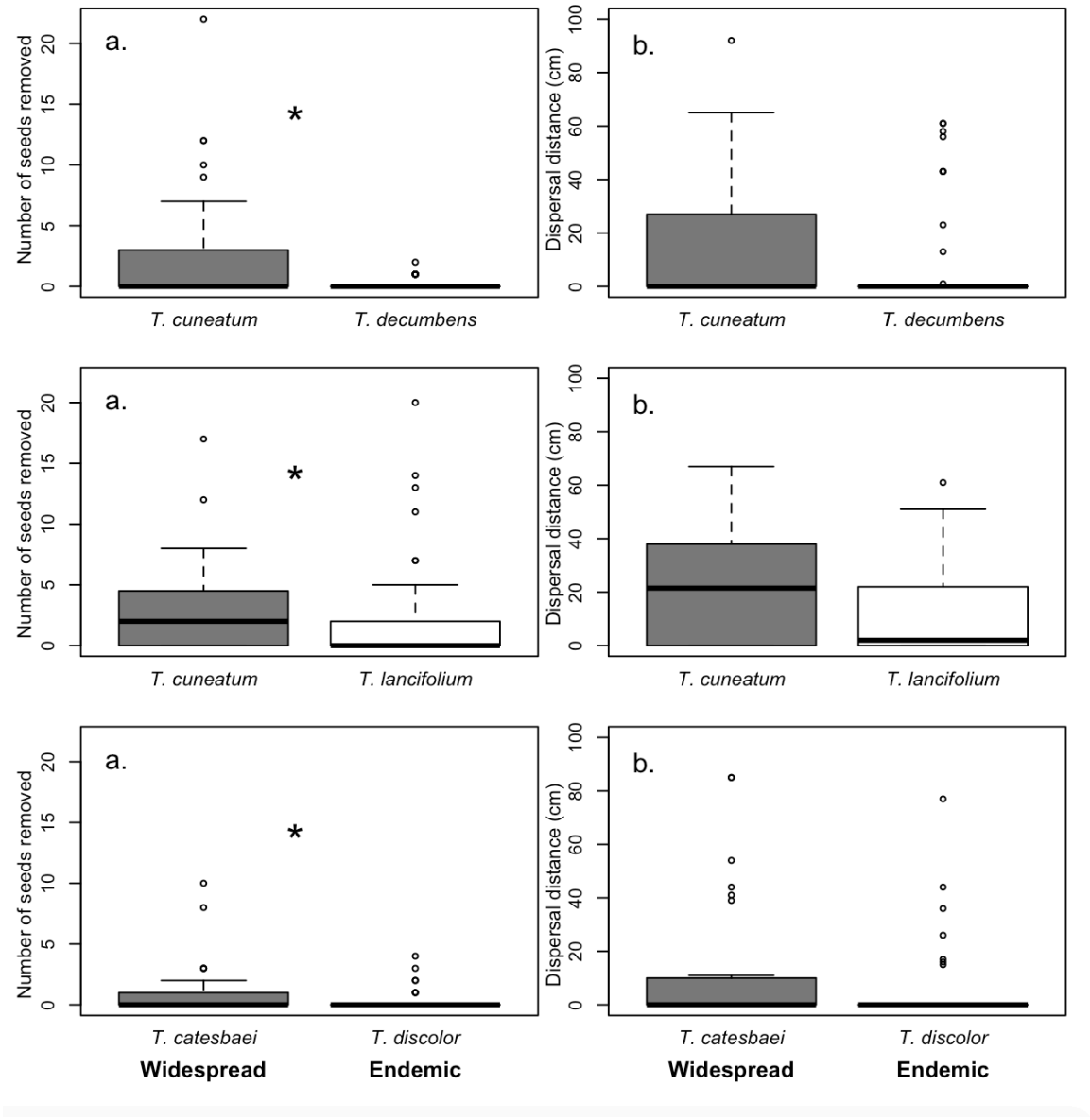


Figure 3: Box plots illustrating median, interquartile range, and outliers for (a.) number of seeds removed and (b.) dispersal distance (cm) for three *Trillium* species pairs. Widespread species are gray; endemic species are white. Asterisks denote significant effects.

After 24 h, which provided captive colonies enough time to discover and interact with seeds in an artificial environment, I found that workers dispersed significantly greater proportions of the seeds of widespread trilliums into nests than endemic congeners ($X^2(1) = 12.380$, $P < 0.001$; Fig. 4a.). I did not find a significant difference in the proportion of seeds removed from the grid but not taken into the nest after 24 h for widespread and endemic trilliums ($X^2(1) = 1.070$, $P = 0.301$; Fig. 4b.), although a trend shows that ants removed a higher proportion of widespread trillium seeds from the grid than endemic congeners. Results after 24 h suggest that *Aphaenogaster picea* prefer the seeds of widespread trillium species over their endemic congeners, a finding that is consistent with lower quantitative and qualitative SDE for endemic species.

Lower SDE for endemic species is also supported by the results of seed behavior scores recorded after 1 h observations, which indicated that seed identity (endemic vs widespread) was significantly related to numbers of seeds scored for each of the four behavioral categories ($X^2(3) = 12.150$, $P = 0.007$). Endemic seeds were more likely to be scored with a "0," indicating no discovery by ants, as indicated by relative row percentages ("0" made up 74.29% of total behavioral scores for endemic seeds, whereas "0" made up only 71.14% of total behavioral scores for widespread seeds). More widespread seeds were scored with a "1", suggesting that ants discovered and handled widespread seeds more often (22.57% of total seed scores and 21.43% of total seed scores, respectively). Endemic seeds were more likely to be moved from the seed grid and abandoned in a location other than the ant nest ("2" made up 3.43 and 1.71% of total

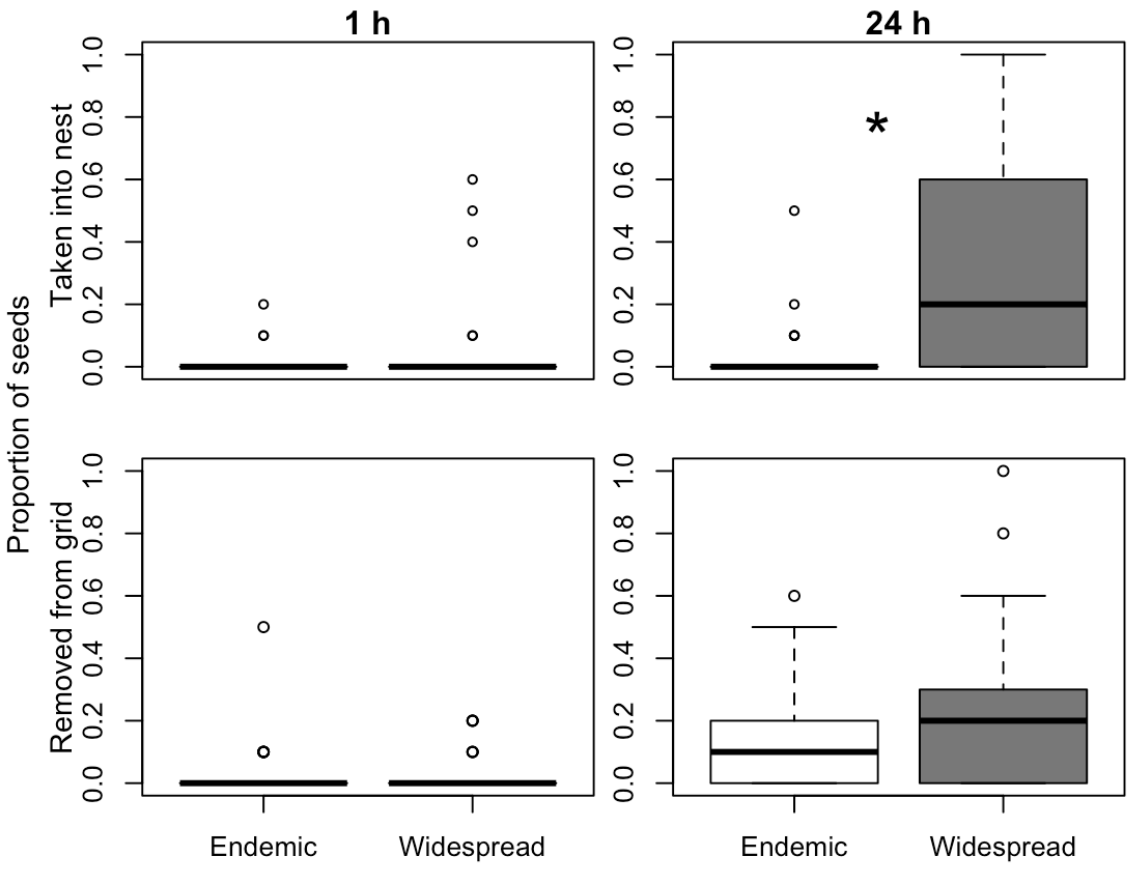


Figure 4: Box plots illustrating median, interquartile range, and outliers for (a.) proportions of seeds taken directly into ant nests by workers in 1 h and 24 h, and (b.) proportions of seeds taken off the seed grid but not into nests in 1 h and 24 h. Asterisks denote significant effects.

behavioral scores for endemic and widespread seeds respectively), and widespread seeds were more likely to be taken directly into nests than endemic congeners (4.57 and 0.86% of total behavioral scores for widespread and endemic seeds, respectively).

I found a non-significant effect of seed identity on average numbers of ant workers handling seeds ($X^2 = 0.022$, $P = 0.882$), suggesting no difference in average numbers of workers handling seeds of endemic and widespread species.

Disperser communities

Among 54 ant specimens collected for the study system, 12 ant species were identified, representing seven genera in the family Formicidae across study sites. *Aphaenogaster rudis* had the highest number of interactions with four of the five *Trillium* species (35), with *Aphaenogaster picea* having the second highest number (6), followed by the invasive Asian Needle Ant, *Brachyponera chinensis* Emery (3). All other ant species were represented by a single observed interaction. *Trillium cuneatum* was the trillium species showing the highest number of interactions in the network (31), followed by *T. lancifolium* (10), *T. catesbaei* (7), *T. decumbens* (5), and *T. discolor* (1). See Table 1 for a complete list of identified ant species and the number of interactions observed for each *Trillium* species.

As members of *Aphaenogaster* are the most effective dispersers in eastern North America and are thus linked to higher quality SDE, I emphasize here the observed number of interactions between widespread and endemic species and *Aphaenogaster.* Widespread trilliums had a higher number of interactions with members of *Aphaenogaster* than endemic trilliums (32 interactions and 9 interactions, respectively). Although interactions with ant dispersers were primarily characterized by *Aphaenogaster,* widespread species also were observed to interact with other non-keystone dispersing ants (6 out of 38 total interactions), including *Camponotus chromaoides* Bolton*, Formica pallidefulva* Latreille*,* and *Nylanderia faisonensis* Trager. In one instance, the widespread *Trillium cuneatum* was found to interact with the invasive *B. chinensis,* which is an ineffective seed disperser that is known to disrupt the structure of native ant assemblages in eastern North America, and specifically has been shown to displace native

Aphaenogaster species in invaded sites (Rodriguez-Cabal et al., 2012). The widespread *T. catesbaei* was observed to only interact with members of *Aphaenogaster*.

Roughly half of the interactions observed for endemic *Trillium* species were with non-*Aphaenogaster* species (7 out of 16 total interactions), including members of *Camponotus* and *Crematogaster*, which are less-effective seed dispersers than *Aphaenogaster* species (Culver and Beattie, 1978). The endemic *T. decumbens*

Table 1: Ant specimens collected during field observations and associated *Trillium* species. *N* = number of specimens collected for each ant–plant pair.

Ant species	Plant species	Site	*N*
Aphaenogaster rudis	*T. catesbaei*	Whitewater Falls (North Carolina)	3
A. rudis	*T. cuneatum*	Old Mine, Pocket Branch, Cave, Dallas, Butler (Georgia)	24
A. rudis	*T. decumbens*	Old Mine, Pocket Branch, Cave (Georgia)	3
A. rudis	*T. lancifolium*	Dallas (Georgia)	5
A. picea	*T. catesbaei*	Whitewater Falls (North Carolina)	4
A. picea	*T. cuneatum*	Old Mine (Georgia)	1
A. picea	*T. lancifolium*	Dallas (Georgia)	1
Camponotus americanus	*T. decumbens*	Cave (Georgia)	1
C. chromaiodes	*T. cuneatum*	Pocket Branch (Georgia)	1
C. pennsylvanicus	*T. lancifolium*	Dallas (Georgia)	1
Crematogaster ashmeadi	*T. decumbens*	Old Mine (Georgia)	1
C. pilosa	*T. lancifolium*	Dallas (Georgia)	1
C. pinicola	*T. lancifolium*	Butler (Georgia)	1
Formica pallidefulva	*T. cuneatum*	Old Mine (Georgia)	1
Temnothorax torrei	*T. lancifolium*	Butler (Georgia)	1
Nylanderia faisonensis	*T. cuneatum*	Old Mine (Georgia)	1
Brachyponera chinensis	*T. cuneatum*	Boat Ramp (Georgia)	3
B. chinensis	*T. discolor*	Jocasse Gorges (South Carolina)	1

interacted primarily with *A. rudis,* and had one additional observed interaction with *Camponotus americanus* Mayr. The endemic *T. lancifolium* interacted with both *A. rudis* and *A. picea,* as well as four other native species that are non-keystone dispersers. The endemic *T. discolor* was observed to interact with the invasive *B. chinensis.*

DISCUSSION

This study yielded results that indicate lower quantitative and qualitative seed dispersal effectiveness in some respects for endemic *Trillium* species in comparison to their sympatric, widespread congeners. For three endemic-widespread species pairs across the southern Appalachian region, endemic species (*T. decumbens, T. lancifolium*, and *T. discolor*) had significantly lower numbers of seeds dispersed by ants than their widespread congeners (*T. cuneatum* and *T. catesbaei*), indicating the initial stage of the quantitative portion of SDE is lower for the endemic species. The results of lab-based choice experiments provide a plausible explanation for the field results. After 1 h of exposure to seeds in the laboratory, *Aphaenogaster picea* behavior was significantly related to seed identity, with widespread seed species experiencing higher levels of discovery, handling, and dispersal by ants into nests than endemic species. After 24 h of exposure, *A. picea* workers dispersed significantly higher proportions of widespread seed species into ant nests. If *Aphaenogaster* species display a consistent preference for the seeds of widespread trilliums over those of their sympatric, endemic congeners in the field, this would result in fewer numbers of seeds dispersed for endemic species. In this case, preference for seeds of widespread species by a keystone group of ants

(*Aphaenogaster*) could also contribute to lower qualitative SDE for endemic species, with fewer of the latter seeds being dispersed at all, or more seeds being available for lower-quality non-*Aphaenogaster* ant species for dispersal.

This is consistent with what I noted through qualitative comparisons of observed ant-trillium interactions in the field, where widespread trilliums appeared to interact more often with members of *Aphaenogaster* than endemic trilliums. Although primarily dispersed by members of *Aphaenogaster,* widespread species had a few connections with other non-keystone dispersing ants. Roughly half of the interactions of endemic species were with non-*Aphaenogaster* species, potentially indicating that endemic trilliums are more often dispersed by less-efficient dispersers than their sympatric, widespread congeners. Widespread species also had a greater total number of interactions than endemic species, providing further support for the hypothesis that endemic species experience lower qualitative SDE stemming from the frequency and composition of interactions with members of the disperser community. When coupled with the quantitative disadvantages that typify SDE of endemic trilliums in my study, this qualitative aspect of SDE may further accentuate the overall SDE experienced by endemic species. However, the scope of the conclusions drawn from ant-trillium interactions is limited, given the constraints of the sampling design. The number of specimens collected was low compared to the total number of observed interactions in the field, due to my reluctance to interfere with otherwise unmanipulated observations of seed dispersal. Additionally, the higher number of interactions recorded for *Trillium cuneatum* may be due to its inclusion in two species pairs. I observed seed dispersal for this species at six sites, whereas the others were observed at three sites (*T. decumbens*

and *T. lancifolium)* or two sites (*T. catesbaei* and *T. discolor).* Finally, my study system-wide comparison of interactions with the disperser community must be considered in light of the geographic separation of the three species pairs, which may explain some of the observed differences. Ideally, interactions with disperser communities would have been compared only within sympatric species pairs; however, due to the low number of specimens collected, some of these comparisons would not have been meaningful. Thus, although comparisons across the entire study system yielded interesting and potentially important insights, some of the results may be an artefact of overlapping distributions of particular ant species and *Trillium* species as reflected by the choice of study sites.

It is possible that the widespread *Trillium* species experienced significantly greater numbers of seeds dispersed in the field than their endemic congeners due to some non-random distribution of individual widespread trilliums and *Aphaenogaster* nests. Attractive myrmecochore seeds will ideally be collected by workers of the nearest nests, and be dispersed to those nests until the colony is satiated; keystone dispersers generally remove > 75% of available seeds (Zelikova et al., 2008; Ness et al., 2009; Warren et al., 2010). In this case, dispersal distances and numbers of seeds dispersed will be dictated in part by the spatial distribution of nests in relation to the spatial distribution of plants. It is unlikely that the locations of plants will be correlated with the locations of ant nests due to past ant-seed dispersal events, because *Aphaengoaster* colonies move nests in the spring, and frequently again in the summer (Lubertazzi, 2011). It is possible that *Aphaenogaster* colonies and these particular widespread *Trillium* species have correlated microhabitat preferences, such as lower soil moisture (Warren et al., 2010), that make their proximity more likely, thus explaining higher SDE for these species. Because I did

not record locations of *Aphaenogaster* nests nor any microhabitat characteristics, I was unable to test this hypothesis. However, many of the results do not support this explanation. First, dispersal distances for widespread and endemic species in the field were not significantly different, suggesting that there was not a consistent trend for widespread trilliums to be located closer to *Aphaenogaster* nests. Second, the results of laboratory experiments, which showed *A. picea* preferred the seeds of widespread species in the absence of differing proximities to ant nests and microhabitat characteristics, make the above explanation unlikely.

Aspects of diaspores (seeds, elaiosomes, or their combination) may explain the findings of lower SDE for endemic trilliums. Because ant-seed foraging is stimulated by the presence of oleic acid, which elicits a carrying response in ants, differences in quantities of oleic acid present in elaiosomes can affect disperser preference (Boulay et al., 2006; Turner and Frederickson, 2013), and thus dispersal success (Boulay et al., 2007). Total lipid content (Boulay et al., 2006) is a representative measure of elaiosome nutritional quality and may also play a role in shaping ant preference, as seed dispersing ants have displayed the ability to detect and preferentially disperse diaspores with greater nutritional rewards (Leal et al., 2014). Physical characteristics, such as elaiosome-seed mass, length, and width ratios, have also been shown to explain preference in seed dispersing ants (Mark and Olesen, 1996), and generally these metrics are positively correlated with seed removal rates and dispersal distances (Davidson and Morton, 1981; Oostermeijer, 1989; Ruhren and Dudash, 1996; Gorb and Gorb, 2003; Peters et al., 2003; Bas et al., 2009; Warren et al., 2014). Gunther and Lanza (1989) specifically noted this for members of *Trillium*. Additionally, some ant-dispersed plants display cheating

behaviors by producing large quantities of cheap signaling compounds and provide little in the way of nutritional reward to ants (Turner and Frederickson, 2013; Pfeiffer et al., 2010). It is possible that the widespread species investigated in this study are characterized by physical and/or nutritional profiles that are more attractive; that they display cheating behavior more often than endemic species; or that they exhibit a combination of these characteristics leading to the observed preference in *A. picea.* Exploring the chemical and physical variability in seed traits for co-occurring, related rare and common myrmecochores will contribute to our understanding of this mutualism in the context of plant rarity, and may establish a mechanism explaining disperser preference for certain myrmecochore species. In turn, this may be linked to lower SDE for myrmecochores that produce diaspores and elaiosomes that are less attractive to ant partners.

Other studies have found that rare species are characterized by poorer measures of ecological and demographic success than widespread relatives. For example, studies have correlated plant rarity with poor competitive ability (Rabinowitz and Rapp, 1981; Prober and Austin, 1993; Baskin et al., 1997; Walck et al., 1999; Lynch et al., 1999); poor ability to survive disturbances before reaching reproductive maturity (Enright et al., 1996); and strong effects of habitat loss and fragmentation on seed-set and seed viability through impacts of pollinator availability (Cunningham, 2000) and pollen limitation (Brown and Kephart, 1999). The finding that endemic *Trillium* species may experience lower SDE than common, geographically widespread relatives adds another dimension to the literature comparing ecological attributes of rare and common, and specifically endemic and widespread, plant species.

The seed life history stage has the potential to contribute to patterns of endemism. Establishment limitation, which may be tied to aspects of seed dispersal, has been documented in some endemic plants (Dinsdale et al., 2000). In addition, endemic plants in the family Fabaceae have been noted to be more susceptible to predispersal seed predation than widespread, sympatric congeners, as well as being characterized by lower seed production (Combs et al., 2013). Similar findings elsewhere (Pirie et al., 2000) suggest that endemic plant species may be both intrinsically and extrinsically seed-limited, a conclusion potentially corroborated by my results. I found that some endemic trilliums exhibit lower fecundity by producing fewer seeds than their common congeners; a result that may or may not be consistent with lower intrinsic seed-limitation, and which is in need of further testing, especially in light of the life-history strategies of narrow endemics. I also note the exception of Species Pair 3 to this conclusion, which may be credited to *T. catesbaei* being a declinate-flowered, pedicellate species and *T. discolor* being sessile (Ohara and Utech, 1986; Ohara and Utech, 1988). My results also provide a potential basis for extrinsic seed-limitation in endemic *Trillium* species as a consequence of poorer quantitative and/or qualitative aspects of the ant-seed dispersal mutualism, resulting in lower SDE. The latter result is, to my knowledge, novel in the context of rare-common plant comparisons.

Although I did not address whether post-seed dispersal demography differed for endemic and widespread trilliums, these results on seed dispersal patterns warrant further work in this area. In this study system, endemic trilliums had fewer seeds dispersed relative to sympatric widespread congeners, thereby having fewer seeds' elaiosomes removed—a seed treatment that has been shown to positively impact seed germination

(Horvitz and Beattie, 1980; Peternelli et al., 2003; Cuautle et al., 2005; Salazar-Rojas et al., 2012; Prior et al., 2014). Whether endemic species with low quantitative and qualitative aspects of SDE relative to other plant species experience lower emigration and colonization of new sites at local scales (e.g., meters) or larger regional scales (e.g., kilometers) is far from clear. In my study system, this may be overcome by other, long-distance forms of seed dispersal, such as dispersal by yellow jackets (Jules, 1996; Zettler et al., 2001; Bale et al., 2003) or white-tailed deer (Vellend et al., 2003; Myers et al., 2004). Regardless, successful dispersal of sufficient seeds to safe sites is of crucial importance to the persistence of metapopulations (Washitani et al., 1997), and therefore should be considered in conservation efforts for endemic plant species.

CHAPTER 2: EFFECTS OF SEED MORPHOLOGY AND ELAIOSOME CHEMICAL COMPOSITION ON ATTRACTIVENESS OF FIVE *TRILLIUM* SPECIES TO SEED-DISPERSING ANTS

ABSTRACT

Morphological and chemical attributes of diaspores in myrmecochorous plants have been shown to affect seed dispersal by ants, but the relative importance of these attributes in determining seed attractiveness and dispersal success is poorly understood. I explored whether differences in diaspore morphology, elaiosome fatty acids, or elaiosome phytochemical profiles explain the differential attractiveness of five species in the genus *Trillium* to eastern North American forest ants. Species were ranked from least- to most-attractive based on empirically-derived seed dispersal probabilities in my study system, and I compared diaspore traits to test the hypotheses that more attractive species will have larger diaspores, greater concentrations of elaiosome fatty acids, and distinct elaiosome phytochemistry compared to the less attractive species. Diaspore length, width, mass, and elaiosome length were significantly greater in the more attractive species. Using gas-chromatography mass spectrometry, I found significantly higher concentrations of oleic, linoleic, hexadecenoic, stearic, palmitoleic, and total fatty acids in elaiosomes of the more attractive species. Multivariate assessments revealed that elaiosome phytochemical profiles, identified through liquid-chromatography mass spectrometry, were more homogeneous for the more attractive species. Random forest classification models (RFCM) identified several elaiosome phytochemicals that differed significantly among species. Random forest regression models revealed that some of the

compounds identified by RFCM, including methylhistidine (α-amino acid) and d-glucarate (carbohydrate), were positively related to seed dispersal probabilities, while others, including salicylate (salicylic acid) and citrulline (L-α-amino acid), were negatively related. These results supported the hypotheses that the more attractive species of *Trillium*—which are geographically widespread compared to their less-attractive, endemic congeners—are characterized by larger diaspores, greater concentrations of fatty acids, and distinct elaiosome phytochemistry. Further advances in my understanding of seed dispersal effectiveness in myrmecochorous systems will benefit from a portrayal of dispersal-unit chemical and physical traits, and their combined responses to selection pressures.

INTRODUCTION

The frequency and quality of seed dispersal by ants in the mutualism known as myrmecochory is influenced by a number of plant and disperser attributes. From the perspective of plants, traits characterizing the diaspore—the dispersal unit of a myrmecochorous plant, comprised of the seed and the elaiosome (a fleshy seed-coat appendage rich in lipids, amino acids, and monosaccharides; Brew et al., 1989; Fischer et al., 2008, Boieiro et al., 2012)— influence seed dispersal effectiveness (Schupp, 1993; Schupp et al., 2010). Present in all myrmecochore species (Lisci et al, 1996), elaiosomes attract ants to pick up and carry diaspores back to the nest (Van der Pijl, 1982); much work has documented this for beneficial and effective seed-dispersing ants such as those in the genera *Aphaenogaster* (Asia and North America; Ness et al., 2009; Takahashi and

Itino, 2015), *Rhytidoponera* (Australia; Gove et al., 2007), *Myrmica,* and *Formica* (Europe; Gorb and Gorb, 1995; Manzaneda and Rey, 2009).

Diaspore size, elaiosome biomass, and the elaiosome-seed mass ratio are all known to influence the attractiveness of diaspores to ants (Mark and Olesen, 1995; Leal et al., 2014). Interspecific variation in elaiosome size and elaiosome-seed mass ratios is substantial and these traits are often correlated (Edwards et al., 2006; Reifenrath et al., 2012; Leal et al., 2014; Levine et al., 2019), meaning any interpretation of the effects of one trait should be considered in light of the other. In many systems, effective ant dispersers prefer the seeds of species that produce heavier elaiosomes (Ness et al., 2004; Takahashi and Itino, 2015; Levine et al., 2019), although Manzaneda et al. (2009) found that ants exerted significant selection on seed size, but not on elaiosome size when both traits were considered simultaneously. The nutritional quality of elaiosomes also likely plays a role in ant preference (Fischer et al., 2008; Caut et al., 2013), given that ants disassemble the diaspore and feed the elaiosome to their larvae upon return to the nest (Fischer et al., 2005). However, diaspore morphology and elaiosome nutritional rewards may not be the most important determinants of the first, critical step of the dispersal process: diaspore retrieval and dispersal to the nest. Instead, ant foraging behavior is strongly influenced by non-nutritive elaiosome chemical cues (Nelson et al., 2019).

Elaiosome lipid chemistry has been well described (Bresinsky, 1963; Skidmore and Heithaus, 1988; Brew et al., 1989; Kusmenoglu et al., 1989; Lanza et al., 1992; Hughes and Westoby, 1994; Lisci et al, 1996; Boulay et al., 2006, 2007; Gammans et al., 2006; Fischer et al., 2008; Leal et al., 2014; Wu et al., 2019) and is known to influence both the quantitative and qualitative aspects of seed dispersal effectiveness (*sensu*

Schupp, 1993; Schupp et al., 2010) by ants. Seed-dispersing ants display a preference for food items containing high concentrations of oleic acid, an important signaling compound that stimulates foraging behavior, and thus select for seeds containing high concentrations of this compound (Brew et al., 1989; Lanza et al., 1992; Boulay et al., 2006, 2007; Fischer et al., 2008; Pfeiffer et al., 2010; Turner and Frederickson, 2013). Other fatty acids, including palmitic acid, stearic acid, linoleic acid, and linolenic acid may also be attractive to ants and help initiate seed dispersal (O'Dowd and Hay, 1980).

Non-lipid chemical cues stemming from the broader phytochemical profiles of elaiosomes may also govern ant preference in myrmecochorous systems. Non-lipid elaiosome phytochemistry likely includes diverse nutritional rewards in addition to other plant primary and secondary metabolites. Such compounds may function to mediate interactions between myrmecochore seeds and other species in their environment, including mutualistic partners and antagonists such as granivores or pathogenic microbes. For example, capsaicin, a secondary metabolite responsible for heat in chilli (*Capsicum,* Solanaceae) fruits, selectively deters vertebrate predators and microbial pathogens without deterring effective seed-dispersing birds (Tewksbury and Nabhan, 2001; Tewksbury et al., 2008; Fricke et al., 2016); it is possible that secondary metabolites present in elaiosomes have similar mediating effects for myrmecochorous plants. Although we do not know which non-lipid phytochemical compounds are important to ant seed dispersers, we do know that ants utilize chemical reception and therefore could be affected by a suite of seed phytochemicals (Nelson et al., 2019).

Because elaiosome chemistry is under selection by ant dispersers (Boulay et al., 2006; 2007), it is possible that ants exert stabilizing selection on the entire phytochemical

profile. Such stabilizing selection on the chemotype of plant-derived food bodies by ant mutualists has been shown for myrmecophytic symbioses in the tropics (see the review by Mayer et al., 2014). Extra-floral nectaries (EFN) and pearl bodies produced by myrmecophytic plants are rich in carbohydrates/amino acids and lipids, respectively (Rickson, 1976; Heil et al., 1998, 2004; Fischer et al., 2002; Bluthgen et al., 2004; Webber et al., 2007; Buono et al., 2008; Gonzalez-Teuber and Heil, 2009; Paiva et al., 2009; Heil, 2011; Shenoy et al., 2012), and these food resources are typically not as attractive to generalist ants as they are to the plants' specific ant mutualists (Davidson et al., 1991; Heil et al., 1998, 2004; Bluthgen and Fiedler, 2004; Gonzalez-Teuber and Heil, 2009). Furthermore, ant preferences for specific chemical compounds in EFN are highly species-specific, suggesting mutualists exert strong stabilizing selection on EFN chemotypes. Although the myrmecochory is a facultative mutualism rather than an obligate mutualism as in tropical myrmecophytic systems, it is possible that seed-dispersing ants nevertheless exercise a similar stabilizing selection on elaiosome phytochemical profiles such that a certain narrow chemotype is defined. This hypothesis has not been tested in myrmecochorous systems; in particular, studies are lacking that characterize the metabolomic profiles of elaiosome phytochemistry and assess whether single compounds or entire phytochemical profiles affect diaspore retrieval and dispersal to the nest by ants.

Despite the fact that three main aspects of diaspores (morphological, nutritional, chemical) have been linked to seed dispersal by ants, rarely have multiple diaspore traits been investigated in the same study (but see Leal et al., 2014), and the most important mechanisms governing ant preference remain unclear. I explore whether differences in

diaspore morphology, elaiosome fatty acids, or elaiosome phytochemical profiles explain the differential attractiveness of diaspores of plants in the genus *Trillium* (Order Liliales, Family Melanthiaceae) to ants, primarily of the keystone seed-dispersing genus *Aphaenogaster.* Although a number of ant species disperse seeds of myrmecochores in eastern North America (Gaddy, 1986; Zelikova et al., 2011), *Aphaenogaster* ants are the primary seed dispersal vector for most temperate ant-dispersed flora in this region, and are responsible for approximately 74% of myrmecochore seed collection in forests where *Aphaenogaster* have been reported (Ness et al., 2009).

My study system is comprised of five eastern North American *Trillium* species, two of which are geographically widespread and three of which are narrowly endemic; herein, pairs of widespread and endemic species co-occur in multiple sites across the southern Appalachian region of the United States. In this system, *Aphaenogaster* ants display a preference for dispersing the diaspores of the geographically widespread trilliums compared to their co-occurring, endemic congeners (Miller and Kwit, 2018). Using field-based observations of seed dispersal probabilities from Miller and Kwit (2018), I ranked each of the five study species from least attractive to most attractive (see *Materials and Methods* below) and use this scale of attractiveness as a comparative framework for exploring diaspore morphology and elaiosome chemistry throughout the present study.

My first objective was to quantify interspecific differences in diaspore morphology among my study species. I predicted that the more attractive species of *Trillium* would have larger diaspores overall, and higher elaiosome mass and elaiosome-seed mass ratios than less attractive species. My second objective was to assess

interspecific differences in concentrations of elaiosome fatty acids. I predicted that concentrations of key fatty acids, including oleic acid, would be higher in the elaiosomes of the more attractive species. My third objective was to assess interspecific differences in elaiosome phytochemical profiles. I predicted that elaiosome phytochemical profiles of the more attractive species would be more tightly clustered in multivariate space than those of less attractive species, reflecting greater stabilizing selection by ant dispersers.

MATERIALS AND METHODS

Study species and sites

Eastern North America is a biodiversity hotspot for the plants in the genus *Trillium*—a group of myrmecochorous (ant-seed dispersed) perennial understory forest herbs—with at least 29 species occurring in this region (Freeman, 1975; Ohara, 1989). In the southern Appalachian region, many species of *Trillium* are sympatric. Although co-occurring myrmecochore species often temporally stagger fruiting (Warren et al., 2014; Gordon et al., 2019), the sympatric species studied here have overlapping flowering and fruiting phenology (Miller and Kwit, 2018). Due to the spatial proximity of plants at my study sites, foraging ants occasionally come across mature diaspores of co-occurring congeners at the same time (C. N. M., personal observation), resulting in potential interspecific competition for dispersal services. I investigated the importance of morphological and chemical diaspore attributes in determining seed attractiveness to foraging ants using the five species of *Trillium* studied in Miller and Kwit (2018). These included the geographically widespread species *Trillium catesbaei* Elliott and *T.*

cuneatum Raf., as well as the range-restricted, endemic species *T. lancifolium* Raf., *T. discolor* Wray ex Hook, and *T. decumbens* Harbison. Pairs of these species co-occur in multiple locations throughout the southern Appalachians (see Miller and Kwit, 2018).

I located six study sites in spatially distinct forest stands (average size of 130 ha; stands separated from one another by at least 3 km) containing sympatric populations of species pairs, and one additional site containing only *T. lancifolium* to supplement a lack of available mature fruits for this species at the other sites (n = 7 sites; Table 2). Five sites were located in northeast Georgia, and two sites were located near the borders of Georgia, North Carolina, and South Carolina. All sites were low-lying, mesic, deciduous forests with moderate to thick canopy cover. Sites in northeast Georgia were located in the Limestone Valley soil province, and sites at the Georgia-North Carolina-South Carolina borders were located in the Blue Ridge soil province; both provinces are characterized by mildly acidic soils high in phosphate. Sites in northeast Georgia were dominated by tulip poplar (*Liriodendron tulipifera*), American beech (*Fagus grandiflora*), and oak species (*Quercus* spp.). Sites at the Georgia-North Carolina-South Carolina borders were characterized by tulip poplar, American holly (*Ilex opaca*), and eastern hemlock (*Tsuga canadensis*). All sites had understory layers characterized by red buckeye (*Aesculus pavia*), goldenseal (*Hydrastis canadensis*), jack-in-the-pulpit (*Arisaeama triphyllum*), and numerous herbaceous myrmecochorous species including trout lily (*Erythronium americanum*), violets (*Viola* spp.), bloodroot (*Sanguinaria canadensis*), wood anemone (*Anemone quinquefolia*), and sharp-lobed hepatica (*Anemone acutiloba*).

Table 2: Study species collected at field sites in June 2018, and their relative status as attractive or less attractive to ant dispersers.

Species	Status	Site	State
T. cuneatum	Attractive	Cave	Georgia
T. cuneatum	Attractive	Pocket branch	Georgia
T. cuneatum	Attractive	Old mine	Georgia
T. cuneatum	Attractive	Blue hole	Georgia
T. catesbaei	Attractive	Jocassee Gorges	South Carolina
T. catesbaei	Attractive	Whitewater falls	North Carolina
T. decumbens	Less Attractive	Cave	Georgia
T. decumbens	Less Attractive	Pocket branch	Georgia
T. discolor	Less Attractive	Jocassee Gorges	South Carolina
T. discolor	Less Attractive	Whitewater falls	North Carolina
T. lancifolium	Less Attractive	Tilton bridge	Georgia
T. lancifolium	Less Attractive	Blue hole	Georgia

I collected mature fruits just prior to natural dehiscence from the study species at their respective sites during summer 2018 for use in morphological and chemical analyses. Upon return to the laboratory, I stored fruits at -20°C for 1 month prior to diaspore morphological analyses and 4 months prior to chemical analyses. To organize the species by their overall attractiveness to ants, I used seed dispersal probabilities calculated for each species based on *in situ* observations conducted during a previous study in the same study system (Miller and Kwit, 2018). In that study, the proportion of seeds removed by ants (primarily of the genus *Aphaenogaster*) in 1 h from natural seed depots was averaged for 15 individuals of each species at three study sites; the proportion across sites for each species was then averaged (n = 45 hours of observation/species). I interpret these averaged proportions as the probability of seed dispersal, a measure of

attractiveness for each species of *Trillium* and a comparative framework for the present study. *Trillium catesbaei* had the highest average proportion of seeds removed at 16.8% ("most attractive") followed by *T. cuneatum* at 12.5%; *T. lancifolium* at 10.7%; *T. discolor* at 2.5%, and *T. decumbens* at 2.2% ("least attractive"; Fig. 5).

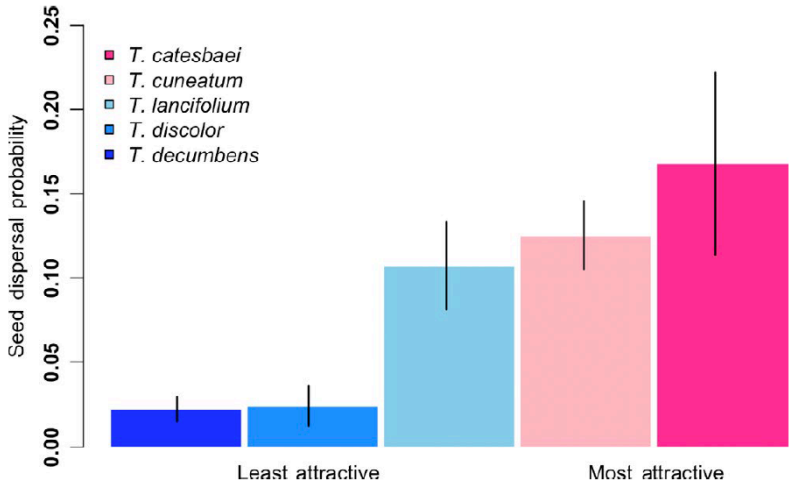


Figure 5: Average probability of seed dispersal by ants for five species of *Trillium* at seven sites in the southeastern U.S. Bars represent standard error. Blue indicates lower seed dispersal probabilities (i.e., "least attractive" species) while pink indicates highest seed dispersal probabilities (i.e., "most attractive" species).

Diaspore morphology

To address my first study objective, quantifying interspecific differences in diaspore morphology, I measured diaspore and elaiosome length, width, and mass for 26 diaspores of each species, representing multiple individuals from each study site (n = 130). I took the fresh mass of entire diaspores (g) and then measured the length and width of diaspores (mm) using digital calipers (Mitutoyo Digimatic Caliper, 0.01 mm

resolution, Sakado, Japan). I then removed the elaiosome from each seed using a straight razor and repeated the above measurements for the elaiosome. Elaiosome-seed mass ratios were calculated by dividing the mass of the elaiosome by the mass of the entire diaspore.

After evaluating each response variable (diaspore length, diaspore width, diaspore mass, elaiosome length, elaiosome width, elaiosome mass, and elaiosome-seed mass ratio) for normality by generating normal probability (Q-Q) plots and histograms to visualize model residuals, I determined that the residuals of all response variables were normally distributed. Each response variable was then compared for the five study species using a linear mixed effects model in the package lmerTest (Kuznetsova et al., 2017) in R (R Core Team, 2017). This test uses the Satterthwaite's method for approximating degrees of freedom for *t* and *F* tests, which is more conservative than residual degrees of freedom approximations; therefore, the assumption of homogeneity of variances across samples can be relaxed (Keselman et al., 1999). Each model included one morphological trait as the response variable, species of *Trillium* as the fixed effect, and study site as a random effect to account for the potential effect of environmental heterogeneity or intraspecific population. Tukey post hoc tests were performed to identify the pairwise direction of effects using the R package multcomp (Hothorn et al., 2008). To visualize trends, which were impacted by the random effect of study site, I generated model residuals from linear models with the effect of site partialled out. Boxplots display these model residuals, which depict the isolated effect of species on diaspore morphology traits.

Elaiosome fatty acid profiles

To address my second study objective, quantifying interspecific differences in concentrations of elaiosome fatty acids among species of *Trillium*, I assessed triglyceride, diglyceride, and free fatty acid forms using gas-chromatography mass spectrometry (GC-MS). Detailed methodology relating to the GC-MS analysis can be found in Appendix I. I produced standard curves by running five concentrations (10, 1, 0.5, 0.1, 0.05, and 0.025 mg/mL) of a fatty acid methyl ester standard mixture of methyl linoleate (20 wt %), methyl linolenate (20 wt %), methyl oleate (20 wt %), methyl palmitate (20 wt %), and methyl stearate (20 wt %) (CRM1891, Millipore Sigma, Merck KGaA, Darmstadt, Germany). Linearity across the relevant range of concentrations was evaluated visually and by calculating the slope of the line and R^2 values for each compound. GC-MS produced TIC chromatograms for each elaiosome fraction (free, di-, and triglycerides). Peaks were identified based on comparisons of retention times and spectra with authentic standards. Compounds for which I did not have standards were tentatively identified using the NIST 14 library (Scientific Instrument Services, 1027 Old York Rd., Ringoes, New Jersey, U.S.A). Peak areas were determined using the Integrator Trapezoid feature in OpenChrom (Wenig and Odermatt, 2010) and concentrations were determined for known compounds based on response factors of standards relative to the internal standard. Concentrations for tentatively identified compounds were estimated as internal standard equivalents. All concentrations were converted to a percentage of elaiosome fresh weight for final reporting and statistical analyses.

Total fatty acid concentrations and concentrations of fatty acids in free, di-, and triglyceride forms were compared for the five study species using linear mixed effects

models in the package lmerTest in R, as above. Concentrations (% fresh weight) of five individual fatty acids and of total fatty acids were evaluated for normality prior to running linear mixed effects models; in each case, the assumption of normality was violated. Therefore, I applied data transformations to raw concentrations to approximate normality (logarithmic or square-root; see Keene, 1995; Osborne, 2002). As in the *Diaspore morphology* methods above, my use of the Satterthwaite's method for approximating degrees of freedom allowed the assumption of homogeneity of variances to be relaxed. Study site was included as a random effect in each model to account for the effects of environmental heterogeneity or interspecific population. Species and fatty acid form (free, di-, or triglyceride) were included as fixed effects. Interactions between species and form were tested for each model; in no case was the interaction term significant, so interactions were dropped from all models. Tukey post hoc tests were performed to identify the pairwise direction of effects between species and between fatty acid forms. To visualize trends, I generated boxplots using model residuals as in the *Diaspore morphology* methods above.

Elaiosome phytochemical profiles

To address my third study objective, assessing interspecific differences in elaiosome phytochemistry, I characterized and compared profiles of both known and unknown elaiosome metabolites using liquid-chromatography mass spectrometry (LC-MS). Each of the five study species was represented by six sample replicates (N = 30). Samples were taken from 2 – 6 individuals from the representative study sites (Table 2). Detailed methodology relating to the LC-MS analysis can be found in Appendix I. LC-

MS produced relative concentrations of phytochemical compounds across samples, which were compared using z-transformed peak areas from extracted ion chromatograms.

To visualize clustering of elaiosome phytochemical profiles across species, I performed a constrained redundancy analysis (RDA) using the vegan package (Oksanen et al., 2017) in R with species of *Trillium* as the explanatory variable and peak areas of the partial phytochemical data (known compounds; $n = 122$) as the response. RDA, a method to summarize the variation in a set of response variables that can be explained by a set of explanatory variables, is a constrained version of principal components analysis wherein canonical axes (i.e. linear combinations of response variables) must also be linear combinations of the explanatory variables (Legendre and Legendre, 1998). I did not explicitly account for variation in the phytochemistry data that may have been explained by study site because RDA partitions the total variance of the data into constrained variances (i.e. variation in the response matrix that is redundant with the variation in the explanatory matrix) and unconstrained variances (i.e. variation in the response matrix that is not redundant with the variation in the explanatory matrix), which can be compared to determine how much of the variation in the response is accounted for by the explanatory variable (Legendre and Legendre, 1998). I used all six sample replicates per species, for a total of 30 samples. The model ran for 1,000 permutations and the significance of the explanatory variables and the first two RDA axes were assessed using ANOVA. I repeated this procedure for the full phytochemical data (known compounds + unknown features; $n = 7,552$) detected by LC-MS.

To better understand the major chemical drivers of the multivariate spread of the partial (i.e. known compounds) and full (i.e. known + unknown compounds)

phytochemical data, I constructed four random forest models using the package randomForest (Liaw and Wiener, 2002) in R. Random forest is a supervised machine learning technique that builds a classification tree by repeatedly splitting the data based on whether or not they fall above or below a threshold value of each explanatory variable in the model (Bielby et al., 2010; Biau, 2012). Random forest ranks the relative importance of different predictors in distinguishing among the levels, and provides a measure of prediction accuracy, cross-validation, for correctly classifying unknown samples into groups according to the predictor variables. By training the model using a subset of samples, random forest is able to classify the remaining samples with higher accuracy than would be achieved by always guessing the most common category.

In my analysis, random forest classification models (RFCMs) were implemented to identify compounds with the highest relative importance in distinguishing among the five species of *Trillium,* a categorical level. Two RFCMs were used evaluate the partial and full phytochemical data sets, respectively. Peak areas of the known compounds were used as the set of possible predictors for the partial RFCM (n = 122), whereas peak areas of all compounds (known + unknown) were used as the set of possible predictors for the full RFCM (n = 7,552). Following implementation of the RFCMs, two random forest regression models (RFRMs) were used to identify compounds with the highest relative importance in distinguishing among the average probability of seed dispersal for each species, a continuous level (16.8% - *T. catesbaei*, 12.5% - *T. cuneatum*, 10.7% - *T. lancifolium*, 2.5% - *T. discolor*, and 2.2% - *T. decumbens*). Peak areas of the known compounds were used as predictors for the partial RFRM, and peak areas of all compounds were used as the set of predictors for the full RFRM, as in the RFCMs above.

Boruta variable selection was used in all four models to select relevant compounds associated with the levels (i.e. *Trillium* species in RFCMs, and seed dispersal probabilities in RFRMs). Unlike regular random forest variable selection procedures, which choose only the smallest number of variables that allow accurate classification of unknown samples to groups, the Boruta algorithm compares the importance of each variable with a set of 'shadow' variables created by randomly shuffling the values of the original variable to list all of the compounds that are likely to differ among groups (Kursa and Rudnicki, 2010). For the selected set of known candidate compounds in models, I assigned molecular formulas and compound classes using the ClassyFire software (Feunang et al., 2016). I followed RFCMs with a MANOVA and then separate ANOVAs, and RFRMs with linear regressions on each set of selected compounds. Predictors were considered significant at $\alpha = 0.05$.

RESULTS

Diaspore morphology

Diaspore length, width, and mass differed significantly among the five *Trillium* species ($F_{4,15} = 4.98$, $P = 0.009$; $F_{4,13} = 7.79$, $P = 0.002$; $F_{4,15} = 9.03$, $P < 0.001$, respectively). Elaiosome length also differed significantly among species ($F_{4,15} = 5.59$, $P = 0.006$). Post hoc comparisons revealed that the second most attractive species, *T. cuneatum,* had significantly greater diaspore length than the least attractive species, *T. decumbens;* that the two most attractive species, *T. catesbaei* and *T. cuneatum,* had significantly greater diaspore width than the least attractive species; and that the two most

attractive species had significantly greater diaspore mass than the least attractive species (Appendix I: Table 9; Fig. 6). Post hoc comparisons revealed that the most attractive species, *T. catesbaei,* had significantly greater elaiosome length than all of the other species except for *T. lancifolium*. Elaiosome width, mass, and elaiosome-seed mass ratio did not differ among species at the $\alpha = 0.05$ threshold ($F_{4,12} = 0.92$, $P = 0.49$; $F_{4,12} = 0.02$, $P = 0.99$; $F_{4,10} = 2.33$, $P = 0.13$; Table 3).

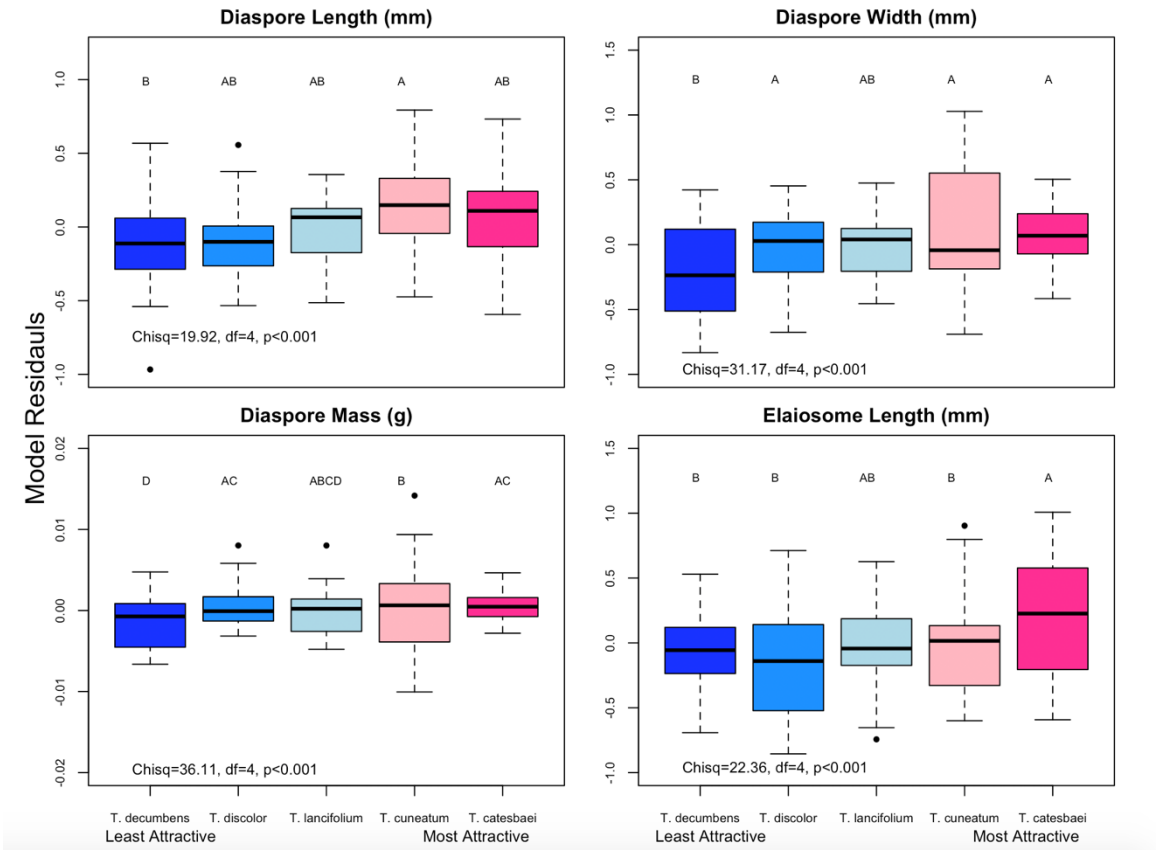


Figure 6: Median, interquartile range, and outliers for model residuals depicting differences in diaspore length, width, mass, and elaiosome length among five species of *Trillium* (blue = least attractive species, pink = most attractive species), with the effects of site partialled out. Statistical results of linear mixed effects models are included at the bottom left of each plot. Letters indicate Tukey post-hoc pairwise differences.

Table 3: Mean ± standard error for diaspore and elaiosome morphological metrics.

Metric (mean +- S.E.)	*T. catesbaei*	*T. cuneatum*	*T. lancifolium*	*T. discolor*	*T. decumbens*
Diaspore mass (g)	0.012 ± 0.001	0.02 ± 0.001	0.02 ± 0.001	0.02 ± 0.002	0.02 ± 0.001
Diaspore width (mm)	2.16 ± 0.09	2.65 ± 0.09	2.30 ± 0.05	2.29 ± 0.11	2.25 ± 0.08
Diaspore length (mm)	3.35 ± 0.12	4.09 ± 0.07	3.52 ± 0.04	3.55 ± 0.15	3.63 ± 0.07
Elaiosome mass (g)	0.001 ± 0.0001	0.004 ± 0.0002	0.003 ± 0.0002	0.003 ± 0.0004	0.004 ± 0.0002
Elaiosome width (mm)	1.03 ± 0.06	1.73 ± 0.06	1.63 ± 0.06	1.39 ± 0.11	1.64 ± 0.08
Elaiosome length (mm)	2.62 ± 0.13	3.18 ± 0.08	2.86 ± 0.07	2.79 ± 0.20	2.92 ± 0.06
Elaiosome-seed mass ratio (%)	0.14 ± 0.01	0.20 ± 0.02	0.19 ± 0.01	0.16 ± 0.02	0.23 ± 0.01

Elaiosome fatty acid profiles

Gas chromatography-mass spectrometry identified a number of short-chain fatty acids (≤ C18). The major compounds included oleic acid (C 18:1), linoleic acid (C 18:2), hexadecenoic acid (C 16:2), stearic acid (C 18:0), and palmitoleic acid (C 16:2); these were detected in all three forms (free, di-, and triglycerides; Table 4). Concentrations of oleic, linoleic, hexadecenoic, stearic, palmitoleic, and total fatty acids differed significantly among the five *Trillium* species ($F_{4,9} = 4.53$, $P = 0.03$; $F_{4,9} = 8.75$, $P = 0.004$; $F_{4,7} = 8.68$, $P = 0.006$; $F_{4,9} = 9.00$, $P = 0.006$; $F_{4,12} = 5.68$, $P = 0.008$; $F_{4,9} = 7.27$, $P = 0.007$, respectively). Post hoc tests revealed that the most preferred species, *T. catesbaei,* had significantly greater concentrations of all compounds and of total fatty acids than the least preferred species, *T. decumbens* (Appendix I: Table 9; Fig. 7). All raw data associated with this study are deposited in Dryad digital repository (d.o.i: https://doi.org/10.5061/dryad.hhmgqnkcz).

Table 4: Mean ± s.e. of fatty acid concentrations across species (avg. across elaiosome forms; top). Mean ± s.e. of fatty acid concentrations within elaiosome forms (bottom).

Trillium	*cuneatum*		*catesbaei*		*lancifolium*		*discolor*		*decumbens*	
Compound	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
Hexadecanoic acid	0.136	0.031	0.365	0.131	0.082	0.032	0.166	0.092	0.048	0.014
Oleic acid	0.624	0.230	1.456	0.687	0.420	0.246	0.721	0.480	0.128	0.047
Stearic acid	0.043	0.018	0.059	0.021	0.040	0.024	0.014	0.006	0.012	0.008
Palmitoleic acid	0.007	0.007	0.082	0.058	0.000	0.000	0.002	0.002	0.000	0.000
Linoleic acid	0.971	0.271	1.345	0.368	0.539	0.300	1.084	0.681	0.271	0.097
Total fatty acids	2.028	0.539	3.371	1.131	1.202	0.649	2.112	1.271	0.656	0.230

Form	Diglyceride		Triglyceride		Free	
Compound	Mean	SE	Mean	SE	Mean	SE
Hexadecanoic acid	0.054	0.016	0.256	0.082	0.157	0.032
Oleic acid	0.281	0.119	1.355	0.426	0.351	0.101
Stearic acid	0.003	0.002	0.024	0.01	0.079	0.017
Palmitoleic acid	0.002	0.002	0.044	0.031	0.004	0.003
Linoleic acid	0.304	0.105	1.503	0.374	0.783	0.223
Total fatty acids	0.78	0.271	3.281	0.869	1.637	0.359

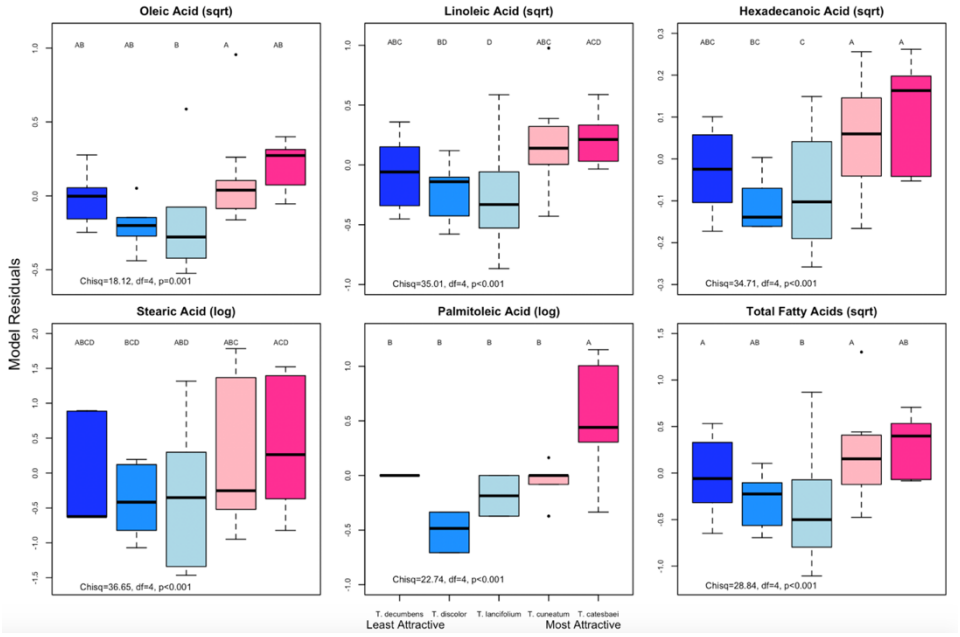


Figure 7: Boxplots illustrating median, interquartile range, and outliers for model residuals depicting differences in concentrations (% fresh weight) of five individual fatty acids and total fatty acids among the species of *Trillium* (blue colors = least attractive species, pink colors = most attractive species), with the effects of site partialled out. Statistical results of linear mixed effects models are included at the bottom left of each plot. Letters indicate Tukey post-hoc pairwise differences.

Elaiosome phytochemical profiles

Across the thirty elaiosome samples, 122 known compounds and 7,431 unknown features (paired *m/z* and retention times, which could correspond to individual compounds, adducts, or other artifacts; Clasquin et al., 2012) were detected by LC-MS. The RDA for known phytochemical compounds was significant ($P < 0.001$; Fig. 8), with 32.03% of the variation in the data explained by species of *Trillium*. The first and second constrained axes explained 14.28% and 8.76% of the variation respectively and were both significant ($P = 0.005$; $P = 0.029$, respectively). The phytochemical profiles of the two most attractive species, *T. catesbaei* and *T. cuneatum,* did not overlap, but the spread of the data for both species was tightly constrained along the second axis. The RDA for the full phytochemical profiles of elaiosomes (including both the 122 known compounds and 7,431 unknown features) was significant ($P = 0.03$), with 15.52% of the variation in the data explained by the constrained RDA. Species of *Trillium* was thus maintained as a significant explanatory variable for elaiosome phytochemical profiles when considering the full set of LC-MS results; this finding was further supported by the clear grouping of species in the full RDA (Appendix I: Fig. 15).

The composition of known phytochemicals in elaiosomes was distinguishable among species of *Trillium* based on the partial RFCM. The out-of-bag error rate was 0.10, meaning the model had 90% sample classification accuracy. Boruta variable selection identified 23 of the possible 122 compounds as important classifiers for species of *Trillium.* A MANOVA showed a significant overall difference among species for these 23 compounds ($F_{20,96} = 6.51$, $P < 0.001$). Follow-up ANOVAs revealed that 20 of the 23 compounds were significantly different among the five species (Appendix I: Table 10).

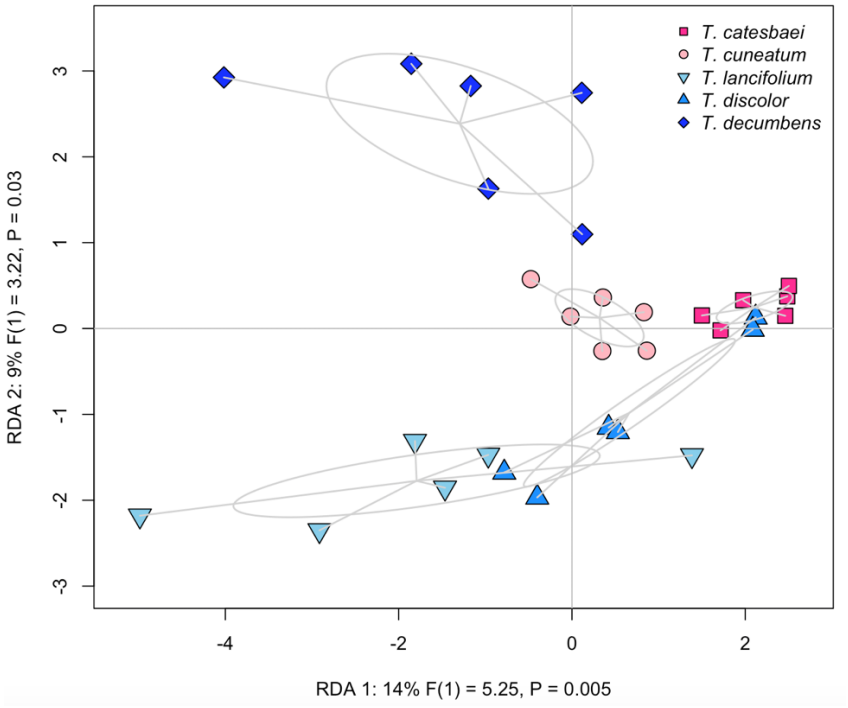


Figure 8: Constrained RDA illustrating the multivariate spread of elaiosome phytochemical profiles (n = 6 samples per species; n = 30). Ellipses represent standard deviation around the centroid for each species. More attractive species are indicated by pink, and less attractive species are indicated by blue.

The composition of known phytochemicals in elaiosomes was also distinguishable among averaged seed dispersal probabilities associated with each of the species of *Trillium* based on the partial RFRM. The model explained 57.27% of the variation in the data, and identified 12 of the possible 122 compounds as important in distinguishing among dispersal probabilities. Follow-up regressions revealed that 8 of the 12 variables were significantly different among the average seed dispersal probabilities.

Of the 12 compounds identified by the partial RFRM, 10 were also identified by the partial RFCM; these include salicylate, xanthine, histidine, 3, 4-

dihydroxyphenylacetate, methylhistidine, citrulline, glucosamine, creatinol o-phosphate, d-glucarate, and pantothenate (Table 5). Salicylate, xanthine, citrulline, glucosamine, and pantothenate had significant negative relationships with seed dispersal probability, while methylhistidine and d-glucarate had significant positive relationships with dispersal probability (Appendix I: Table 10). Results of the full (known + unknown) RFCM and RFRM are included in Appendix I. All four models identified methylhistidine, citrulline, creatinol o-phosphate, and d-glucarate as important (Appendix I: Fig. 16). Follow-up ANOVAs revealed significant differences in relative concentrations of three of the four compounds (methylhistidine, citrulline, and d-glucarate) among the five study species ($F_{4,8.93} = 9.63$, $P < 0.001$; $F_{4,11.26} = 7.39$, $P < 0.001$; $F_{4,12.87} = 18.23$, $P < 0.001$, respectively) and a marginally significant difference in relative concentration of creatinol o-phosphate among the study species ($F_{4,8.98} = 2.46$, $P = 0.07$; Fig. 9).

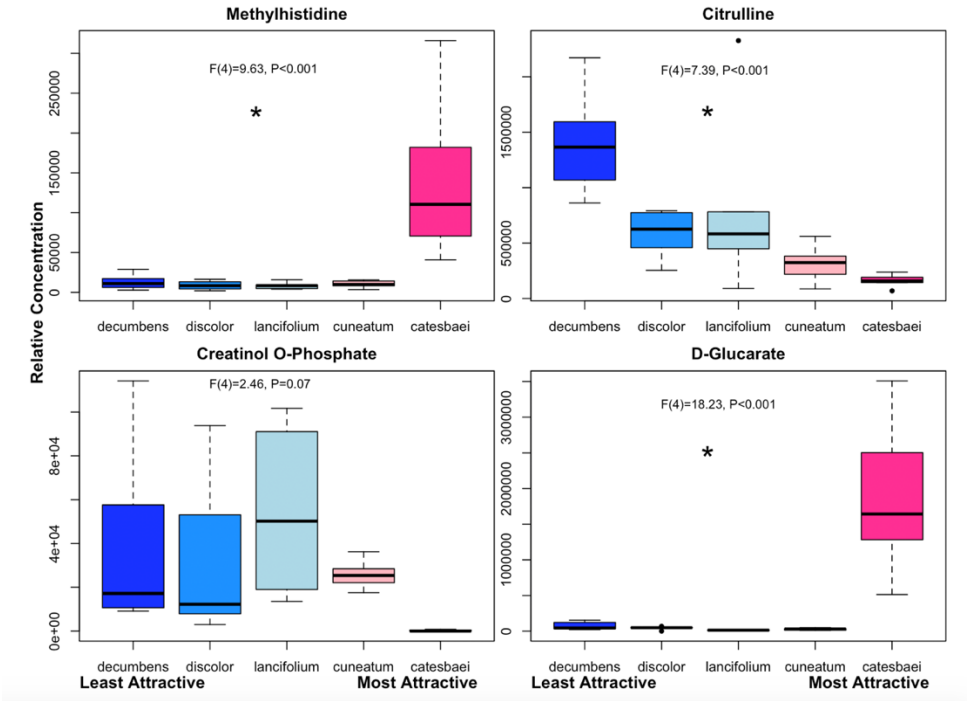


Figure 9: Boxplots illustrating median, interquartile range, and outliers for relative concentrations of elaiosome phytochemicals selected by all random forest models as important (blue = least attractive species, pink = most attractive species). Statistical results of follow-up ANOVA models are included at the top of each plot.

Table 5: Known non-lipid phytochemical compounds selected by both the partial RFCM and the partial RFRM and their chemical formulas and compound classes. Asterisks indicate the compounds selected by all four RF models.

Compound	Formula	Compound class
Salicylate	$C_7H_5O_3$	Salicylic acids
Xanthine	$C_5H_4N_4O_2$	Purines
Histidine	$C_6H_9N_3O_2$	α -amino acids
3,4-Dihydroxyphenylacetic acid	$C_8H_8O_4$	Benzenediols
Methylhistidine[a]	$C_7H_{11}N_3O_2$	α -amino acids
Citrulline[a]	$C_6H_{13}N_3O_3$	*L*- α -amino acids
Glucosamine	$C_6H_{13}NO_5$	Monosaccharides
Creatinol O-Phosphate[a]	$C_4H_{12}N_3O_4P$	Phosphate esters
D-Glucarate[a]	$C_6H_8O_8$	Carbohydrates
Pantothenate	$C_9H_{16}NO_5$	β-amino acids

DISCUSSION

In this study, I investigated the differences in diaspore morphology, elaiosome fatty acids, and elaiosome phytochemisty among five species of *Trillium* with different levels of attractiveness to seed-dispersing ants in the southern Appalachian region of North America. Of the morphology metrics considered, diaspore length, diaspore width, diaspore mass, and elaiosome length are significantly different among species, and post-hoc tests revealed that values of these traits tended to increase with seed attractiveness. These findings provide support for the established hypothesis that seed size is a key trait determining the probability of seed dispersal (Manzaneda et al., 2009), with larger diaspores overall being preferred by seed-dispersing ants (Hughes and Istoby, 1992; Takahashi and Itino, 2015).

Contrary to my prediction, elaiosome mass, elaiosome width, and elaiosome-seed mass ratio were not significantly different among species. These results are not consistent with the results of several studies showing that ants prefer larger elaiosome biomass and/or elaiosome-seed mass ratios (Mark and Olesen, 1995; Leal et al., 2014; Levine et al., 2019), although they do support the finding that ants do not exert significant selection on elaiosome size in *Helleborous foetidus* (Manzaneda et al., 2009). Considering that the most attractive species in my study, *T. catesbaei,* had smaller average elaiosome mass and smaller elaiosome-seed mass ratios than all of its congeners, these aspects of elaiosome morphology do not appear to be the most important factors contributing to the attractiveness of *Trillium* seeds to ants. Larger diaspores likely enhance attractiveness of seeds to ants, but ants may not always prefer seed species with larger elaiosome mass or elaiosome-seed mass ratios.

The results of the GC-MS analysis indicate that elaiosome fatty acid composition is generally homogeneous across these species of *Trillium.* This result is supported by the conclusions of Fischer et al. (2008), who found that elaiosomes of myrmecochores in different families had homogeneous fatty acids dominated by oleic, palmitic, linoleic, and γ-linolenic acids (see also: Soukup and Holman, 1987; Hughes et al., 1994; Fischer et al., 2005; Boulay et al., 2006; Pfeiffer et al., 2010; Boieiro et al., 2012.; Leal et al., 2014). Despite the homogeneity of fatty acid composition, interspecific differences in concentrations of total and individual fatty acids appear to be associated with seed attractiveness to ants. My prediction that the more attractive species of *Trillium* would contain higher total concentrations of fatty acids was supported, corroborating the majority of studies of elaiosome chemistry (Soukup and Holman, 1987; Brew et al., 1989; Kusmenoglu et al., 1989; Lanza et al., 1992; Hughes and Istoby, 1994; Morrone et al., 2000; Boulay et al., 2006; Gammans et al., 2006; Fokuhl et al., 2007; Fischer et al., 2008; Boieiro et al., 2012; Leal et al., 2014; Wu et al., 2019).

The most attractive species also had significantly higher concentrations of oleic acid than the less attractive species, providing support for the hypothesis that this compound acts as a behavior-releasing signal that stimulates ants to pick up and carry items to or from the nest (Marshall et al., 1979; Skidmore and Heithaus, 1988; Brew et al., 1989; Qiu et al., 2015). Oleic acid is the most abundant fatty acid in plant and animal tissue and the biosynthetic precursor of linoleic and linolenic acids (Christie, 2005), essential nutrients that are not synthesized by hymenopterans (Dadd, 1973; Hagen et al., 1984; Barbehenn et al., 1999; Canavoso et al., 2001). As the main constituent in insect hemolymph, oleic acid in the form of diolein is of particular nutritional importance for

ant larvae (Thompson, 1973; Municio et al., 1975; Fischer et al., 2008). However, not all studies are in agreement that oleic acid is responsible for inducing carrying behavior in myrmecochorous systems. For example, O'Dowd and Hay (1980) found that elaiosomes of *Datura discolor* (Solanaceae) were conspicuously absent of oleic acid, whereas palmitic acid, stearic acid, linoleic acid, and linolenic acid were all present. I found that linoleic, hexadecanoic, stearic, and palmitoleic acids were also present in higher concentrations in the more attractive species of *Trillium*, so these fatty acids likely contribute to the overall attractiveness of seeds to ants.

My prediction that elaiosome phytochemical profiles of the more attractive species of *Trillium* would be distinct from their less-attractive congeners was supported. Although the partial RDA shows that all study species were independently clustered in multivariate space, the phytochemical profiles of the more attractive species, *T. catesbaei* and *T. cuneatum,* grouped out closer to one another than any of the less-attractive species, suggesting greater similarity of their chemical profiles. The finding that sympatric endemic and widespread species have distinct elaiosome phytochemical profiles is noteworthy for two reasons. First, these species are closely related, minimizing differences in phytochemical composition that could be attributed to phylogeny. Second, when sympatric, these species may engage in interspecific competition for dispersal services; this suggests that the shared need to attract high-quality ant dispersers may exert different strengths of selection on different species, or that other agents of selection differentially impact elaiosome phytochemistry in sympatric *Trillium* species. Other studies have produced similar findings for fatty acid chemistry in intraspecific populations of myrmecochores. Boulay et al. (2006) found that distant populations of

Helleborus foetidus had significantly different fatty acid composition, which influenced preference in three major ant dispersers, and Boieiro et al. (2012) found that the content of oleic acid differed markedly among populations of *Euphorbia characias* in the Mediterranean. Environmental heterogeneity, lower rates of gene flow, or geographic variations in co-adaptation between plant species and ant dispersers (i.e. less stabilizing selection exerted by ants) may explain the large disparity in the chemical profiles of the less-attractive species. Narrowly endemic populations of myrmecochores may experience infrequent dispersal away from maternal patches, resulting in more heterogeneous chemical composition of their elaiosomes across populations due to low levels of gene flow. Indeed, the lower dispersal probabilities experienced by these endemic species (Miller and Kwit, 2018) supports this interpretation. In contrast, elaiosome phytochemistry of the widespread species might be more homogeneous due to higher levels of gene flow stemming from greater connectivity and higher rates of dispersal, potentially rendering their seeds more recognizable, and thus attractive, to ants.

Salicylate, xanthine, histidine, citrulline, and pantothenate were significantly negatively related to probability of seed dispersal according to follow-up linear models to the partial RFRM. This finding is of interest, because very little research thus far has considered the impacts of repellent chemicals on myrmecochory. These metabolites, belonging to diverse classes including salicylic acids, purines, α-amino acids, L-α-amino acids, and β-amino acids, might function as deterrents to foraging ants making decisions about which items to pick up and carry back to the nest. In their study of European ant-dispersed plants, Fischer, et al. (2008) found citrulline to be present at very low concentrations in elaiosomes of all fifteen species considered. Low concentrations of

citrulline in elaiosomes may therefore be favorable for ant-mediated seed dispersal; alternatively, the presence of citrulline in elaiosomes may simply be a product of the Krebs-Henseleit cycle, wherein arginine (a major constituent of seed protein), citrulline, and ornithine are metabolically interconverted (Hill-Cottingham and Lloyd-Jones, 1973). Davidson, et al. (1990) discovered methyl-6-methyl salicylate, a compound closely related to salicylate, on the seeds of nine taxonomically unrelated ant garden epiphytes in Western Amazonia; this compound was found to be mildly repellent to ants at high concentrations, but to stimulate ant excitement, seed handling, and seed carrying in lower concentrations in *Camponotus femoratus*. My results corroborate these findings in that lower concentrations of salicylate were found in seeds that had a higher probability of being dispersed by ants. The potential roles played by the remaining compounds selected by random forest models is unknown, but could be the focus of further investigations.

The defense trade-off hypothesis posited by Cipollini and Levey (1997), and further explored by Schaefer et al. (2003), may explain the presence of phytochemicals that appear to reduce the attractiveness of trillium seeds. Plants must balance attracting dispersers and repelling the mortality agents that exposed fruits/seeds come into contact with prior to and following dispersal. Whereas some compounds present in elaiosomes might deter ant dispersers, this cost may be off-set by better defenses against granivores or microbial pathogens. In myrmecochores, this could be of particular importance, given the likelihood that rodents will prey on seeds that are not removed by ants within a few hours of dehiscence (Heithaus, 1981). The importance of microbial pathogens has not been investigated thoroughly in myrmecochore systems, so deterrent compounds in elaiosomes might also play a role in defending the seed against fungal infections. The

defense trade-off hypothesis would imply that mortality selection agents (i.e. granivores, microbial pathogens) have had a greater impact on the less-attractive, endemic species of *Trillium* considered here, which tend to contain higher concentrations of these deterrent compounds, whereas the widespread, attractive species might invest in nutrients and attractants resulting in their higher dispersal probabilities.

This study provides evidence that multiple physical and chemical diaspore traits influence how attractive the seeds of eastern North American species of *Trillium* are to ants, primarily of the keystone seed-dispersing genus *Aphaenogaster*. Species within the *fulva-rudis-texana* complex, consisting of *A. fulva, A. rudis, A. texana,* and several other morphologically similar species including *A. picea* (Umphrey, 1996), are difficult to reliably distinguish from one another and may be characterized by polyphyly (DeMarco and Cognato, 2016). My study system encompasses the ecotone between *A. picea* and *A. rudis* (Warren et al., 2016). Although both species are effective seed dispersers, they may be marked with differences in seed dispersal effectiveness that are not accounted for in this paper (see Warren et al., 2011); as such, I acknowledge that some of the variability in seed dispersal probabilities across species of *Trillium* may be due to differences in the seed disperser assemblages at each study site. However, that pairs of the study species co-occur within multiple sites and thus are exposed to identical dispersal assemblages during the *in situ* observations of seed dispersal likely minimizes these differences.

Based on these results, the trillium seeds that are most attractive to ants are characterized by larger diaspores overall (but not necessarily by larger elaiosome biomass or elaiosome-seed mass ratios), and by elaiosomes with high concentrations of the α-amino acid methylhistidine and the carbohydrate d-glucarate, low concentrations of

salicylate, xanthine, histidine, citrulline, and pantothenate, and high concentrations of free oleic, linoleic, hexadecenoic, stearic, palmitoleic, and total fatty acids. Many of these traits are likely correlated, and thus there is some redundancy in this multi-dimensional description of an attractive myrmecochore seed. The results regarding diaspore and elaiosome morphology and elaiosome fatty acids compliment the work of many previous studies, but here I provide novel insights into the potential roles played by previously-unknown components of the broader elaiosome phytochemical profile. Previous work in other *Trillium* species corroborates the conclusion that seed-dispersing ants respond to a complex of characters, including morphological and chemical seed attributes (Gunther and Lanza, 1989; Lanza et al., 1992). Further advances in our understanding of seed dispersal effectiveness in myrmecochorous as well as other animal-mediated seed dispersal systems will require a portrayal of dispersal-unit chemical and physical traits, and their combined responses to selection pressures.

CHAPTER 3: GEOGRAPHIC DISTRIBUTION OF SPECIES IN THE GENUS *TRILLIUM* IN EASTERN NORTH AMERICA

ABSTRACT

Species distribution models (SDMs) statistically combine occurrences with climate variables to predict fundamental niches and produce range maps based on abiotic and environmental conditions. However, SDMs are limited in their ability to capture the effects of species interactions and dispersal limitations on species distributions. One way to incorporate these effects into model projections is to regress SDM fit statistics against a set of species traits related to species interactions or dispersal ability; this can determine whether traits affect the degree to which the realized niche matches the fundamental niche. To determine the overlap between the realized and fundamental niche and to evaluate the performance of SDMs for species with different conservation rankings, range sizes, and reproductive life history traits, I produced SDMs using Maxent for 21 species in the genus *Trillium* in eastern North America. In this system, I investigated whether range size, conservation status, and reproductive life history traits related to biotic requirements and dispersal would predict measures of model performance using generalized linear regression. I found that the average proportional occupancy of the predicted suitable area (PSA) was low (mean = 31%), meaning that for most species of *Trillium* in this study, the realized distribution does not approximate the fundamental niche. Additionally, trends illustrate that rarer species of greater conservation concern tend to have lower proportional occupancy of the PSA than more common species of

lower conservation concern. A significant interaction between conservation status and PSA predicted standardized testing AUC; model performance was positively related to PSA for species of greater conservation concern and negatively related to PSA for species of lower conservation concern. Out of all 21 species, SDMs performed best for those of higher conservation concern with larger predicted suitable areas (PSAs). Flower type and ovule number were the only reproductive traits that were significant predictors of standardized testing AUC. Pedicellate-flowered species had lower model performance than sessile-flowered species, and ovule number was negatively related to model performance. These findings could be explained by phylogeny or differences in reproductive output that may lead to larger rates of dispersal or greater numbers of occurrences. Together, these results suggest that factors not included in SDMs, such as dispersal limitations, reproductive traits, or species interactions, play a larger role in shaping the geographic extent of rare eastern North American species of *Trillium* than their widespread congeners.

INTRODUCTION

The geographic distributions of species have been of interest to ecologists and biogeographers for more than a century (Grinnell, 1917; Dobzhansky, 1950; Hutchinson, 1957; Hutchinson, 1978; MacArthur, 1984; Peterson, 2001; Lomolino et al., 2005). Three broad groups of factors determine species distributions. Abiotic conditions (i.e. environmental variables such as temperature, precipitation, elevation, distance to the nearest body of water, etc.) constitute the first group; these factors define the fundamental

niche. The second group of factors shaping species distributions include biotic conditions, such as vegetative characteristics comprising the landscape, and species interactions, including competitive exclusion, predation, parasitism, or mutualism; these factors shape the realized niche. The third group of factors shaping species distributions consists of dispersal limitations, which govern the ability of species to reach areas that might otherwise be suitable (Soberon & Peterson, 2005). Dispersal limitations include both extrinsic barriers preventing dispersal, as well as intrinsic dispersal limitations, such as low quantitative or poor qualitative dispersal of offspring from the location of the parent. To gain a complete understanding of the constraints on a species' distribution, all three sets of factors must be considered.

Species distribution models (SMDs), also referred to as ecological niche models, have been used extensively over the past two decades to relate species distributions to abiotic factors and characteristics of the environment (Guisan & Zimmermann, 2000; Phillips et al., 2006; Elith et al., 2006; Elith & Leathwick, 2009; Elith et al., 2011; Araújo et al., 2019). By combining known occurrences with a set of climate variables, models can predict areas of suitable habitat, or expected geographic distributions, for the species in question. While these methods have proved useful for predicting the fundamental niche of species and projecting potential range maps, typical SDMs are limited in their ability to quantify the effects of biotic factors and dispersal limitations on species distributions. Remote sensing variables relating to features of the landscape, such as vegetative indices or percent tree cover, can be incorporated into SDMs much like climate variables (Papeş et al., 2012). Studies that have incorporated such remote sensing variables into SDMs have found that they generally improve model fit (Heikkinen et al.,

2007; Meier et al., 2011; Pellissier et al., 2010), significantly alter projections of future distributions (Araújo & Luoto, 2007; Hof et al., 2012), and are among the most supported variables predicting species population density (Lewis et al., 2017). However, temporal mismatch between remote sensing variables, which are mostly restricted to the recent past, and climate variables and occurrence records, which can span hundreds of years, often prevents use of these data in SDM.

One way to incorporate the effects of biotic factors on species distributions into model projections is to regress SDM model fit statistics against a set of species traits. This can determine whether species traits affect the degree to which distributions are determined by climatic conditions; i.e., the degree to which the realized niche matches the fundamental niche. Previous studies in this arena have found that species traits can impact the accuracy, and in some cases, transferability (e.g. Wogan, 2016), of SDMs. Examples include fire disturbance response, life history traits, and rarity in California plants (Syphard & Franklin, 2010); life span and level of human disturbance in central European plants (Hanspach et al., 2010); the relationship between fecundity and body size in bivalves (Montalto et al., 2015); range, migration, habitat, and rarity in birds (Wogan, 2016); and dispersal ability in endemic Mexican mammals (Munguía et al., 2008). These studies can identify, based on a set of traits, species for which traditional SDMs are likely to yield poor results. Comparing model fit statistics to species traits can also inform our understanding of the proportion of the fundamental niche, represented by the potentially suitable area (PSA), that is occupied by the species (Munguía et al., 2008). This provides context for whether abiotic factors or other factors affiliated with species

traits (which may encompass biotic factors or dispersal limitations) exert greater influence over a species' geographic distribution.

In this study, I produced SDMs for 21 species of *Trillium* in eastern North America using a maximum entropy algorithm, Maxent (Phillips et al., 2006). I then regressed two measures of model fit against a set of species traits encompassing known range, conservation status, and reproductive life history, to determine whether traditional SDMs underperform for some of our study species. I had two study objectives. First, I sought to produce predictions of suitable areas based on contemporary climate conditions for 21 of 29 species in the genus *Trillium* in eastern North America (Ohara, 1989), a trillium biodiversity hotspot (Case & Case, 1997). This genus consists of horticulturally important, charismatic spring ephemeral species, which have been the subject of scientific and public interest for 150 years (Meehan, 1870; Gates, 1917; Darlington & La Cour, 1940; Haga & Kurabayashi, 1953; Nohara et al., 1975; Freeman, 1975; Anderson, 1994; Jules & Rathcke, 1999; Tomimatsu & Ohara, 2003; Vellend et al., 2003; Koh et al., 2010; Singh et al., 2017; Miller & Kwit, 2018; Miller et al., 2020). This objective was motivated by my desire to identify factors contributing to the wide variation in the geographic ranges of these species, which have similar dispersal mechanisms and life histories and which are closely related, thus minimizing ecological differences accounted for by phylogeny. My aim was to determine the proportional occupancy of the fundamental niche for each species of *Trillium* by comparing the area of predicted suitability generated by the SDM to the known range. This enabled approximation of the relative importance of climatological factors—as compared to other factors not included in SDMs—in shaping the realized niche for each species.

My second study objective was to integrate aspects of species interactions and intrinsic dispersal limitations into our understanding of *Trillium* species distributions. To address this objective, I regressed a set of species traits (known range, conservation status, and reproductive life history traits) against model fit statistics. Since traditional SDMs incorporate only abiotic variables into predictions of suitable area, they are likely to produce inaccurate predictions—which may be captured by poor model fit—for species with distributions constrained in large part by factors other than abiotic. Indeed, research has shown that the majority of SDM algorithms perform poorly for species with high ecological specialization (i.e. endemic species and/or those with strong biotic interactions) and small sample sizes (i.e. rarer species; Hernandez et al., 2006).

Regarding the first study objective, I hypothesized that the species of *Trillium* that are considered to be endemic or range-restricted (i.e. those with smaller ranges, of higher conservation concern) will have lower proportional occupancy of the fundamental niche than species that are considered to be more widespread (i.e. those with larger ranges, of lower conservation concern). Regarding the second study objective, I hypothesized that species considered to be widespread will have higher model fit than species considered to be endemic or range-restricted. Furthermore, I hypothesized that reproductive life history traits relating to microhabitat requirements and/or the likelihood, frequency, and distance of dispersal will predict model fit, thus providing information about constraints on the distributions of species with different reproductive life histories. The evaluation of these hypotheses will provide information about the constraints on geographic distributions of endemic, as compared to widespread, species in the horticulturally-important genus *Trillium.* Additionally, this study will help distinguish species of *Trillium* that are good

candidates for traditional SDM from those that will require additional modeling efforts to accurately predict the ecological niche.

MATERIALS AND METHODS

Study species

The study system is comprised of plants in the genus *Trillium*. Trilliums are perennial monocot rhizomatous herbs found in temperate deciduous forests of eastern Asia, western North America, and eastern North America (Freeman, 1975). Trilliums flower from March to June, with flowering lasting 2-4 weeks, followed by the production of a single ovary containing a few to dozens of diaspores (i.e. the dispersal unit of the plant; the seed-elaiosome complex). Ripe ovaries often dehisce and drop mature diaspores beneath the maternal plant, or make diaspores available within the fruit on the plant. The seeds of *Trillium*, with their elaiosomes, fit the syndrome of being dispersed by ants (i.e., myrmecochory), though other seed dispersers have been noted for some species (i.e. yellow jackets [*Vespula* spp.]; see Jules, 1996; Zettler et al., 2001). Species of *Trillium* in eastern North America share a mutualism primarily with ants of the *Aphaenogaster fulva-rudis-texana* species group, a complex consisting of *A. fulva, A. rudis, A. texana,* and several other morphologically similar species (Umphrey, 1996; DeMarco & Cognato, 2016). Species in this complex are keystone ant seed dispersers for myrmecochores in eastern North America (Ness et al., 2009).

At least 29 species of *Trillium* are native to eastern North America (Ohara, 1989). Fifteen of these are characterized as rare and are ranked as G3 or higher conservation

priority (NatureServe, 2017). Some rare trilliums are also endemic in the sense that they inhabit restricted geographic regions or specific habitat types; many of these rare, endemic species co-occur with geographically widespread congeners. In this study, I used the extent of occurrence (EOO), number of states/provinces occupied by a species, and representation throughout counties in states/provinces to assign species of *Trillium* a relative commonness status (RCS). EOO was calculated by applying a convex hull technique in the package dismo in R (Hijmans et al., 2012) to all spatially rarified occurrence points for each species. Convex hull draws a polygon (km^2) to delineate the area contained within the shortest boundary encompassing all training occurrences; this method is acknowledged by IUCN as a valid measure of the EOO (IUCN, 2001). Species were categorized with "widespread" RCS if they were found across multiple U.S. states and/or Canadian provinces, represented continuously throughout several counties within each state/province comprising their range, and with a known EOO greater than or equal to 1,300,000 km^2. An RCS of "Medium-range" was assigned to species found in several counties within the state(s)/province(s) comprising their range and with an EOO between 650,000 and 1,300,000 km^2. An RCS of "rare" was assigned to species found only in restricted geographic localities (i.e. found in one or a handful of adjoining counties) or in specific habitats (for example, only known to occur in mesic forests of the Blue Ridge Mountains) and with EOO equal to or less than 650,000 km^2. EOO and RCS of each study species was confirmed by consulting maps of endemism generated by the Biota of North America Program (BONAP; Kartesz, 2015) and IUCN, and through discussions with *Trillium* experts (E. Schilling, University of Tennessee, Knoxville; Aaron Floden,

Missouri Botanical Garden; T. Patrick, Georgia Department of Natural Resources, personal communications).

I included 21 of the 29 native eastern North American species of *Trillium* in this study. Species were selected based on taxonomic certainty and number and type of records. I excluded cryptic species and/or those that are phylogenetically unresolved, such as *Trillium lancifolium* and *T. simile* (E. Schilling, personal communication), as well as *T. pusillum* (Cabe, 1995; Cabe & Werth, 1995; Osaloo et al., 1999; Timmerman-Erskine et al., 2002; Farmer, 2007). Species with protected occurrence records, such as *T. rugelii*, were not included. Finally, study species were selected based on a minimum of 20 occurrence records, as SDMs typically do not perform well with less than 10 training and testing records (Wisz et al., 2008). This eliminated several of the most "rare" species, such as *T. tennesseense, T. oostingii,* and *T. persistens.* I did not include *T. viride, T. texanum,* or *T. gracile,* as these species have ranges that extend west of the Mississippi River, the western-most boundary of which is the extent of the study area.

The 21 species included in the study are listed in Table 6. Nine of the study species are pedicellate, meaning their single flower is attached to the inflorescence by a pedicel. The other 12 species are sessile, meaning their single flower is attached directly to the inflorescence. Pedicellate- and sessile-flowered species of *Trillium* are phylogenetically distinct, with pedicellate-flowered species comprising the paraphyletic subgenus *Trillium* L. (Farmer & Schilling, 2002; Case, 2002a.) and sessile-flowered species comprising the monophyletic subgenus *Sessilium* Raf., formerly *Phyllantherum* Raf. (Farmer & Schilling, 2002; Case, 2002b.). Species in these groups are also characterized by vegetative and reproductive differences, such as mottling of the leaves

Table 6: List of species of *Trillium* included in the study, subgenus, NatureServe conservation status, flower type, relative commonness status (RCS), and reproductive life history characteristics. Life history traits (biomass, ovule number, number of seeds per plant, seed setting rate, and seed mass) obtained from Ohara (1989). *PD = pedicellate-declinate; PE = pedicellate-erect.

Species	Subgenus	Nature Serve Status	Flower Type*	RCS	Biomass	Ovule #	# of seeds/ plant	Seed setting rate	Seed mass (mg)
T. catesbaei	*Trillium*	G4	PD	rare	2.68	53.2	16.1	30.26	3.1
T. cernuum	*Trillium*	G5	PD	widespread	NA	NA	NA	NA	NA
T. cuneatum	*Sessilium*	G4G5	Sessile	rare	6.85	160.3	47.7	29.76	4.02
T. decipiens	*Sessilium*	G4G5	Sessile	rare	3.95	94.6	47.9	50.63	8.86
T. decumbens	*Sessilium*	G4	Sessile	rare	2.56	49	19.5	39.8	6.83
T. discolor	*Sessilium*	G4	Sessile	rare	2.91	44	21	47.73	3.62
T. erectum	*Trillium*	G5	PE	widespread	6.95	105.1	80.3	76.4	5.04
T. flexipes	*Trillium*	G5	PD	medium	6.74	128.7	43.9	34.11	4.2
T. foetidissimum	*Sessilium*	G4	Sessile	rare	1.86	57	21	36.84	16
T. grandiflorum	*Trillium*	G5	PE	widespread	4.47	38.4	26	67.71	6.42
T. luteum	*Sessilium*	G4	Sessile	rare	5.77	47.3	32.2	68.08	3.58
T. ludocivianum	*Sessilium*	G4	Sessile	rare	4.07	57.3	25.3	44.15	8
T. maculatum	*Sessilium*	G4	Sessile	rare	3.15	155	24.2	15.61	8.17
T. nivale	*Trillium*	G4	PE	rare	0.52	27.3	14.3	64.1	2.7
T. recurvatum	*Sessilium*	G5	Sessile	medium	2.67	100.4	7.8	7.77	10.5
T. sessile	*Sessilium*	G4G5	Sessile	medium	2.65	124.9	33.1	26.5	7.8
T. stamineum	*Sessilium*	G3	Sessile	rare	2.31	48.7	14.3	29.36	12.17
T. sulcatum	*Trillium*	G4	PE	rare	NA	NA	71	NA	NA
T. vaseyi	*Trillium*	G4	PD	rare	9.5	41.2	18.5	44.9	3
T. viride	*Sessilium*	G4G5	Sessile	widespread	3.28	67.6	18.5	27.37	5.79
T. underwoodii	*Sessilium*	G4	Sessile	rare	2.47	75.8	20.7	27.31	10.28
T. undulatum	*Trillium*	G5	PE	rare	3.41	34.2	29.2	85.38	4.15

(Case, 2002b.), the number of seeds produced per plant, seed set, and seed weight (Appendix II: Fig. 17).

Occurrence data and study extent

To the best of my knowledge, I obtained every publicly-available presence record for each of the study species. The databases I searched included the Global Biodiversity Information Facility (GBIF; https://www.gbif.org/, accessed July 28 - Aug. 1, 2018), the SouthEast Regional Network of Expertise and Collections (SERNEC; http://sernecportal.org/portal/, accessed Aug 1, 26, 30, and Sept. 1 - 4, 2018; July 12, 2019), Tropicos (https://www.tropicos.org/, March 28, 2019), and online regional herbaria databases, such as the University of Tennessee Herbarium (TENN; https://herbarium.utk.edu/, accessed Aug. 10 – 15, 2018) and the Arnold Arboretum of Harvard University (https://www.arboretum.harvard.edu/, accessed Aug. 17, 2018).

Occurrence data originated from a variety of sources, and the majority of occurrences were opportunistic collections. Approximately 2.5% of the occurrences originated from iNaturalist (https://www.inaturalist.org/; both GBIF and SERNEC provided occurrences derived originally from iNaturalist), representing citizen scientist collections. Approximately 25% of the occurrences originated from herbaria/museum records, and are therefore more likely to be reliable scientific specimens with accurate locality data, which may or may not have originated from standardized surveys. The remainder of the data have unknown sources. The vast majority of occurrences contained no information about sampling method (i.e. standardized scientific survey vs. opportunistic sampling) or sampling date. As such, it was not possible to estimate

sampling bias or the temporal range represented by the data set. However, because the data were derived from a variety of sources, it is expected that the likelihood that they exhibit a consistent form of sampling bias is low, and the risk that sampling bias would drive distribution models is probably negligible. I estimate that the majority of the occurrences originated within the last century.

Just under half of the records were descriptive localities without latitude/longitude coordinates (n ~ 4,000). To assign geographic coordinates to descriptive localities, I used the GEOLocate software (Rios et al., 2010; https://www.geo-locate.org/, accessed from August, 2018 – September, 2019). A centroid of uncertainty with an area of 3 km^2 was automatically assigned to each georeferenced locality by the GEOLocate software. Minimum uncertainty was adjusted manually based on clarity and specificity of record descriptions. Descriptive localities were georeferenced by one of three researchers (myself and two assistants); I checked and confirmed all final latitude/longitude coordinates and associated uncertainties.

The study extent was restricted to eastern North America, including all U.S. states and Canadian provinces east of -92.9 degrees longitude, the western extent of the Mississippi River. I used 1 km^2 spatial resolution (0.0083 decimal degrees; 30 arc-seconds) because this is the original resolution of the climate variable layers (see *Climate variables* below), and the assumption is that this resolution will capture differences in the ranges of endemic and widespread *Trillium* species. All occurrence records within the study extent were included in the models. Occurrence records for each species were spatially rarified to include only one unique record per 1 km^2 using the spatial rarefication tool in SDMtoolbox 2.0 (Brown, 2014; Brown et al., 2017) in ArcMap (Version 10.7,

Redlands, CA). Finally, I plotted all the points for each species on a map of the spatial extent, excluded any points falling outside the verified distributions generated by the Biota of North America Program (BONAP; Kartesz, 2015), and ensured the final set of occurrence records conformed to the known range of each species according to *Trillium* experts Aaron Floden (Missouri Botanical Garden) and Edward Schilling (University of Tennessee, Knoxville). This resulted in a total of 10,116 records across all 21 species of *Trillium*.

Climate variables

Climate variables were generated by Wang et al. (2016) at 1 km^2 resolution (ClimateNA v5.10 software package; available at http://tinyurl.com/ClimateNA), accessed September, 2017. These data include 27 monthly, seasonal, and annual climate variables calibrated at the extent of North America from 1961 – 1990; they include growing and chilling degree days, heating and cooling degree days, mean annual temperature and precipitation, and a climate moisture index (Wang et al., 2016). Environmental variables used in our models were extracted for each of the 1 km^2 cells in the study extent using GIS layers. I did not attempt to minimize correlation among the predictor variables prior to modeling, because removing correlated variables does not significantly alter model fit statistics when implementing machine learning algorithms, as these algorithms have built-in procedures to minimize multidimensionality and collinearity of predictor variables (Tanner et al., 2017; Feng et al., 2019). Instead, I relied on the modeling algorithm to rank the variables that contributed to model accuracy gain (Table 8; see *Species distribution models* below).

Species distribution models

I used Maxent 3.4.1 software (Phillips et al., 2006) for modeling species ecological niches and projecting them as potential geographic distributions. Maxent is a machine learning modeling approach based on maximum entropy that estimates the probability distribution for a species' occurrence based on the given environmental constraints, represented by the climate variables used in the model. Maxent has been shown to perform at a high standard compared to other algorithms for modeling presence-only data (Elith et al., 2006), and has also been proved valid in modeling species distributions with scarce presence-only data (Phillips et al., 2006; Elith et al., 2006; Rebelo & Jones, 2010; Elith et al., 2011). This makes it particularly useful for modeling rare and endemic species with relatively few occurrences, as is the case for several of the study species. Maxent uses presence-only data and a defined number of randomly selected background points (10,000 in our models), or 'pseudo-absences,' across the specified environmental background and combines these with climate variables to construct a continuous index of environmental suitability for each cell. Specific environmental backgrounds based on the EOO were assigned for each species; random points were generated from the EOO, which approximate the environmental conditions and locations that are accessible to each species via dispersal (Soberon & Peterson, 2005; Barve et al., 2011).

To split the occurrence data for each species into model training and testing sets, I used one of two methods. The first method involved selecting a subset of manually georeferenced localities with < 3 km^2 uncertainty as the training data for the model. This method was employed for 9 of the 21 species, and resulted in a training split ranging

from 21 – 70% of the total occurrences (Table 7). This method was initially used for all 21 species, but after checking model outputs and fit, I determined that 12 of the models did not perform well with this method of data splitting (i.e. the AUC scores and omission error were unacceptably high, or the projected geographic range was wildly unrealistic). To rectify this, I allowed the models for each of these 9 species to perform a random 30 – 70 or 50 – 50 testing-training split of all available occurrences, regardless of georeferencing uncertainty and depending on model performance (Table 7); this resulted in improved model fit statistics and predictions. To avoid over-fitting the test data, I set the regularization multiplier value to 1 (Phillips et al., 2004), but lowered it to 0.5 for some models when the AUC scores were unacceptably low. The resulting models obtained did not have high omission error rates, so overfitting was not a concern (Tanner et al., 2017). Linear, quadratic, and hinge features were used to allow variation in the complexity of models (Phillips & Dudik, 2008). Models were initially implemented with all 27 climate variables; after calculating the relative contribution of each variable to the generated Maxent distribution, the top variables contributing to a combined 95% of gain in model fit were retained for use in final models (Table 8).

Model performance was assessed by determining how well each model discriminated between unsuitable and suitable environmental conditions over a range of thresholds (Fielding & Bell, 1997) within the EOO. For any threshold of suitability index, presence locations are either correctly classified as being in suitable conditions (true positives), or incorrectly classified as being in unsuitable conditions (false negatives). False positives (i.e. absences classified as suitable) cannot be estimated by presence-only models, so Maxent estimates instead the fractional predicted area, or the proportion of

cells predicted to have suitable conditions for the species (Phillips et al., 2006). I plotted the receiver operator characteristic curve (ROC) for each model (for both training and testing data) which compares the model sensitivity (true positives) to 1 – specificity (false positives) over the range of thresholds (Fielding & Bell, 1997). The area under this curve (AUC) represents the probability that a randomly chosen presence site will be ranked as more suitable than a randomly chosen pseudo-absence site. A model that performs at random has an AUC of 0.5, whereas a model with perfect discrimination has an AUC of 1. Any model with an AUC ≥ 0.7 is considered to have good discriminatory ability. I used the Maxent logistic value corresponding to 0.1 omission error of training data as a threshold to convert the model suitability values to binary form (suitable above threshold, unsuitable below threshold) and calculate omission error of the testing data points; models with ≤ 0.3 testing omission error were considered reliable. These two model fit statistics (testing AUC and omission error) were the response variables we evaluated for our second study objective, described in *Model fit analyses* below.

AUC is strongly reliant on the extent of the training region (i.e., the EOO). Increasing the extent of the training region typically includes absences that are more distant environmentally from the presences; this artificially increases AUC values, since it is easy to parameterize models with good discriminatory ability, but which are not particularly useful (Barve et al., 2011; VanDerWal et al., 2009). It is therefore not advisable to make direct comparisons of SDMs using AUC scores across species with different training extents (Barve et al., 2011). To compare AUC scores across species in our study, I standardized raw testing AUC scores by the extent of occurrence (AUC/EOO; Table 7). This procedure penalized the overly optimistic AUC scores

obtained for species with larger EOOs. Standardized AUC scores can be interpreted similarly to raw AUC scores in the sense that larger values indicate better model performance.

To test the hypothesis that species of *Trillium* of higher conservation concern with smaller ranges (i.e. endemic or range-restricted species) will have lower proportional occupancy of the fundamental niche than species of lower conservation concern with larger ranges (i.e. widespread species), I performed two generalized linear models (GLMs) with binomial error structures using R (version 3.4.0; R Core Team, 2017). GLMs were used to evaluate the effects of relative commonness status (RCS) and NatureServe conservation status, respectively, on the proportional occupancy of the potentially suitable area (i.e. the proportion of the PSA represented by the overlap between the PSA and the EOO) for each species. RCS is a categorical variable consisting of three levels that contains information about the known range of the set of *Trillium* species, and was derived from binning measures of the EOO. Species with EOO greater than 1,300,000 km^2 were binned as "widespread"; species with EOO between 650,000 - 1,300,000 km^2 were binned as "medium-range"; and species with EOO less than 650,000 km^2 were binned as "rare." NatureServe conservation status, a measurement of a species' conservation risk (NatureServe, 2017), is a categorical variable consisting of two levels. I used the "G" (global) ranking; given that the study extent included several U.S. states and Canadian provinces, the "S" (state-level) ranking was not appropriate. Global status is ranked from G5 (lowest conservation risk) to G1 (highest conservation risk), with intermediate rankings (i.e. G1G2, G4G5) representing species for which a single ranking consensus could not be reached (NatureServe, 2017). The set of study species were

represented by ranks G3, G4, G4G5, and G5 (Table 6). For the purposes of retaining power, I binned species with NatureServe status G3 or G4 into one category (G3/G4) and G4G5 or G5 into a second category (G4G5/G5). Significance of predictors in these, and all subsequent GLMs, was evaluated using likelihood ratio chi-squared tests in the R package car (Fox & Weisberg, 2011) at $a = 0.05$.

Model fit analyses

Range and conservation status

To test the hypothesis that species of lower conservation concern with larger ranges would have higher model performance, I implemented analysis of covariance tests (ANCOVAs) in R. Standardized testing AUC and training presence omission error were included in separate ANCOVA models as the respective response variables. I tested for interactions between PSA (km^2) and NatureServe conservation status, and PSA and RCS, in separate ANCOVAs. Species were again binned into two NatureServe categories (G3/G4 and G4G5/G5) for the purposes of retaining power. I evaluated ANCOVA residuals using frequency histograms and normal quantile-quantile plots, and found that residuals violated the assumption of normality for both response variables. Because both sets of response variables were continuous, positive, and heavily skewed, I performed GLMs with a Gamma error distribution and log link function.

Reproductive life history traits

To test the hypothesis that reproductive life history traits would predict model fit, I performed two GLMs with a Gamma error distribution and log link function, with standardized testing AUC and omission error as the respective response variables (as in

Range and conservation status above). The following five continuous reproductive life history traits were considered as predictors: biomass (g), number of ovules per flower, seed setting rate (%), number of seeds per plant, and seed weight (mg). An additional categorical predictor flower type (which contained three levels: pedicellate-erect, pedicellate-declinate, and sessile), was also considered. Values of all predictor variables were obtained from Ohara (1989). Because the Ohara (1989) data set only contained information for 19 of our 21 study species (excluding *T. cernuum* and *T. sulcatum*), I did not include either of these species in models for this portion of the study.

Prior to implementing GLMs, I evaluated the complete dataset for multicollinearity using the mctest package in R (Imdadullah et al., 2016; Imdad & Aslam, 2018) and the variance inflation factor (VIF) of each predictor using the car package in R (Fox & Weisberg, 2011). Mctest evaluates multicollinearity using six common tests; three of the tests detected collinearity among ovule number, seed setting rate, and number of seeds per plant. Pearson correlation tests indicated ovule number and seed setting rate, ovule number and number of seeds per plant, and seed setting rate and number of seeds per plant were significantly correlated. Additionally, the VIF for ovule number, seed setting rate, and the number of seeds per plant indicated that these predictors were more closely related to one another than to the response variables. I therefore eliminated seed setting rate and number of seeds per plant from the data set and tested for multicollinearity and VIF again; none of the six tests of collinearity detected issues with collinearity among biomass, ovule number, seed weight, or flower type (Imdadullah et al., 2016; Imdad & Aslam, 2018; Imdad et al., 2019; Appendix II: Table 11), nor did VIF scores indicate problems with collinearity among these four predictors. Following

implementation of GLMs, I evaluated effect size and power using the pwr package in R (Champely, 2017). I performed an additional GLM with a binomial error structure to evaluate the effects of these four reproductive life history traits on the proportional occupancy of the potentially suitable area.

RESULTS

Species distribution models

Overall, SDMs had high performance metrics. Success rates for classification of training points (testing AUC) were high for most SDMs, ranging from 0.558 (*Trillium discolor*) to 0.845 (*T. vaseyi*), with an average of 0.733 (Table 7). Training omission error was low for most SDMs, ranging from 0.103 (*T. undulatum*) to 0.315 (*T. foetidissimum*), with an average of 0.199. The climate variables that averaged ≥ 5% contribution across all 21 models included CMD (Hargreave's climatic moisture index), TD (difference between mean temperature of the coldest month and mean temperature of the warmest month, °C), DD18 (degree-days above 18°C), EXT (extreme maximum temperature over 30 years), NFFD (the number of frost-free days), MAR (mean annual solar radiation [MJ m-2 d-1]), and MWMT (mean temperature of the warmest month, °C; Table 8).

Resulting distribution models and range maps for 21 species in the genus *Trillium* native to eastern North America have been provided (see Fig. 10 and Appendix II: Fig. 21). Table 7 shows the best model for each species. Figures of models depict the PSA (km^2), a binary map of potentially suitable environment, with the EOO (km^2), an estimate of the known range, depicted by a polygon. The area of overlap between the PSA and EOO is illustrated in green; this represents the proportional occupancy of the PSA. PSA

ranged widely, from 59,307 km^2 (*T. ludovicianum*) to 1,657,908 km^2 (*T. undulatum*), with an average of 702,834.7 km^2 (Table 7). EOO ranged from 10,831 km^2 (*T. discolor*) to 1,981,141 km^2 (*T. cernuum*), with an average of 569,180.8 km^2 (Table 7). Simple linear models, evaluated using ANOVA, yielded significant, positive relationships between PSA and EOO ($F_{1,19}$ = 16.58, P < 0.001, R^2 = 0.47; Appendix II: Fig. 19a.), between PSA and the number of presence-only occurrences used in each SDM ($F_{1,19}$ = 7.97, P = 0.01, R^2 = 0.30; Appendix II: Fig. 19b.), and between EOO and the number of presence-only occurrences used in each SDM ($F_{1,19}$ = 33.53, P < 0.001, R^2 = 0.64; Appendix II: Fig. 19c.).

The proportional occupancy of the PSA was low for most species. Proportions of the PSA represented by the intersection between the EOO ranged from only 0.3% (*T. discolor*) to 89% (*T. erectum*), with an average of 31% (Appendix II: Fig. 18). Although binomially-distributed GLMs revealed no significant effects of either RCS or NatureServe conservation status on the proportional occupancy of the PSA at the a = 0.05 level (X^2_2 = 3.14, P < 0.20; X^2_2 = 1.95, P < 0.16, respectively), trends indicated that rare species and those of higher conservation concern tended to have lower proportional occupancy of the PSA than medium or widespread species of lower conservation concern (Fig. 11a.; Fig. 11b., respectively).

Table 7: A list of SDM characteristics for each species. Total occ. = total number of occurrences/species (total = 10,116). Train occ., Test occ. = number of occurrences used for training and testing, respectively. % Train, % Test = percent of the total number of occurrences used for training and testing, respectively. Asterisks indicate species for which the first method of data splitting was used. Sensitivity/specificity for testing data (AUC), Sensitivity/specificity for training data (AUC), and 10 % training presence test omission rate = model fit statistics. Standardized testing AUC = sensitivity/specificity for testing data (AUC) / EOO. PSA, EOO = areas (km^2) of predicted suitable area and known range, respectively. Intersect = area of overlap between the PSA & EOO. Intersect/PSA = percent of area of overlap divided by PSA.

Species	Total occ.	Train occ.	Test occ.	% Train	% Test	Sensitivity/ specificity for testing data (AUC)	Standardized testing AUC	Sensitivity/ specificity for training data (AUC)	10 % training presence test omission rate	PSA (km^2)	EOO (km^2)	Intersect (km^2)	Intersect /PSA (%)
T. catesbaei	274	57	217	0.21*	0.79*	0.72	2.69E-06	0.871	0.3	440,098	267,241	97,504	22
T. cernuum	612	306	306	0.5	0.5	0.844	4.26E-07	0.871	0.153	1,367,839	1981,141	596,166	44
T. cuneatum	402	201	201	0.5	0.5	0.697	1.73E-06	0.738	0.132	1,206,993	403,195	270,078	22
T. decipiens	29	9	20	0.3	0.7	0.680	1.14E-05	0.900	0.250	311,082	59,835	20,696	7
T. decumbens	50	16	34	0.32*	0.68*	0.689	2.30E-05	0.799	0.147	713,702	29,990	17,392	2
T. discolor	20	14	6	0.7	0.3	0.558	5.15E-05	0.675	0.167	896,703	10,831	2,499	.03
T. erectum	1,991	996	995	0.5	0.5	0.776	4.69E-07	0.784	0.117	801,654	1,653,056	714,869	89
T. flexipes	438	219	219	0.5	0.5	0.752	6.12E-07	0.806	0.169	834,348	1,228,625	586,008	70
T. foetidissimum	112	39	73	0.35*	0.65*	0.763	2.07E-05	0.9	0.315	85,788	36,927	9,595	11
T. grandiflorum	2,399	1,200	1,199	0.5	0.5	0.752	4.37E-07	0.774	0.127	1,178,332	1,722,397	819,539	70
T. ludovicianum	47	31	16	0.66*	0.34*	0.787	1.39E-05	0.878	0.312	59,307	56,617	12,403	21
T. luteum	147	67	80	0.45*	0.55*	0.719	8.60E-06	0.713	0.125	1,033,378	83,599	27,690	3
T. maculatum	75	38	37	0.5	0.5	0.619	3.08E-06	0.719	0.135	653,683	200,894	137,883	21
T. nivale	195	59	136	0.3	0.7	0.733	1.20E-06	0.863	0.235	506,524	611,950	252,255	50
T. recurvatum	1,118	378	740	0.34*	0.66*	0.713	1.00E-06	0.804	0.247	951,733	709,791	308,960	33
T. sessile	597	299	298	0.5	0.5	0.765	7.80E-07	0.79	0.117	796,876	980,410	462,198	58
T. stamineum	118	36	82	0.3	0.7	0.675	6.01E-06	0.768	0.268	369,648	112,373	73,918	20
T. sulcatum	97	68	29	0.7*	0.3*	0.697	4.55E-06	0.805	0.3	120,728	153,345	64,358	53
T. underwoodii	85	31	54	0.36*	0.64*	0.766	8.07E-06	0.849	0.296	416,141	94,925	63,529	15
T. undulatum	1,187	594	593	0.5	0.5	0.84	5.64E-07	0.851	0.103	1,657,908	1,488,725	497,193	30
T. vaseyi	123	48	75	0.39*	0.61*	0.845	1.26E-05	0.903	0.173	357,063	66,929	17,804	5

Table 8: Percent contribution of environmental variables to SDMs. Tmin (min. temp., ºC), Tmax (max. temp., ºC), Tave (mean temp., ºC), Ppt (total precip., mm), MAT (mean annual temp., °C), MWMT (mean temp. of warmest month, °C), MCMT (mean temp. of coldest month, °C), TD (difference between MCMT & MWMT, °C), MAP (mean annual precip., mm), MSP (May - Sep precip., mm), AHM (annual heat moisture index), SHM (summer heat moisture index), DD_0 (degree-days below 0°C), DD5 (degree-days above 5°C), DD_18 (degree-days below 18°C), DD18 (degree-days above 18°C), NFFD (number of frost-free days), bFFP (julian date on which FFP begins), eFFP (julian date on which FFP ends), FFP (frost-free period), PAS (precipitation as snow [mm]), EMT (extreme min. temp. over 30 years), EXT (extreme max. temp. over 30 years), Eref (Hargreave's reference evaporation), CMD (Hargreave's climatic moisture index), MAR (mean annual solar radiation), RH (mean annual relative humidity, %), Tave_wt (Dec - Feb mean temp., °C), Tave_sm (Jun to Aug mean temp., °C), PPT_wt (Dec - Feb precip., mm), and PPT_sm (Jun - Aug precip., mm) (Hamann et al., 2013).

Species	ahm	bffp	cmd	dd_0	dd_18	dd5	dd18	effp	emt	eref	ext	ffp	map
T. catesbaei	0.044	0	0.318	0.102	0	0	0.045	0	0	0	0.075	0	0
T. cernuum	0	0.024	0.078	0	0	0	0.026	0	0.029	0.026	0.017	0	0.013
T. cuneatum	0.014	0.007	0.327	0	0	0	0	0.031	0.058	0	0.024	0.021	0.116
T. decipiens	0	0.128	0.026	0	0	0	0	0.526	0	0.063	0	0.375	0
T. decumbens	0	0	0	0	0	0	0	0	0	0	0	0	0.266
T. discolor	0	0	0	0	0	0	0	0	0	0	0	0	0
T. erectum	0	0	0.049	0	0	0	0	0.023	0	0.047	0.257	0	0.015
T. flexipes	0.102	0	0	0	0	0	0.384	0.043	0	0.035	0.065	0	0
T. foetidissimum	0	0.305	0.07	0	0	0	0	0.048	0	0.037	0	0	0
T. grandiflorum	0.025	0	0.026	0	0	0	0.454	0	0	0	0.137	0	0.041
T. ludocivianum	0	0	0	0	0	0	0	0	0	0	0.17	0	0
T. luteum	0.035	0	0	0.192	0	0	0	0	0	0	0	0	0
T. maculatum	0.008	0	0.011	0	0.008	0.074	0.021	0.03	0	0	0.039	0	0.094
T. nivale	0	0	0.029	0.048	0	0	0	0	0.007	0.022	0.07	0	0.01
T. recurvatum	0	0	0.004	0.019	0	0	0.061	0.006	0	0.03	0.044	0	0
T. sessile	0.002	0	0.003	0.175	0	0	0.181	0.003	0	0.117	0.071	0	0.012
T. stamineum	0	0	0.049	0	0	0	0.209	0	0.038	0.018	0.217	0	0.09
T. sulcatum	0	0.02	0.18	0	0	0	0	0.102	0.03	0	0.071	0.206	0
T. underwoodii	0.042	0	0.226	0	0	0	0	0	0	0	0	0	0.189
T. undulatum	0.163	0	0.101	0	0	0	0	0	0.268	0.034	0.072	0	0.109
T. vaseyi	0	0	0.635	0	0	0	0	0	0	0	0	0	0

(Table 8 continued.)

Species	mar	mat	mcmt	msp	mwmt	nffd	pas	ppt_sm	ppt_wt	rh	shm	tave_sm	tave_wt	td
T. catesbaei	0	0	0	0	0	0	0.125	0.064	0.229	0	0	0	0	0
T. cernuum	0.013	0.361	0	0	0.02	0	0	0.058	0.209	0	0.056	0	0	0.047
T. cuneatum	0.018	0	0.071	0.012	0.016	0	0.045	0	0.04	0	0	0	0	0.2
T. decipiens	0	0	0	0	0	0.182	0	0.026	0	0.199	0	0	0	0
T. decumbens	0	0	0	0	0.437	0	0.116	0	0	0	0	0	0	0.274
T. discolor	0.307	0	0	0.019	0	0	0	0.111	0	0.163	0.04	0	0	0
T. erectum	0	0	0.291	0	0	0	0	0.147	0.146	0	0	0	0	0.025
T. flexipes	0.02	0	0	0	0	0.087	0	0.023	0.038	0.077	0	0	0	0.125
T. foetidissimum	0	0.025	0	0	0.048	0.327	0.03	0.036	0.012	0	0.026	0.035	0	0.001
T. grandiflorum	0.018	0	0	0.019	0.103	0	0.085	0	0.03	0.022	0.017	0	0	0.024
T. ludovicianum	0.027	0	0	0	0.484	0	0	0.21	0.046	0	0	0	0	0.063
T. luteum	0	0	0	0	0	0.312	0.072	0.095	0	0.101	0.108	0	0	0.084
T. maculatum	0.323	0	0	0.203	0.061	0	0	0	0	0	0	0.02	0	0.108
T. nivale	0.148	0	0	0	0.02	0.224	0.145	0	0.03	0.177	0.068	0	0	0.001
T. recurvatum	0.022	0	0.056	0	0.019	0	0.362	0.053	0.081	0.223	0.021	0	0	0
T. sessile	0	0	0	0.003	0	0	0.058	0.001	0	0.044	0.053	0.014	0	0.162
T. stamineum	0	0	0	0	0.009	0	0	0	0.011	0	0	0.102	0	0.257
T. sulcatum	0	0	0.035	0.298	0	0	0.033	0	0	0.026	0	0	0	0
T. underwoodii	0.265	0	0	0	0	0.046	0	0	0	0	0.232	0	0	0
T. undulatum	0	0	0.101	0	0.033	0	0	0.113	0	0.006	0	0	0	0
T. vaseyi	0	0	0.08	0	0	0	0	0.049	0	0	0.184	0	0	0.051

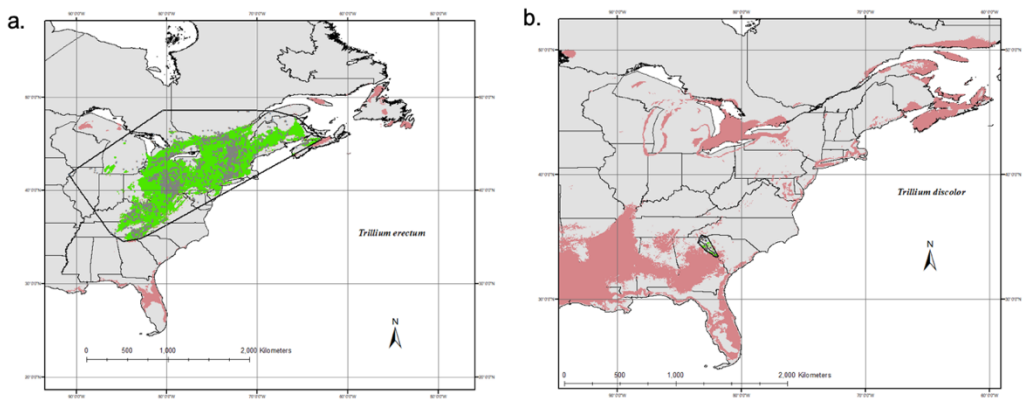


Figure 10: Maps illustrating the predicted suitable area (pink) for a. *Trillium erectum* and b. *T. discolor.* Convex hulls depict the EOO. Areas in green represent the overlap between the PSA and the EOO. For *T. erectum,* proportional occupancy of the PSA is 89%; for *T. discolor,* 0.3%.

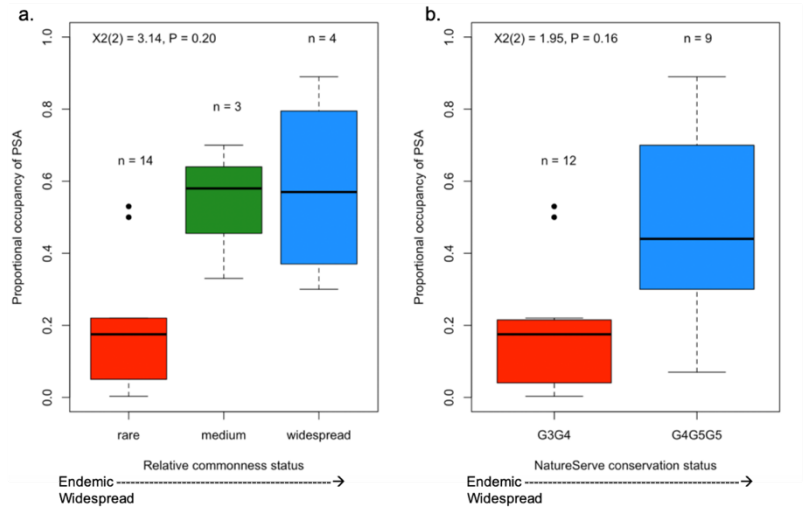


Figure 11: Boxplots illustrating median, interquartile range, and outliers for the proportional occupancy of the PSA (i.e. area of PSA represented by the overlap with the EOO / total area of PSA) of species according to a. relative commonness status and b. NatureServe conservation status. Test statistics are at the top left of each plot. Species lower on the endemism gradient (red) have lower proportional occupancy of the PSA than those higher on the endemism gradient (green/blue). The number of species represented by each category is indicated above each boxplot.

Model fit analyses

Range and conservation status

A Gamma-distributed GLM yielded a significant interaction between PSA and NatureServe conservation status on standardized testing AUC ($X^2_1 = 19.80$, $P < 0.001$). *Trillium* species of greater conservation concern (G3/G4) with smaller PSA had lower standardized testing AUC scores (and therefore lower model performance) than species of greater conservation concern with larger PSA. The opposite trend was observed for species of lower conservation concern (G4G5/G5); those with smaller PSA had slightly higher standardized testing AUC scores, meaning slightly higher model fit, than species of lower conservation concern with larger PSA (Fig. 12a.). The results of a second GLM yielded a significant effect of PSA on omission error ($X^2_1 = 16.09$, $P < 0.001$), while neither NatureServe status nor the interaction between PSA and NatureServe status were significant ($X^2_1 = 1.29$, $P = 0.25$; $X^2_1 = 2.05$, $P = 0.15$, respectively). As PSA increased, omission error decreased (and thus model performance increased) for all species, regardless of their conservation status (Fig. 12b.). Simple linear models, evaluated using ANOVA, yielded significant relationships between EOO and NatureServe conservation status ($F_{3,17} = 22.97$, $P < 0.001$; Appendix II: Fig. 19d.) and between PSA and NatureServe conservation status ($F_{3,17} = 4.99$, $P = 0.01$; Appendix II: Fig. 19e.).

The results of a GLM testing the effects of PSA and RCS revealed that neither PSA, RCS, nor their interaction were significant predictors of standardized testing AUC ($X^2_1 = 0.08$, $P = 0.78$; $X^2_2 = 2.75$, $P = 0.25$; $X^2_2 = 0.07$ $P = 0.96$, respectively). However, I found a significant interaction between PSA and RCS on omission error ($X^2_2 = 15.27$, $P <$

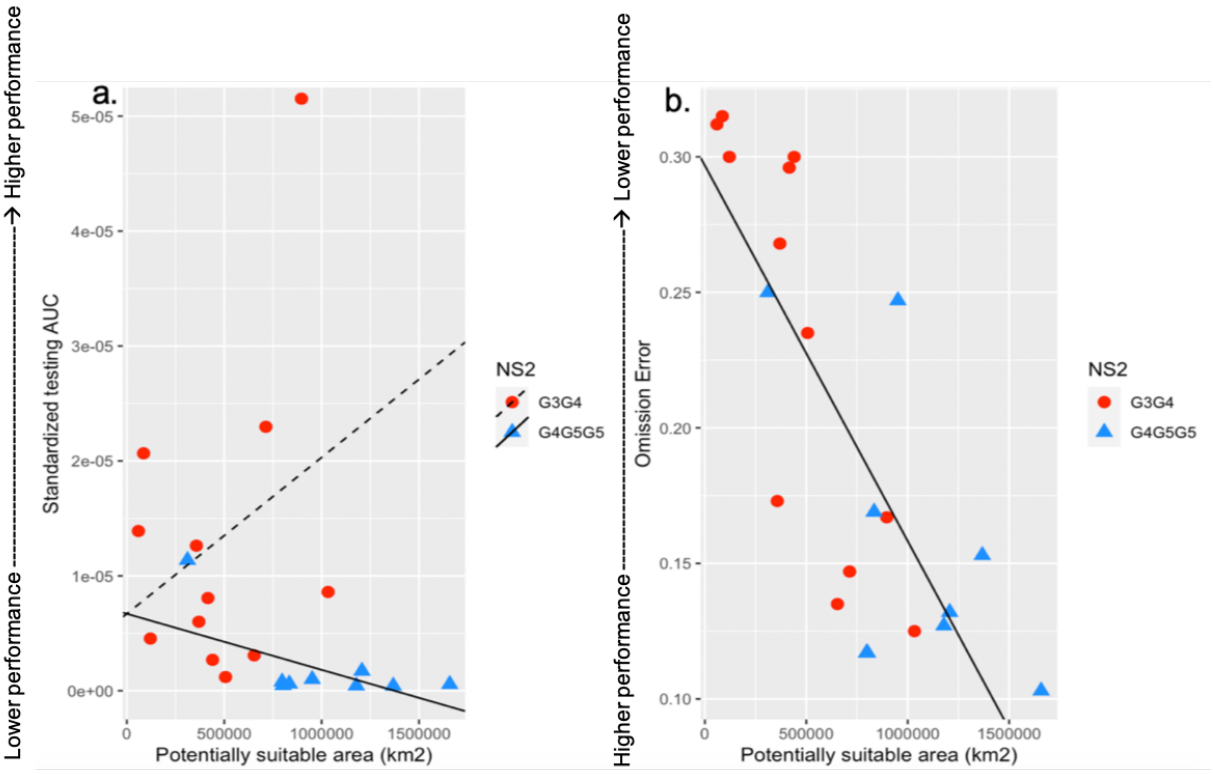


Figure 12: Scatterplots illustrating the effects of potentially suitable area (PSA) and NatureServe conservation status on two model fit statistics, a. standardized testing AUC, and b. omission error, for 21 species of *Trillium.* Red dots indicate species at higher conservation risk (G3/G4 NatureServe status) and blue triangles indicate species at lower conservation risk (G4G5/G5 NatureServe status). The plot on the left depicts the interacting effects of PSA and NatureServe status on standardized testing AUC. With the same intercept, standardized testing AUC increases for G3/G4 species as PSA increases (Y = 1.360e-11X + 6.706e-06, $P = 0.16$), while standardized testing AUC decreases for G4G5/G5 species as PSA increases (Y = -4.880e-12X + 6.705991e-06, $P = 0.38$). The inverse relationship between standardized testing AUC and raw testing AUC means model performance is higher for G3/G4 species with lower PSA and for G4G5/G5 species with higher PSA. The plot on the right illustrates the significant effect of PSA on omission error; as PSA increases, omission error decreases significantly (Y = -1.385e-07X + 2.967e-01, $P < 0.001$, $R^2 = 0.6$). Legend: "NS2" refers to grouped NatureServe conservation status.

0.001). Rare *Trillium* species with larger PSA had lower omission error (better model performance) than rare species with smaller PSA. Medium-range species varied only slightly in their PSA (796,876 – 951,733 km^2), but omission error increased considerably with increasing PSA, meaning models performed worse for medium-range species with larger PSA. Widespread species did not exhibit change in omission error with increasing PSA (Fig. 13).

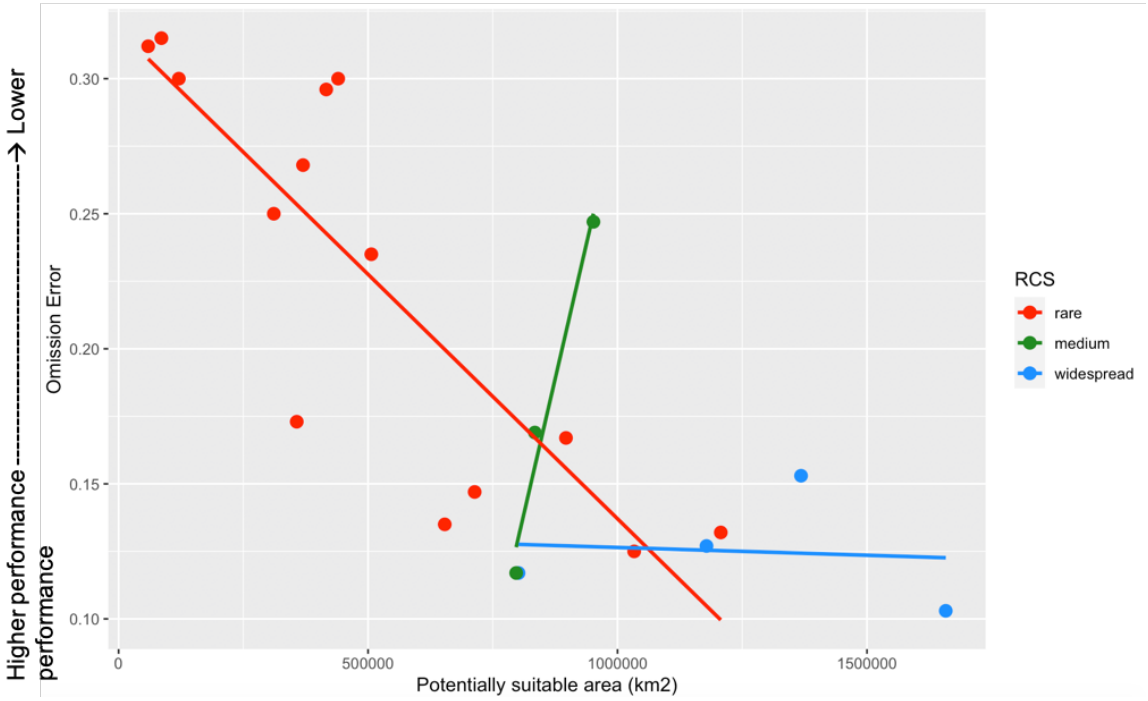


Figure 13: Scatterplot illustrating the interacting effects of potentially suitable area (PSA) and relative commonness status (RCS) on omission error for 21 species of *Trillium.* Red dots indicate rare species; green dots indicate medium-range species; blue dots indicate widespread species. Rare species of *Trillium* with larger PSA had significantly lower omission error (Y = -1.810e-07X + 3.181e-01, R^2 = 0.69, P < 0.001). Medium-range species with larger PSA had higher omission error (Y = 7.976e-07X - 5.090e-01, R^2 = 0.94, P = 0.11). Omission error did not vary significantly with PSA for widespread species (Y = -5.774e-09X + 1.322e-01, R^2 = 0.01, P = 0.90).

Reproductive life history traits

Flower type and ovule number were significant predictors of standardized testing AUC according to the results of a Gamma-distributed GLM ($X^2_1 = 38.464$, $P < 0.001$; $X^2_1 = 23.132$, $P < 0.001$, respectively; Appendix II: Table 12). Standardized testing AUC (and therefore model performance) was lowest for pedicellate-erect species and highest for sessile species, with pedicellate-declinate species having intermediate performance (Fig. 14a.). Standardized testing AUC decreased with increasing ovule number (Fig. 14b.), meaning that species with larger numbers of ovules per flower tended to have lower model performance. The Gamma-distributed GLM for omission error yielded no significant predictors ($P > 0.05$ in all cases; Supplemental Materials Table 2). Power analyses showed that the GLM for standardized testing AUC had a good effect size and high power on 4 and 14 degrees of freedom at significance level 0.05 (effect size (f^2) = 0.92, power = 0.82), while the effect size and power of the GLM for omission error were lower (effect size (f^2) = 0.45, power = 0.47). The low effect size and power for the GLM testing omission error might have impacted our ability to reject the null hypothesis.

According to the results of a binomially-distributed GLM, none of the reproductive life history traits were significant predictors of the proportional occupancy of the PSA ($P > 0.05$ in all cases). Flower type was, however, marginally significant ($X^2_1 = 3.70$, $P = 0.054$), with sessile-flowered species tending to have lower proportional occupancy of their fundamental niche than pedicellate-declinate or pedicellate-erect species (Appendix II: Fig. 20).

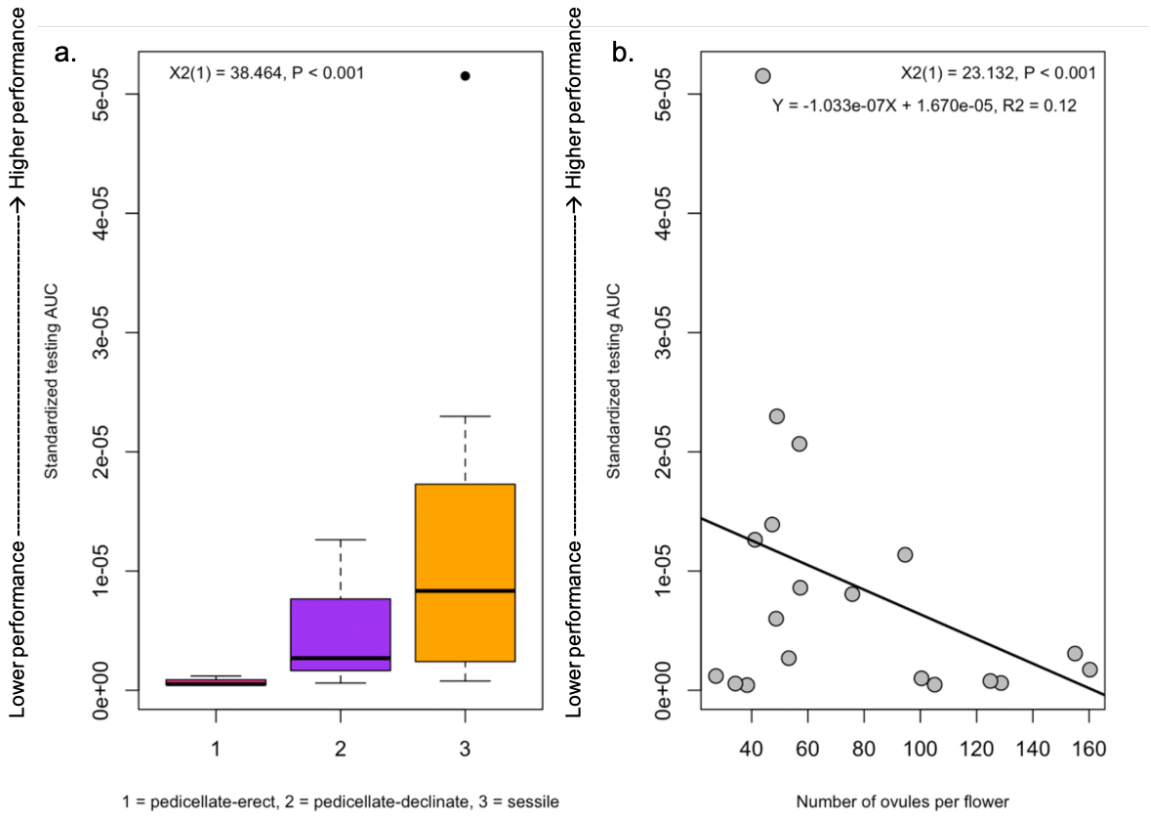


Figure 14: a. Boxplot illustrating median and interquartile range for standardized testing AUC scores of species of *Trillium* with 1. pedicellate-erect flowers, 2. pedicellate-declinate flowers, and 3. sessile flowers. Sessile-flowered species have significantly larger standardized testing AUC scores than pedicellate-erect and pedicellate-declinate species, meaning that model performance in terms of classification of training points is higher for sessile species. b. Scatterplot illustrating the significant effect of ovule number (per flower) on standardized testing AUC scores. As ovule number increases, standardized testing AUC decreases, meaning species with larger numbers of ovules have lower model performance than species with lower ovule numbers. Test statistics (and the equation of the line for b.) are included at the top of each plot.

DISCUSSION

When a species' actual distribution only partially occupies its predicted fundamental niche, dispersal limitations, biotic limitations, or a combination of these factors are expected to limit the geographic extent of the species (Svenning & Skov, 2004; Soberon & Peterson, 2005; Peterson, 2006; Munguía et al., 2008). In this study, I calculated the proportional occupancy of the potentially suitable area (PSA) for each of 21 species of *Trillium* in eastern North America as the overlap between the known range (EOO) and the PSA. I found that the proportional occupancy of the PSA was not significantly related to relative commonness status (RCS) or NatureServe conservation status of *Trillium* species. However, trends indicate that rare species of greater conservation concern tend to have lower proportional occupancy of the fundamental niche than medium-range or widespread species of lower conservation concern. Because more than half of the species in this study were considered rare (n = 14) and/or of greater conservation concern (n = 12), the realized niche for the majority of the study species does not closely approximate the fundamental niche (mean = 31% proportional occupancy). Although EOO is not a concrete measurement of the known range of a species, nor the only means of estimating the known range, IUCN recognizes it as a valid measure of range size (IUCN, 2001; Marcer et al., 2013). The finding that the proportional occupancy of the PSA is low for the majority of species considered in this study (and that the majority of those species are range-restricted or of greater conservation concern) suggests that factors not captured by traditional SDMs play a substantial role in determining the actual geographic distributions of these species. Whether those factors are categorized as species interactions, biotic characteristics of the

environment, edaphic factors, or dispersal limitations, I cannot say. It is possible that the EOO for some species is missing existing but unsampled populations, in which case it is underestimating the realized niche. The PSA might capture some of these unknown occupied areas, and therefore the mismatch between EOO and PSA could also be representative of sampling insufficiencies.

Reproductive life history traits were also not significant predictors of the proportional occupancy of the PSA for trilliums in this study, although there was a trend for sessile-flowered species to have lower proportional occupancy than pedicellate-flowered species. This may be explained by the fact that many of the rare species of greater conservation concern were also sessile, including *T. decumbens, T. discolor, T. foetidissimum, T. ludovicianum, T. luteum, T. maculatum, T. stamineum,* and *T. underwoodii.* Munguía et al. (2008) similarly found that life history traits were poor predictors of the proportional occupancy of species potential distributions for 89 species of endemic Mexican mammals. However, their study found dispersal mechanism to be a significant predictor of the proportional occupancy of the PSA (Munguía et al., 2008). In this study, all of the species assumedly share a dispersal mechanism: myrmecochory (ant-mediated seed dispersal). Previous work in my system has revealed lower quantitative and qualitative seed dispersal effectiveness for endemic species of *Trillium* compared to geographically widespread congeners, likely due to interspecific differences in diaspore morphology and elaiosome chemistry contributing to lower seed attractiveness in endemic species (Miller & Kwit, 2018; Miller et al., 2020). While this suggests that endemic species of *Trillium* may experience dispersal limitation stemming from lower seed dispersal effectiveness compared to their more widespread congeners, Reid's

Paradox posits that the geographic ranges of many of these species—both endemic and widespread, and especially those with ranges extending into northern latitudes—cannot be fully explained by observed average seed dispersal rates and distances (Clark et al., 1998; Griffin & Barrett, 2004).

Given the relationship between current distributions of ant-dispersed woodland herbs and the maximum extent of the glaciers (Cain et al., 1998), the migration of plants such as those in the genus *Trillium* must rely on some unknown form of long-distance seed dispersal (Cain et al., 1998; Griffin & Barrett, 2004). Long-distance dispersal (LDD) by white-tailed deer (*Odocoileus virginianus*) has been observed for *Trillium grandiflorum*, which might account for the widespread geographic range of this species (Vellend et al., 2003; Kartesz, 2015). Whether other species of *Trillium* similarly obtain LDD from either white-tailed deer or some other vector is unknown. I posit that variation in dispersal potential stemming from all sources, including ants as primary dispersers and any potential LDD vectors, is strongly related to range size in plants of the genus *Trillium.* Sources of LDD and of variation in the quantity and quality of LDD services for myrmecochorous plants could be the focus of future investigations.

SDM performance was related to *Trillium* species' conservation status, RCS, and PSA. I found that a model's ability to successfully classify training points (measured by standardized AUC scores) was significantly dependent on the interacting effects of conservation status and the size of the PSA. *Trillium* species of greater conservation concern (G3/G4) with smaller PSA, such as *T. catesbaei, T. ludovicianum, T. nivale,* and *T. stamineum,* had lower model performance than species of greater conservation concern with larger PSA, such as *T. discolor* and *T. decumbens.* The opposite trend was apparent

for species of lower conservation concern (G4G5/G5); those with smaller PSA, such as *T. decipiens*, had slightly higher model performance than those with larger PSA, such as *T. undulatum, T. cernuum,* and *T. cuneatum.* This finding did not support my hypothesis; I anticipated that species of lower conservation concern predicted to occupy a larger spatial extent (i.e. "widespread" species) would have higher model fit. Furthermore, I uncovered a more complex relationship between model performance (as measured by standardized AUC), range size, and conservation status than anticipated. While EOO and PSA are simple measures of known and expected range size, the categorical variable of conservation status encompasses all of the factors NatureServe takes into consideration during species assessment. These include the area in which a species is present; the number of occurrences in the area of interest; the current wild population size of the species; the EOO and area of occupancy (AOO) as described by IUCN; short- and long-term trends in population size, EOO, AOO, number of occurrences, and/or condition of occurrences; the degree to which the species is directly or indirectly threatened; the number of occurrences that are protected and managed for the long-term persistence; the degree to which intrinsic factors make a species vulnerable to natural or anthropogenic stresses; and the vulnerability of the species due to habitat preferences or environmental specificity (Master et al., 2003; Regan et al., 2004). Thus, the categorization of species as "low" (G3/G4) or "high" (G4G5/G5) conservation concern encompasses many additional factors that likely influenced the results of this analysis.

However, if these results are distilled down to the notion that Maxent (as opposed to the majority of other common SDM algorithms) performs at a higher standard for "rarer" species, there is support in the literature. For example, Hernandez et al. (2006)

found that, of four common modeling algorithms, Maxent had the highest performance for rare species. Phillips et al. (2006), Elith et al. (2006), Rebelo and Jones (2010) and Elith et al. (2011) all found that Maxent performs at a high standard for species with fewer presence-only occurrences. In this system, PSA is positively related to EOO, meaning the species of *Trillium* with higher model fit are both known and expected to occupy a smaller spatial extent. Furthermore, these species had fewer occurrence records, confirming that Maxent yields robust predictions for rarer, more at-risk species. Despite the significantly lower standardized AUC scores for more widespread species, overall model performance (including both testing AUC and omission error) was relatively high for all species in this study.

It is possible that the differences in SDM performance captured by standardized AUC scores can be explained by the classification of narrowly endemic species as high conservation priority simply *because* they are narrowly endemic. The system by which NatureServe ranks species for conservation priority relies heavily on known or inferred population sizes, numbers of occurrences, and geographic range sizes; these factors may skew a narrowly endemic species towards high conservation priority without the species necessarily experiencing significant population declines or being intrinsically at-risk to external threats. A useful framework for considering species rarity was put forth by Rabinowitz (1981). The framework describes "seven forms of rarity," whereby a species is categorized according to its geographic range size, habitat specificity, and local population size. Narrowly endemic species are those with both a narrow geographic range and narrow habitat specificity. However, these species can additionally be characterized by large, dominant local population sizes, as is the case for several of the

endemics in this study (e.g. *T. discolor, T. decumbens*). In these cases, it is worth considering whether traits characterizing an endemic species are geared towards successful local population persistence (as in Lavergne et al., 2004) rather than maximizing dispersal distances or colonizing novel habitats. Williams et al. (2009) characterized such species as environmental specialists, and provided evidence that small range size and narrow habitat specificity can be successful traits as long as they are offset by demographic commonness.

Training presence omission error was found to be dependent on the size of the PSA, but not on NatureServe conservation status. As PSA increased, training presence omission error decreased significantly across all species, regardless of conservation status. I detected a significant interaction effect between PSA and RCS on omission error. Rare, medium-ranged, and widespread *Trillium* species displayed three distinct relationships between omission error and PSA: omission error was negatively related to PSA for rare species, positively related to PSA for medium-range species, and apparently unrelated to PSA for widespread species. Because the majority of species in this study were categorized as rare, these species drove the observed overall negative relationship between omission error and PSA. These findings, unlike those pertaining to standardized AUC scores, partially supported my hypothesis. While species with larger ranges had significantly lower omission error (and therefore higher model performance) than species with smaller ranges, I did not find that conservation status was a significant predictor of a model's ability to correctly predict presences as present at the 10% training threshold. The significant interaction between RCS and PSA on omission error again highlights the

fact that the relationships between model performance and species characteristics are not simple, and are likely dependent on a wide range of intrinsic and extrinsic traits.

There was some support for my hypothesis that reproductive life history traits would predict SDM fit. Ovule number and flower type were significant predictors of standardized testing AUC, although none of the reproductive life history traits were significant predictors of omission error. Ovule number was negatively related to model performance as measured by standardized AUC scores. This predictor was significantly, positively correlated with both seed setting rate and number of seeds per plant, so it is likely that all three of these factors influence SDM performance. Number of ovules per flower, number of seeds per plant, and seed setting rate are all measures of female fertility. Larger values of these traits, and high seed : ovule ratios in particular, are indicative of greater female reproductive output, which can be impacted by pollen and resource limitation and plant size, as exemplified by experiments on *T. grandiflorum* (Griffin & Barrett, 2002). Greater female reproductive output could also be related to larger range size (Peat & Fitter, 1994; Eriksson & Jakobsson, 1998); this connection might explain the inverse relationship we found between ovule number and standardized AUC scores. Consistent with the conclusion that Maxent performs better for rarer species, these results suggest that *Trillium* species with lower female reproductive output may be better candidates for species distribution modeling using Maxent.

Pedicellate-erect and pedicellate-declinate species had significantly lower standardized testing AUC scores than sessile species. Pedicellate-flowered species belong to the subgenus *Trillium* L.*,* a paraphyletic group consisting of two informal, ostensibly monophyletic groups, the “erectum group” and the “delostylis group,” spanning North

America and Asia (Farmer & Schilling, 2002; Millam, 2006). In eastern North America, many of these species are geographically widespread (e.g. *T. cernuum, T. erectum, T. grandiflorum,* and *T. flexipes*), while others are narrowly endemic to the southern Appalachian region (e.g. *T. sulcatum, T. vaseyi*). Sessile-flowered species belong to the subgenus *Sessilium* Raf.*,* formerly *Phyllantherum* Raf.*,* and comprise a monophyletic clade found only in North America (Farmer & Schilling, 2002; Millam, 2006). Many sessile species are narrowly endemic to the southern Appalachian region (e.g. *T. discolor, T. decumbens*)*,* or the southern coastal plain region of the U.S. (e.g. *T. underwoodii, T. ludovicianum, T. foetidissimum*). Species of *Trillium* within each subgenus are distinguished by a number of vegetative and floral traits, including odor, color, and presentation of the stigma, stamens, and anthers, and although they all grow in rich hardwood forests, they are characterized by microhabitat differences such as elevation or edaphic factors (Patrick, 1984; Milliam, 2006). Differences driven by phylogeny, which may be tied to macro- and microhabitat, vegetative, floral, or other unknown traits, may explain the finding that model performance differs for species belonging to the three flower types.

It is important to acknowledge a few points about AUC as a measure of model performance. First, while AUC is useful for inter-model comparison across a single species and assessment of a model's ability to discriminate between suitable and unsuitable habitat, this metric is not always useful for specific conservation or management objectives, such as ranking suitable habitat for rare species at specific spatial scales (Gogol-Prokurat, 2011). Second, in its strictest sense, raw AUC scores should not be compared across species with different training regions because increasing

the extent of the training region can cause discrimination measures to artificially increase (Jiménez-Valverde et al., 2008; Barve et al., 2011). It is for this reason that I standardized testing AUC scores by the size of the training extent to generate comparable measures of AUC. However, I am unaware of whether other studies have applied similar transformations to AUC scores prior to making interspecific comparisons of model performance. Indeed, I know of studies that have compared interspecific model fit using AUC without correcting for the spatial extent of the training region (Hanspach et al., 2010; Elith et al., 2011). I recommend against relying heavily on AUC scores to judge the performance or applicability of a model without understanding its limitations.

This study has contributed predicted range maps for 21 species of *Trillium* native to eastern North America based on contemporary climate conditions. Understanding how the contemporary climate defines a species' fundamental niche is a crucial step towards understanding essential environmental conditions for the maintenance of populations, which are being altered by climate change. Contemporary distribution models can identify abiotic variables that have historically been instrumental in determining where species are able to persist across the landscape. This type of information could be used to inform management plans for the conservation of species in the genus *Trillium*—charismatic spring ephemerals of significant horticultural and scientific interest—in eastern North America. Similar approaches to understanding geographic ranges could result in comparable planning efforts in other vascular plant species of conservation interest.

CONCLUSIONS AND RECOMMENDATIONS

This dissertation is comprised of three related studies that contribute significantly to our understanding of the variation within, and the mechanisms underlying, ant-mediated seed dispersal—a fine-scale, biotic interaction that is potentially related to the geographic distributions of plants in the genus *Trillium*—and to our understanding of the coarse-scale, abiotic drivers of *Trillium* species' ranges. Chapter 1 revealed lower quantitative and qualitative seed dispersal effectiveness for endemic *Trillium* species compared to their sympatric, widespread congeners. For three distinct endemic-widespread species pairs across the southern Appalachian region, endemic species had significantly lower numbers of seeds dispersed by ants under natural conditions than their widespread congeners. Laboratory experiments revealed that ants belonging to the keystone seed-dispersing genus in eastern North America (*Aphaenogaster*) demonstrate a consistent, significant preference for widespread seed species, thus resulting in fewer numbers of seeds dispersed for endemic species and, correspondingly, fewer seeds' elaiosomes removed—a treatment that has been shown to positively impact seed germination. These results are consistent with fine-scale dispersal limitation for narrowly endemic species of *Trillium* as compared to their widespread congeners.

Chapter 2 investigated the morphological and chemical mechanisms underlying the observed preference displayed by seed-dispersing ants for widespread trilliums in Chapter 1. This study characterized the morphological and chemical attributes of diaspores and elaiosomes produced by trillium seeds that are most likely to be dispersed by ants in the southern Appalachian region. Specifically, seeds with a higher probability

of dispersal tended to have larger diaspores and higher concentrations of key and total elaiosome fatty acids. The potential roles played by previously-unknown components of the broader elaiosome phytochemical profile in stimulating dispersal by ants are novel contributions to this field of study, and solicit additional future research aimed at characterizing and quantifying elaiosome phytochemical diversity and its impact on myrmecochory.

Chapter 3 diverged from seed dispersal and focused on geographic distributions of *Trillium* species at a regional scale. This study had three primary takeaways. First, it established that the average proportional occupancy of the fundamental niche for most species of *Trillium* is very low, and that rarer species of greater conservation concern tend to have lower proportional occupancy of their potentially suitable area (PSA) than more common species of lower conservation concern. Therefore, factors not included in SDMs play a larger role in shaping the geographic extent of rare eastern North American species of *Trillium* than their widespread congeners. This provides some support for the working hypothesis established by the first and second chapters: that endemic species of *Trillium* may be dispersal limited due to less-attractive seeds. However, this support is limited because it is impossible to disentangle whether the non-abiotic limits on geographic distributions are due to species interactions, biotic characteristics of the environment, or dispersal limitations. Second, this study found that a significant interaction between conservation status and PSA predicted model performance; performance was positively related to the size of the PSA for species of greater conservation concern and negatively related to PSA for species of lower conservation concern. In this system, PSA was significantly, positively related to the known range,

meaning species with higher model performance are both known and expected to occupy a smaller spatial extent. These species also had fewer occurrences, confirming that the species distribution modeling algorithm Maxent yields robust predictions of spatial distributions for rare, more at-risk species. Third, this study determined that flower type and ovule number were significant predictors of model performance. This result might be explained by phylogeny or differences in reproductive output that may lead to larger rates of dispersal or greater numbers of occurrences. This information is pertinent to conservation management of species in the genus *Trillium*, and may be relevant if extrapolated to planning efforts for other range-restricted understory plant species of conservation interest.

In conclusion, I make three principal recommendations based on this body of work. First, because successful dispersal of sufficient seeds to safe sites is of crucial importance to the persistence of metapopulations (Washitani et al., 1997), conservation efforts for endemic species belonging to the genus *Trillium* should consider translocation as a means of compensating for low quantitative and qualitative seed dispersal effectiveness. Translocation planning should include both pre- and post-translocation considerations (Godefroid et al., 2016). Second, to advance our understanding of seed dispersal effectiveness in animal-mediated seed dispersal systems, seed phytochemical diversity should be granted a significant amount of attention. The study of plant phytochemical diversity has progressed considerably in recent years due to the use of species diversity metrics to quantify phytochemical diversity (Wetzel and Whitehead, 2020). However, to my knowledge, very little research thus far has implemented species diversity metrics in the study of the effects of seed phytochemistry on seed dispersal

effectiveness (but see Whitehead et al., 2013); this could constitute a field of untapped knowledge that is ripe for the picking. Third, and finally, I posit that variation in dispersal potential stemming from all sources, including ants as primary dispersers and any potential long-distance dispersal (LDD) vectors, is strongly related to range size in plants of the genus *Trillium.* Sources of LDD and of variation in the quantity and quality of LDD services for myrmecochorous plants should be the focus of future investigations, given that high post-glaciation migration rates (Delcourt and Delcourt, 1991) and the relationship between current distributions of ant-dispersed woodland herbs and the maximum extent of the glaciers (Cain et al., 1998) strongly indicate that the migration of plants such as those in the genus *Trillium* must rely on some unknown form of LDD (Cain et al., 1998; Griffin and Barrett, 2004).

Although this body of work establishes some support for the notion that fine-scale dispersal limitation stemming from low seed dispersal effectiveness limits the geographic distributions of endemic species of *Trillium,* many other factors that were not considered in this dissertation likely play significant roles in shaping *Trillium* species ranges. Such variables may include historic biogeographical factors; species interactions such as parasitism, disease, herbivory, and competition; and human-induced changes such as habitat fragmentation.

LIST OF REFERENCES

AdaptWest Project. (2015). Gridded current and future climate data for North America at 1km resolution, interpolated using the ClimateNA v5.10 software (T. Wang et al., 2015). adaptIst.databasin.org.

Anderson, R. C. (1994). Height of white-flowered trillium (*Trillium grandiflorum*) as an index of deer browsing intensity. *Ecological Applications*, *4*(1), 104-109.

Aguirre-Gutiérrez, J., Serna-Chavez, H. M., Villalobos-Arambula, A. R., Perez de la Rosa, J. A., and Raes, N. (2015). Similar but not equivalent: ecological niche comparison across closely–related Mexican white pines. *Diversity and Distributions*, *21*(3), 245-257.

Araújo, M. B., and Luoto, M. (2007). The importance of biotic interactions for modelling species distributions under climate change. *Global Ecology and Biogeography*, *16*(6), 743-753.

Araújo, M. B., Anderson, R. P., Barbosa, A. M., Beale, C. M., Dormann, C. F., Early, R., ... and O'Hara, R. B. (2019). Standards for distribution models in biodiversity assessments. *Science Advances*, *5*(1), eaat4858.

ArcGIS [GIS software]. Version 10.7. Redlands, CA: Environmental Systems Research Institute, Inc., 2010.

Arnold Arboretum of Harvard University, The. (2020). The President and Fellows of Harvard College. https://www.arboretum.harvard.edu/. [Aug. 17, 2018].

Bale, M. T., Zettler, J. A., Robinson, B. A., Spira, T. P. and Allen, C. R. (2003). Yellow jackets may be an underestimated component of an ant–seed mutualism. *Southeastern Naturalist, 2,* 609-614.

Barbehenn, R. V., Reese, J. C., and Hagens, K. S. (1999). The food of insects. In: HuVaker, C. B., Gutierrez, A. P. (eds.) Ecological entomology, 2nd edn. Wiley, New York, pp 83–121. ISBN: 047124483X, 9780471244837.

Barve, N., Barve, V., Jiménez-Valverde, A., Lira-Noriega, A., Maher, S. P., Peterson, A. T., ... and Villalobos, F. (2011). The crucial role of the accessible area in ecological niche modeling and species distribution modeling. *Ecological Modelling*, *222*(11), 1810-1819.

Bas, J. M., Oliveras, J., and Gómez, C. (2009). Myrmecochory and short-term seed fate in *Rhamnus alaternus:* ant species and seed characteristics. *Acta Oecologica, 35*, 380-384.

Baskin, C. C., and Baskin, J. M. (1988). Germination ecophysiology of herbaceous plant species in a temperate region. *American Journal of Botany, 75*(2), 286-305.

Baskin, J. M., Snyder, K. M., Walck, J. L., and Baskin, C. C. (1997). The comparative autecology of endemic, globally-rare, and geographically-widespread, common plant species: three case studies. *Southeast Naturalist, 42,* 384–99.

Bates, D., Maechler, M., Bolker, B. and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67,* 1–48.

Beattie, A. J. (1985). *The evolutionary ecology of ant-plant mutualisms*. Cambridge University Press.

Bevill, R. L., and Louda, S. M. (1999). Comparisons of related rare and common species in the study of plant rarity. *Conservation Biology*, *13*(3), 493-498.

Bevill, R. L., Louda, S. M., and Stanforth, L. M. (1999). Protection from natural enemies in managing rare plant species. *Conservation Biology*, *13*(6), 1323-1331.

Bielby, J., Cardillo, M., Cooper, N., and Purvis, A. (2010). Modelling extinction risk in multispecies data sets: Phylogenetically independent contrasts versus decision trees. *Biodiversity and Conservation, 19,* 113–127.

Biau, G. (2012). Analysis of a random forests model. *Journal of Machine Learning Research, 13,* 1063–1095.

Bluthgen, N., and Fiedler, K. (2004). Competition for composition: lessons from nectar-feeding ant communities. *Ecology, 85,* 1479–1485.

Bluthgen, N., Gottsberger, G., and Fiedler, K. (2004). Sugar and amino acid composition of ant-attended nectar and honeydew sources from an Australian rainforest. *Austral Ecology, 29*, 418–429.

Boieiro, M., Espadaler, X., Gómez, C., and Eustaquio, A. (2012). Spatial variation in the fatty acid composition of elaiosomes in an ant-dispersed plant: differences within and between individuals and populations. *Flora-Morphology, Distribution, Functional Ecology of Plants, 207*(7), 497-502.

Bond, W. J. (1994). Do mutualisms matter? Assessing the impact of pollinator and disperser disruption on plant extinction. *Philosophical Transactions of the Royal Society, B, Biological Sciences, 344*, 83-90.

Boulay, R., Coll-Toledano, J., and Cerdá, X. (2006). Geographic variations in *Helleborus foetidus* elaiosome lipid composition: implications for dispersal by ants. *Chemoecology, 16,* 1–7.

Boulay, R., Coll-Toledano, J., Manzaneda, A. J., and Cerdá, X. (2007). Geographic variations in seed dispersal by ants: are plant and seed traits decisive?. *Naturwissenschaften, 94(*3), 242-246.

Boyd, R. S. (1996). Ant-mediated seed dispersal of the rare chaparral shrub *Fremontodendron decumbens* (Sterculiaceae). *Madrono*, 299-315.

Bresinsky, A. (1963). Bau, Entwicklungsgeschichte und Inhaltsstoffe der-Elaiosomen.

Brew, C. R., O'Dowd, D. J., and Rae, I. A. (1989). Seed dispersal by ants: behavior-releasing compounds in elaiosomes. *Oecologia, 80*, 490–497.

Brown, E., and Kephart, S. (1999). Variability in pollen load: Implications for reproduction and seedling vigor in a rare plant, *Silene douglasii* var. Oraria. International Journal of Plant Sciences, 160, 1145–52.

Brown, J., Enright, N. J., and Miller, B. P. (2003). Seed production and germination in two rare and three common co-occurring *Acacia* species from southeast Australia. *Austral Ecology 28*(3): 271-280.

Brown, J. L. (2014). SDMtoolbox: a python-based GIS toolkit for landscape genetic, biogeographic and species distribution model analyses. *Methods in Ecology and Evolution*, *5*(7), 694-700.

Brown, J. L., Bennett, J. R., and French, C. M. (2017). SDMtoolbox 2.0: the next generation Python-based GIS toolkit for landscape genetic, biogeographic and species distribution model analyses. *PeerJ*, *5*, e4095.

Buono, R. A., De Oliveira, A. B., Sousa Paiva, E. A. (2008). Anatomy, ultrastructure and chemical composition of food bodies of *Hovenia dulcis* (Rhamnaceae). *Annals of Botany, 101,* 1341–1348.

Byers, D. L., and Meagher, T. R. (1997). A comparison of demographic characteristics in a rare and a common species of *Eupatorium*. *Ecological Applications 7*(2): 519-530.

Cabe, P. R. (1995). The *Trillium pusillum* Michaux (Liliaceae) complex in Virginia. I. Morphological investigations. *Castanea*, 1-14.

Cabe, P. R., and Werth, C. (1995). The *Trillium pusillum* Michaux (Liliaceae) complex in Virginia. II. Isozyme evidence. *Castanea*, 15-29.

Cain, M. L., Damman, H. and Muir, A. (1998.) Seed dispersal and the Holocene migration of woodland herbs. *Ecological Monographs 68*, 325– 347.

Canavoso, L. E., Jouni, Z. E., Karnas, K. J., Pennington, J. E., and Ills, M. A. (2001). Fat metabolism in insects. *Annual Review of Nutrition, 21*, 23–46.

Canner, J., Dunn, R., Giladi, I., and Gross, K. (2012). Redispersal of seeds by a keystone ant augments the spread of common wildflowers. *Acta Oecologica 40*: 31-39.

Case, F. W., and Case, R. B. (1997). *Trillium*. Portland, Ore: Timber Press.

Case, F. W. (2002a.). "*Trillium* subg. *Trillium*". In Flora of North America Editorial Committee (ed.). *Flora of North America North of Mexico (FNA)*. 26. New York and Oxford – via eFloras.org, Missouri Botanical Garden, St. Louis, MO and Harvard University Herbaria, Cambridge, MA.

Case, F. W. (2002b.). *"Trillium subg.Phyllantherum".* In Flora of North America Editorial Committee (ed.). *Flora of North America North of Mexico (FNA)*. 26. New York and Oxford – via eFloras.org, Missouri Botanical Garden, St. Louis, MO and Harvard University Herbaria, Cambridge, MA.

Caut, S., Jowers, M. J., Cerdá, X., and Boulay, R. R. (2013). Questioning the mutual benefits of myrmecochory: a stable isotope-based experimental approach. *Ecological Entomology, 38*(4), 390-399.

Champely, S. (2017). pwr: Basic Functions for Power Analysis. R package version 1.2-1. https://CRAN.R-project.org/package=pwr.

Christie, W. W. (2005). The lipid library. http://www.lipidlibrary.co.uk.

Cipollini, M. L., and D. J. Levey. (1997). Secondary metabolites of fleshy vertebrate-dispersed fruits: adaptive hypotheses and implications for seed dispersal. *The American Naturalist 150(*3): 346-372.

Clark, J. S., Fastie, C., Hurtt, G., Jackson, S. T., Johnson, C., King, G. A., ... and Schupp, E. W. (1998). Reid's paradox of rapid plant migration: dispersal theory and interpretation of paleoecological records. *BioScience*, *48*(1), 13-24.

Clark, J. S., Gelfand, A. E., Woodall, C. W., and Zhu, K. (2014). More than the sum of the parts: forest climate response from joint species distribution models. *Ecological Applications*, *24*(5), 990-999.

Clark, J. S., Nemergut, D., Seyednasrollah, B., Turner, P. J., and Zhang, S. (2017). Generalized joint attribute modeling for biodiversity analysis: Median-zero, multivariate, multifarious data. *Ecological Monographs*, *87*(1), 34-56.

Clasquin, M. F., Melamud, E., and Rabinowitz, J. D. (2012). LC-MS data processing with MAVEN: a metabolomic analysis and visualization engine. *Current Protocols in Bioinformatics, 37*(1), 14-11.

Combs, J. K., Lambert, A. M., and Reichard, S. H. (2013). Predispersal seed predation is higher in a rare species than in its widespread sympatric congeners (*Astralagus*, Fabaceae). *American Journal of Botany, 100*, 2149-2157.

Combs, J. K., Reichard, S. H., Groom, M. J., Wilderman, D. L., and Camp, P. A. (2011). Invasive competitor and native seed predators contribute to rarity of the narrow endemic *Astragalus sinuatus* Piper. *Ecological Applications, 21*(7), 2498-2509.

Cuautle, M., Rico-Gray, V., and Díaz-Castelazo, C. (2005). Effects of ant behavior and presence of extrafloral nectaries on seed dispersal of the Neotropical myrmecochore *Turnera ulmifolia* L. (Turneraceae). *Biological Journal of the Linnean Society, 86*: 67-77.

Culver, D. C., and Beattie, A. J. (1978). Myrmecochory in *Viola*: dynamics of seed–ant interactions in some west Virginia species. *Journal of Ecology, 66,* 53–72.

Cunningham, S. A. (2000). Depressed pollination in habitat fragments causes low fruit set. *Proceedings of the Royal Society, B, Biological Sciences, 267*, 1149–52.

Dadd, R. H. (1973). Insect nutrition: current developments and metabolic implications. *Annual Review of Entomology, 18,* 381–429.

Darlington, C. D., and LaCour, L. (1940). Nucleic acid starvation of chromosomes in *Trillium*. *Journal of Genetics*, *40*(01-02).

Davidson, D. W., and Morton, S. R. (1981). Competition for dispersal in ant-dispersed plants. *Science, 213*, 1259-1261.

Davidson, D. W., Seidel, J. L., and Epstein, W. W. (1990). Neotropical ant gardens II. Bioassays of seed compounds. *Journal of Chemical Ecology, 16*(10), 2993-3013.

Davidson, D. W., Foster, R. B., Snelling, R. R., and Lozada, P. W. (1991). Variable composition of some tropical ant–plant symbioses. In: Price, P. W., Lewinsohn, T. M., Fernandes, G. W., Benson, W. W., eds. Plant-animal interactions: evolutionary ecology in tropical and temperate regions. New York, USA, Wiley, 145–162.

Delcourt, H. R., and Delcourt, P. A. (1991). Quaternary ecology: a paleoecological perspective. Chapman and Hall, New York, New York, USA.

DeMarco, B. B., and Cognato, A. I. (2016). A multiple-gene phylogeny reveals polyphyly among eastern North American *Aphaenogaster* species (Hymenoptera: Formicidae). *Zoologica Scripta, 45*(5), 512-520.

Dinsdale, J. M., Dale, M. P., and Kent, M. (2000). Microhabitat availability and seedling recruitment of *Lobelia urens:* a rare plant species at its geographic limit. *Seed Science Research, 10*: 471–87.

Dobzhansky, T. (1950). Mendelian populations and their evolution. *The American Naturalist, 84,* 401–418.

Dussutour, A., and Simpson, S. J. (2008). Description of a simple synthetic diet for studying nutritional responses in ants. *Insectes Sociaux* 55: 329-333.

Edwards, W., Dunlop, M., and Rodgerson, L. (2006). The evolution of rewards: seed dispersal, seed size and elaiosome size. *Journal of Ecology, 94*, 687–694.

Elith, J., Graham, C. H., Anderson, R. P., Dudík, M., Ferrier, S., Guisan, A., ... and Li, J. (2006). Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, *29*(2), 129-151.

Elith, J., and Leathwick, J. R. (2009). Species distribution models: ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics*, *40*, 677-697.

Elith, J., Phillips, S. J., Hastie, T., Dudík, M., Chee, Y. E., and Yates, C. J. (2011). A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions*, *17*(1), 43-57.

Ellstrand, N. C., and Elam, D. R. (1993). Population genetic consequences of small population size: implications for plant conservation. *Annual Review of Ecology and Systematics*, *24*(1), 217-242.

Engler, R., Guisan, A., and Rechsteiner, L. (2004). An improved approach for predicting the distribution of rare and endangered species from occurrence and pseudo-absence data. *Journal of Applied Ecology*, *41*(2), 263-274.

Enright N. J., Lamont, B. B., and Marsula, R. (1996). Canopy seed bank dynamics and optimum fare regime for the highly serotinous shrub, *Banksia hookeriana. Journal of Ecology, 84*, 9–17.

Eriksson, O., & Jakobsson, A. (1998). Abundance, distribution and life histories of grassland plants: a comparative study of 81 species. *Journal of Ecology*, *86*(6), 922-933.

Falk, D. A., Millar, C. I., and Olwell, M. (1996). Restoring diversity: strategies for reintroduction of endangered plants.

Farmer, S. B., & Schilling, E. E. (2002). Phylogenetic analyses of Trilliaceae based on morphological and molecular data. *Systematic Botany*, *27*(4), 674-692.

Farmer, S. B. (2007). Systematics of *Trillium* subgenus *Delostylis*. PhD Dissertation, University of Tennessee, Knoxville TN.

Feng, X., Park, D. S., Walker, C., Peterson, A. T., Merow, C., and Papeş, M. (2019). A checklist for maximizing reproducibility of ecological niche models. *Nature Ecology and Evolution,* 3, 1382–1395.

Feunang, Y. D., Eisner, R., Knox, C., Chepelev, L., Hastings, J., OIn, G.,... and Greiner, R. (2016). ClassyFire: automated chemical classification with a comprehensive, computable taxonomy. *Journal of Cheminformatics, 8*(1), 61.

Fiedler, P. L. (1987). Life history and population dynamics of rare and common Mariposa lilies (*Calochortus pursh*: Liliaceae). *Journal of Ecology, 75*, 977-995.

Fielding, A. H., and Bell, J. F. (1997). A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, *24*(1), 38-49.

Fischer, M., and Matthies, D. (1998). RAPD variation in relation to population size and plant fitness in the rare *Gentianella germanica* (Gentianaceae). *American Journal of Botany*, *85*(6), 811-819.

Fischer, R. C., Richter, A., Wanek, W., and Mayer, V. (2002). Plants feed ants: food bodies of myrmecophytic *Piper* and their significance for the interaction with *Pheidole bicornis* ants. *Oecologia, 133*, 186–192.

Fischer, R. C., Ölzant, S. M., Wanek, W., and Mayer, V. (2005). The fate of *Corydalis cava* elaiosomes within an ant colony of *Myrmica rubra*: elaiosomes are preferentially fed to larvae. *Insectes Sociaux, 52*(1), 55-62.

Fischer, R. C., Richter, A., Hadacek, F., and Mayer, V. (2008). Chemical differences between seeds and elaiosomes indicate an adaptation to nutritional needs of ants. *Oecologia, 155*, 539–547.

Fokuhl, G., Heinze, J., and Poschlod, P. (2007). Colony growth in *Myrmica rubra* with supplementation of myrmecochorous seeds. *Ecological Research, 22*(5), 845-847.

Fox, J., and Weisberg, S. (2011). An {R} companion to applied regression, 2nd ed. Sage, Thousand Oaks, CA, USA. http://socserv.socsci.mcmaster.ca/jfox/Books/Companion.

Freeman, J. D. (1975). Revision of *Trillium* subgenus Phyllantherum (Liliaceae). *Brittonia, 27*(1), 1-62.

Fricke, E. C., Haak, D. C., Levey, D. J., and Tewksbury, J. J. (2016). Gut passage and secondary metabolites alter the source of post-dispersal predation for bird-dispersed chili seeds. *Oecologia, 181*(3), 905-910.

Gaddy, L. (1986). Twelve new ant-dispersed species in the Southern Appalachians. *Bulletin of the Torrey Botanical Club, 113*, 247–251.

Gammans, N., Bullock, J. M., Gibbons, H., and Schönrogge, K. (2006). Reaction of mutualistic and granivorous ants to *Ulex* elaiosome chemicals. *Journal of Chemical Ecology, 32*(9), 1935-1947.

Gates, R. R. (1917). A systematic study of the North American genus *Trillium*, its variability, and its relation to *Paris* and *Medeola*. *Annals of the Missouri Botanical Garden*, *4*(1), 43-92.

GBIF: The Global Biodiversity Information Facility (2020). What is GBIF?. Available from https://www.gbif.org/what-is-gbif. [July 28 - Aug. 1, 2018].

Godefroid, S., Le Pajolec, S., and Van Rossum, F. (2016). Pre-translocation considerations in rare plant reintroductions: implications for designing protocols. *Plant Ecology*, *217*(2), 169-182.

Gogol-Prokurat, M. (2011). Predicting habitat suitability for rare plants at local spatial scales using a species distribution model. *Ecological Applications*, *21*(1), 33-47.

Gonzalez-Teuber, M., and Heil, M. (2009). The role of extrafloral nectar amino acids for the preferences of facultative and obligate ant mutualists. *Journal of Chemical Ecology, 35*, 459–468.

Gordon, S. C., Meadley-Dunphy, S. A., Prior, K. M., and Frederickson, M. E. (2019). Asynchrony between ant seed dispersal activity and fruit dehiscence of myrmecochorous plants. *American Journal of Botany, 106*(1), 71-80.

Gorb, S. N., and Gorb, E. V. (1995). Removal rates of seeds of five myrmecochorous plants by the ant *Formica polyctena* (Hymenoptera: Formicidae). *Oikos,* 367-374.

Gorb, E. V., and Gorb, S. N. (2003). Seed dispersal by ants in a deciduous forest ecosystem, 1st ed. Kluwer, Dordrecht, Netherlands.

Gove, A. D., Majer, J. D., and Dunn, R. R. (2007). A keystone ant species promotes seed dispersal in a "diffuse" mutualism. *Oecologia, 153*(3), 687-697.

Griffin, S. R., & Barrett, S. C. (2002). Factors affecting low seed: ovule ratios in a spring woodland herb, *Trillium grandiflorum* (Melanthiaceae). *International Journal of Plant Sciences*, *163*(4), 581-590.

Griffin, S. R., and Barrett, S. C. (2004). Post-glacial history of *Trillium grandiflorum* (Melanthiaceae) in eastern North America: Inferences from phylogeography. *American Journal of Botany*, *91*(3), 465-473.

Grinnell, J. (1917). The niche-relationships of the California Thrasher. *Auk, 34*(4), 427-433.

Guisan, A., and Zimmermann, N. E. (2000). Predictive habitat distribution models in ecology. *Ecological Modelling*, *135*(2-3), 147-186.

Guisan, A., and Thuiller, W. (2005). Predicting species distribution: offering more than simple habitat models. *Ecology Letters*, *8*(9), 993-1009.

Gunther, R. W., and Lanza, J. (1989). Variation in attractiveness of *Trillium* diaspores to a seed-dispersing ant. *American Midland Naturalist,* 321-328.

Haga, T., and Kurabayashi, M. (1953). Genom and Polyploidy in the Genus *Trillium.* IV. *Cytologia*, *18*(1), 13-28.

Hagen, K. S., Dadd, R. H., and Reese, J. (1984). The food of insects. In: HuVaker C. B., Rabb R. L. (eds.) *Ecological Entomology.* Wiley, New York, pp. 79–111.

Hamann, A., Wang, T., Spittlehouse, D. L., and Murdock, T. Q. (2013). A comprehensive, high-resolution database of historical and projected climate surfaces for Western North America. *Bulletin of the American Meteorological Society, 94*(9), 1307-1309.

Hand, R., Grossmann, A., and Lauterbach, D. (2017). Endemics and their common congener plant species on an East Mediterranean island: a comparative functional trait approach. *Plant Ecology, 218*, 139-150.

Handel, S. N., and Beattie, A. J. (1990). Seed dispersal by ants. Scientific American, 263, 76-83B.

Hanspach, J., Kühn, I., Pompe, S., and Klotz, S. (2010). Predictive performance of plant species distribution models depends on species traits. *Perspectives in Plant Ecology, Evolution and Systematics*, *12*(3), 219-225.

Harrison, P., Berry, P. M., Butt, N., and New, M. (2006). Modelling climate change impacts on species' distributions at the European scale: implications for conservation policy. *Environmental Science and Policy*, *9*(2), 116-128.

Heikkinen, R. K., Luoto, M., Virkkala, R., Pearson, R. G., and Körber, J. H. (2007). Biotic interactions improve prediction of boreal bird distributions at macro-scales. *Global Ecology and Biogeography*, *16*(6), 754-763.

Heil, M., Fiala, B., Kaiser, W., and Linsenmair, K. E. (1998). Chemical contents of Macaranga food bodies: adaptations to their role in ant attraction and nutrition. *Functional Ecology, 12*, 117–122.

Heil, M., Baumann, B., Kruger, R., and Linsenmair, K. E. (2004). Main nutrient compounds in food bodies of Mexican *Acacia* ant-plants. *Chemoecology, 14*, 45–52.

Heil, M. (2011). Nectar: generation, regulation, and ecological functions. *Trends in Plant Science, 16*, 191–200.

Heithaus, E. R. (1981). Seed predation by rodents on three ant-dispersed plants. *Ecology, 62*(1), 136-145.

Hernandez, P. A., Graham, C. H., Master, L. L., and Albert, D. L. (2006). The effect of sample size and species characteristics on performance of different species distribution modeling methods. *Ecography*, *29*(5), 773-785.

Hijmans, R. J., Phillips, S., Leathwick, J., Elith, J., and Hijmans, M. R. J. (2017). Package ‘dismo’. *Circles*, *9*(1), 1-68.

Hill-Cottingham, D. G., and Lloyd-Jones, C. P. (1973). Metabolism of carbon-14 labelled arginine, citrulline and ornithine in intact apple stems. *Physiologia Plantarum, 29*(1), 125-128.

Hof, A. R., Jansson, R., and Nilsson, C. (2012). How biotic interactions may alter future predictions of species distributions: future threats to the persistence of the arctic fox in Fennoscandia. *Diversity and Distributions*, *18*(6), 554-562.

Honnay, O., and Jacquemyn, H. (2007). Susceptibility of common and rare plant species to the genetic consequences of habitat fragmentation. *Conservation Biology*, *21*(3), 823-831.

Hooker, W. J. (1872). *Icones Plantarum,* Or Figures, with Brief Descriptive Characters and Remarks, of New or Rare Plants, Selected from the Author's Herbarium (Vol. 12). Longman.

Horvitz, C., and Beattie, A. J. (1980). Ant dispersal of *Calathea* (Marantaceae) seeds by carnivorous ponerines (Formicidae) in a tropical rainforest. *American Journal of Botany, 67,* 321-326.

Horvitz, C., and Schemske, D. (1995). Spatiotemporal variation in demographic transitions of a tropical understory herb: projection matrix analysis. *Ecological Monographs, 65,* 155–192.

Hothorn, T., Bretz, F., and Westfall, P. (2008). Simultaneous Inference in General Parametric Models. *Biometrical Journal, 50*(3), 346-363.

Hughes, L., and Westoby, M. (1990). Removal rates of seeds adapted for dispersal by ants. *Ecology, 71*(1), 138-148.

Hughes, L., and Westoby, M. (1992). Effect of diaspore characteristics on removal of seeds adapted for dispersal by ants. *Ecology, 73,* 1300 – 1312.

Hughes, L., Istoby, M. T., and Jurado, E. (1994). Convergence of elaiosomes and insect prey: evidence from ant foraging behavior and fatty acid composition. *Functional Ecology,* 358-365.

Hui, F. K., Warton, D. I., Foster, S. D., and Dunstan, P. K. (2013). To mix or not to mix: comparing the predictive performance of mixture models vs. separate species distribution models. *Ecology*, *94*(9), 1913-1919.

Huntley, B., Berry, P. M., Cramer, W., and McDonald, A. P. (1995). Special paper: modelling present and potential future ranges of some European higher plants using climate response surfaces. *Journal of Biogeography*, 967-1001.

Hutchinson, G. E. (1957). Concluding remarks. *Cold Spring Harbor Symp Quant Biol, 22*, 415–427.

Hutchinson, G. E. (1978). An Introduction to Population Biology (Yale Univ Press, New Haven, CT).

Imdad, M. U. and Aslam, M. (2018). mctest: Multicollinearity Diagnostic Measures. https://CRAN.R-project.org/package=mctest, R package version 1.2.5.

Imdad, M. U., Aslam, M., Altaf, S., and Ahmed, M. (2019). Some New Diagnostics of Multicollinearity in Linear Regression Model. *Sains Malaysiana,* 48(9), 2051-2060. http://dx.doi.org/10.17576/jsm-2019-4809-26.

Imdadullah, M., Aslam, M., and Altaf, S., (2016). mctest: An R Package for Detection of Collinearity among Regressors. The R Journal, 8(2), 499-509. https://Journal.r-project.org/archive/2016/RJ-2016-062/index.html.

iNaturalist.org. (2020). http://www.inaturalist.org.

IUCN. (2001). IUCN Red List Categories and Criteria: Version 3.1. Technical Report, IUCN Species Survival Commission, IUCN, Gland, Switzerland and Cambridge, UK.

Jalali, A., and Jamzad, Z. (1999). Red Data Book of Iran. Research Institute of Forests and Rangelands (RIFR). pp: 1-4.

Jiménez-Valverde, A., Lobo, J. M., and Hortal, J. (2008). Not as good as they seem: the importance of concepts in species distribution modelling. *Diversity and Distributions*, *14*(6), 885-890.

Jules, E. S. (1996). Yellow jackets (*Vespula vulgaris*) as a second seed disperser for the myrmecochorous plant, *Trillium ovatum*. *American Midland Naturalist,* 367-369.

Jules, E. S., and Rathcke, B. J. (1999). Mechanisms of reduced *Trillium* recruitment along edges of old-growth forest fragments. *Conservation Biology*, *13*(4), 784-793.

Kalisz, S., Hanzawa, F. M., Tonsor, S. J., Thiede, D. A., and Voigt, S. (1999). Ant-mediated seed dispersal alters pattern of relatedness in a population of *Trillium grandiflorum. Ecology*, *80*(8), 2620-2634.

Karron, J. (1987). A comparison of levels of genetic polymorphism and self-compatibility in geographically restricted and widespread plant congeners. *Evolutionary Ecology*, 1: 47-58.

Karron, J. D. (1991). Breeding Systems in Rare Plant Species. *Genetics and conservation of rare plants*, 87.

Kartesz, J., and BONAP [Biota of North America Program]. (2015). North American plant atlas. http://bonap.net/napa. Chapel Hill, North Carolina [maps generated from J. Kartesz. (2015). Floristic synthesis of North America, version 1.0 (in press). Biota of North America Program (BONAP).]

Kissling, W. D., Dormann, C. F., Groeneveld, J., Hickler, T., Kühn, I., McInerny, G. J., ... and Singer, A. (2012). Towards novel approaches to modelling biotic interactions in multispecies assemblages at large spatial extents. *Journal of Biogeography*, *39*(12), 2163-2178.

Kleijn, D., Bekker, R. M., Bobbink, R., De Graaf, M. C., and Roelofs, J. G. (2008). In search for key biogeochemical factors affecting plant species persistence in heathland and acidic grasslands: a comparison of common and rare species. *Journal of Applied Ecology*, *45*(2), 680-687.

Klironomos, J. N. (2002). Feedback with soil biota contributes to plant rarity and invasiveness in communities. *Nature*, *417*(6884), 67-70.

Koh, S., Bazely, D. R., Tanentzap, A. J., Voigt, D. R., and Da Silva, E. (2010). Trillium grandiflorum height is an indicator of white-tailed deer density at local and regional scales. *Forest Ecology and Management, 259*(8), 1472-1479.

Kruckeberg, A. R., and Rabinowitz, D. (1985). Biological aspects of endemism in higher plants. *Annual Review of Ecology and Systematics, 16*, 447-479.

Kursa, M. B., and Rudnicki, W. R. (2010). Feature selection with the Boruta package. *Journal of Statistical Software, 36*(11), 1-13.

Kusmenoglu, S., Rockwood, L. L., and Gretz, M. R. (1989). Fatty acids and diacylglycerols from elaiosomes of some ant-dispersed seeds. *Phytochemistry, 28*(10), 2601-2602.

Lamoreaux, J., Morrison, J., Ricketts, T., Olson, D., Dinerstein, E., McKnight, M., and Shugart, H. (2006). Global tests of biodiversity concordance and the importance of endemism. *Nature, 440*, 212.

Lanza, J., Schmitt, M. A., and Awad, A. B. (1992). Comparative chemistry of elaiosomes of three species of *Trillium*. *Journal of Chemical Ecology, 18*(2), 209-221.

Lavergne, S., Thompson, J. D., Garnier, E., & Debussche, M. (2004). The biology and ecology of narrow endemic and widespread plants: a comparative study of trait variation in 20 congeneric pairs. *Oikos*, *107*(3), 505-518.

Lavergne, S., Thuiller, W., Molina, J., and Debussche, M. (2005). Environmental and human factors influencing rare plant local occurrence, extinction and persistence: a 115-year study in the Mediterranean region. *Journal of Biogeography*, *32*(5), 799-811.

Leal, L., Lima Neto, M., de Oliveira, A., Andersen, A., and Leal, I. (2014). Myrmecochores can target high-quality disperser ants: variation in elaiosome traits and ant preferences for myrmecochorous Euphorbiaceae in Brazilian Caatinga. *Oecologia, 174*, 493–500.

Leege, L. M., Thompson, J. S., and Parris, D. J. (2010). The responses of rare and common trilliums (*Trillium reliquum*, *T. cuneatum*, and *T. maculatum*) to deer herbivory and invasive honeysuckle removal. *Castanea, 75*(4), 433-443.

Legendre, P., and Legendre, L. (1998). Numerical Ecology, 2nd ed. Amsterdam: Elsevier. ISBN 978-0444892508.

Leitao, R. P., Zuanon, J., Villeger, J., Williams, S. E., Baraloto, C., Fortunel, C., Mendonca, F. P., and Mouillot, D. (2016). Rare species contribute disproportionately to the functional structure of species assemblages. *Proceedings of the Royal Society, B, Biological Sciences,* 283:1828.

Lengyel, S., Gove, A. D., Latimer, A. M., Majer, J. D., and Dunn, R. R. (2010). Convergent evolution of seed dispersal by ants, and phylogeny and biogeography in flowering plants: a global survey. *Perspectives in Plant Ecology, Evolution and Systematics*, *12*(1), 43-55.

Levin, D. A., Francisco-Ortega, J., and Jansen, R. K. (1996). Hybridization and the extinction of rare plant species. *Conservation Biology*, *10*(1), 10-16.

Levine, N., Ben-Zevi, G., Seifan, M., and Giladi, I. (2019). Investment in reward by ant-dispersed plants consistently selects for better partners along a geographic gradient. *AoB PLANTS, 11*(3), 1-13.

Lewis, J. S., Farnsworth, M. L., Burdett, C. L., Theobald, D. M., Gray, M., and Miller, R. S. (2017). Biotic and abiotic factors predicting the global distribution and population density of an invasive large mammal. *Scientific Reports*, *7*, 44152.

Liaw, A. and Wiener, M. (2002). Classification and Regression by randomForest. *R News, 2*(3), 18-22.

Lisci, M., Bianchini, M., and Pacini, E. (1996). Structure and function of the elaiosome in some angiosperm species. *Flora, 191*(2), 131-141.

Lomolino, M. V., Riddle, B. R. and Brown, J. H. (2005). *Biogeography*, 3rd edn. Sinauer Associates, Sunderland, MA.

Lubertazzi, D. (2011). The biology and natural history of *Aphaenogaster rudis. Psyche, 2011*, 1–11.

Lynch A. J., Gilfedder, L., and Kirkpatrick, J. B. (1999). The Tasmanian endemic shrub *Acacia axillaris:* conservation ecology applied to the question of rarity or vulnerability. *Australian Journal of Botany, 47*, 97–109.

MacArthur, R. H. (1984). Geographical ecology: patterns in the distribution of species. Princeton University Press.

MacGown, J. (2014). Ants (Formicidae) of the southeastern United States. Mississippi Entomological Museum, Mississippi State University, Mississippi State, MS, USA.http://mississippientomologicalmuseum.org.msstate.edu/Researchtaxapages/Formicidaepages/Myrmicinaekeys/Myrmicinekey.html#.WXUCJhMrKCS.

Manzaneda, A. J., and Rey, P. J. (2009). Assessing ecological specialization of an ant–seed dispersal mutualism through a wide geographic range. *Ecology, 90*(11), 3009-3022.

Manzaneda, A. J., Rey, P. J., and Alcántara, J. M. (2009). Conflicting selection on diaspore traits limits the evolutionary potential of seed dispersal by ants. *Journal of Evolutionary Biology, 22*(7), 1407-1417.

Marcer, A., Sáez, L., Molowny-Horas, R., Pons, X., and Pino, J. (2013). Using species distribution modelling to disentangle realised versus potential distributions for rare species conservation. *Biological Conservation*, *166*, 221-230.

Mark, S., and Olesen, J. (1995). Importance of elaiosome size to removal of ant-dispersed seeds. *Oecologia, 107*, 95–101.

Marshall, D. L., Beattie, A. J., and Bollenbacher, W. E. (1979). Evidence for diglycerides as attractants in an ant-seed interaction. *Journal of Chemical Ecology, 5*(3), 335-344.

Master, L.L., Morse, L.E., Weakley, A.S., Hammerson, G.A., Faber-Langendoen, D., 2003. NatureServe Conservation Status Criteria. NatureServe Arlington, VA, USA.

Mawdsley, N. A., Compton, S. G., and Whittaker, R. J. (1998). Population persistence, pollination mutualisms, and figs in fragmented tropical landscapes. *Conservation Biology*, *12*(6), 1416-1420.

Mayer, V. E., Frederickson, M. E., McKey, D., and Blatrix, R. (2014). Current issues in the evolutionary ecology of ant–plant symbioses. *New Phytologist, 202*(3), 749-764.

Meehan, T. (1870). Variations in *Trillium* and *Wisteria*. *The American Naturalist*, *4*(8), 472-473.

Millam, K. C. (2006). *The Trillium erectum species-complex (Melanthiaceae): Insights from molecular systematics and biogeography*. The University of Wisconsin-Madison.

Miller, R. I. (1986). Predicting rare plant distribution patterns in the southern Appalachians of the south-eastern USA. *Journal of Biogeography*, 293-311.

Miller, C. N., and Kwit, C. (2018). Overall seed dispersal effectiveness is lower in endemic *Trillium* species than in their widespread congeners. *American Journal of Botany, 105*(11), 1–11.

Miller, C. N., Whitehead, S. R., and Kwit, C. (2020). Effects of seed morphology and elaiosome chemical composition on attractiveness of five *Trillium* species to seed-dispersing ants. *Ecology and Evolution*, 00: 1– 14. https://doi.org/10.1002/ece3.6101.

Montalto, V., Rinaldi, A., and Sarà, G. (2015). Life history traits to predict biogeographic species distributions in bivalves. *The Science of Nature*, *102*(9-10), 61.

Morrone, J. (1994). On the identification of areas of endemism. *Systematic Biology, 43*: 438-441.

Morrone, O., Vega, A. S., and Maier, M. (2000). Elaiosomes in *Urochloa paucispicata* (Poaceae: Panicoideae: Paniceae): anatomy and chemical composition. *Flora, 195*(4), 303-310.

Munguía, M., Townsend Peterson, A., and Sánchez-Cordero, V. (2008). Dispersal limitation and geographical distributions of mammal species. *Journal of Biogeography*, *35*(10), 1879-1887.

Municio, A. A., Odriozola, J. M., and Pérez-Albarsanz, A. (1975). Biochemistry of development in insects. *European Journal of Biochemistry, 60,* 123–128.

Myers, N., Mittermeier, R. A., Mittermeier, C. G., da Fonseca, G. A. B., and Kent, J. (2000). Biodiversity hotspots for conservation priorities. *Nature, 403*: 853-858.

Myers, J. A., Vellend, M., Gardescu, S., and Marks, P. L. (2004). Seed dispersal by white-tailed deer: implications for long-distance dispersal, invasion, and migration of plants in eastern North America. *Oecologia, 139*(1), 35-44.

NatureServe. (2017). NatureServe Explorer: an online encyclopedia of life. NatureServe, Arlington, VA, USA. http://explorer.natureserve.org/.

Ness, J. H, Bronstein, J. L., Andersen, A. N., and Holland, J. N. (2004). Ant body size predicts dispersal distance of ant-adapted seeds: implications of small-ant invasions. *Ecology, 85,* 1244–1250.

Ness, J. H., Morin, D., and Giladi, I. (2009). Uncommon specialization in a mutualism between a temperate herbaceous plant guild and an ant: Are *Aphaenogaster* ants keystone mutualists? *Oikos, 118,* 1793–1804.

Nelson, A. S., Acosta, N. C., and Mooney, K. A. (2019). Plant chemical mediation of ant behavior. *Current Opinion in Insect Science.*

Nohara, T., Miyahara, K., and Kawasaki, T. (1975). Steroid saponins and sapogenins of underground parts of Trillium kamtschaticum Pall. II. Pennogenin-and kryptogenin 3-O-glycosides and related compounds. *Chemical and Pharmaceutical Bulletin, 23*(4), 872-885.

Oberrath R., and Böhning-Gaese, K. (2002). Phenological adaptation of ant-dispersed plants to seasonal variation in ant activity. *Ecology, 83,* 1412–1420.

O'Dowd, D. J., and Hay, M. E. (1980). Mutualism between harvester ants and a desert ephemeral: seed escape from rodents. *Ecology, 61*(3), 531-540.

Ohara, M., and Utech, F. H. (1986). Life history studies on the genus *Trillium* (Liliaceae) III. Reproductive biology of six sessile‐flowered species occurring in the southeastern United States with special reference to vegetative reproduction. *Plant Species Biology, 1,* 135-145.

Ohara, M., and Utech, F. H. (1988). Life history studies on the genus *Trillium* (Liliaceae) V. Reproductive biology and survivorship of three declinate‐flowered species. *Plant Species Biology, 3,* 35-41.

Ohara, M. (1989). Life history evolution in the genus *Trillium*. *Plant Species Biology, 4,* 1–28.

Ohkawara, K., Ohara, M., and Higashi, S. (1997). The evolution of ant‐dispersal in a spring‐ephemeral *Corydalis ambigua* (Papaveraceae): timing of seed‐fall and effects of ants and ground beetles. *Ecography, 20*(3), 217-223.

Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'hara, R. B.,... and Wagner, H. (2017). vegan: Community Ecology Package. R package version 2.3–2015.

Oostermeijer, J. G. B. (1989). Myrmecochory in *Polygala vulgaris* L., *Luzula campestris* (L.) DC. and *Viola curtisii* Forster in a Dutch dune area. *Oecologia, 78,* 302-311.

Osunkoya, O. O., and Swanborough, P. W. (2001). Reproductive and ecophysiological attributes of the rare *Gardenia actinocarpa* (Rubiaceae) compared with its common congener, *G. ovularis*. *Australian Journal of Botany, 49,* 471-478.

Paiva, E. A. S., Buono, R. A., and Lombardi, J. A. (2009). Food bodies in *Cissus verticillata* (Vitaceae): ontogenesis, structure and functional aspects. *Annals of Botany, 103*, 517–524.

Palmer, M. E. (1987). A critical look at rare plant monitoring in the United States. *Biological Conservation*, *39*(2), 113-127.

Papeş, M., Peterson, A. T., and Powell, G. V. (2012). Vegetation dynamics and avian seasonal migration: clues from remotely sensed vegetation indices and ecological niche modelling. *Journal of Biogeography*, *39*(4), 652-664.

Patrick, T. S. (1984). *Trillium sulcatum* (Liliaceae), a new species of the southern Appalachians. *Brittonia*, *36*(1), 26-36.

Peat, H. J. & Fitter, A. H. (1994). Comparative analyses of ecological characteristics of British angiosperms. *Biological Reviews 69:* 95–115.

Pellissier, L., Anne Bråthen, K., Pottier, J., Randin, C. F., Vittoz, P., Dubuis, A., ... and Guisan, A. (2010). Species distribution models reveal apparent competitive and facilitative effects of a dominant species on the distribution of tundra plants. *Ecography*, *33*(6), 1004-1014.

Peternelli, E. F. O., Della Lucia, T. M. C., Peternelli, L. A., Martins, S. V., and Borges, E. E. (2003). The interaction among workers of the genera *Atta* and *Acromyrmex* (Hymenoptera: Formicidae) and seeds of *Mabea fistulifera* (Euphorbiaceae) a pioneer tree species in Brazil. *Sociobiology, 42*, 597–603.

Peters, M., Oberrath, R., and Böhning-Gaese, K. (2003). Seed dispersal by ants: are seed preferences influenced by foraging strategies or historical constraints?. Flora-Morphology, Distribution, Functional Ecology of Plants, 198(6), 413-420.

Peterson, A. T. (2001). Predicting species' geographic distributions based on ecological niche modeling. *The Condor*, *103*(3), 599-605.

Peterson, A. T. (2006) Uses and requirements of ecological niche models and related distributional models. *Biodiversity Informatics, 3,* 59–72.

Pfeiffer, M., Huttenlocher, H., and Ayasse, M. (2010). Myrmecochorous plants use chemical mimicry to cheat seed-dispersing ants. Functional Ecology, 24, 545–555.

Phillips, S. J., Dudík, M., and Schapare, R. E. (2004). A maximum entropy approach to species distribution modeling. *Proceedings of the twenty-first international conference on Machine learning* (p. 83).

Phillips, S. J., Anderson, R. P., and Schapare, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, *190*(3-4), 231-259.

Phillips, S. J., and Dudík, M. (2008). Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography*, *31*(2), 161-175.

Pirie, C. D., Walmsley, S., Ingle, R., Pérez Jiménez, A., Solis Magallanes, A., and Kelly, C. K. (2000). Investigations in plant commonness and rarity: a comparison of seed removal patterns in the widespread *Jatropha standleyi* and the endemic *J. chamelensis* (Euphorbiaceae). *Biological Journal of the Linnean Society, 71*: 501-512.

Pollock, L. J., Tingley, R., Morris, W. K., Golding, N., O'Hara, R. B., Parris, K. M., ... and McCarthy, M. A. (2014). Understanding co-occurrence by modelling species simultaneously with a Joint Species Distribution Model (JSDM). *Methods in Ecology and Evolution*, *5*(5), 397-406.

Prior, K. M., Saxena, K., and Frederickson, M. E. (2014). Seed handling behaviors of native and invasive seed-dispersing ants differentially influence seedling emergence in an introduced plant. *Ecological Entomology, 39*, 66-74.

Prober, S. M., and Austin, M. P. (1993). Habitat peculiarity in the rare *Eucalyptus paliforrmis. Ecology, 82*, 189–205.

Pijl Van Der, L. (1982). Principles of dispersal in higher plants. Springer-Verlag, New York.

Qiu, H. L., Lu, L. H., Shi, Q. X., Tu, C. C., Lin, T., and He, Y. R. (2015). Differential necrophoric behavior of the ant *Solenopsis invicta* towards fungal-infected corpses of workers and pupae. *Bulletin of Entomological Research, 105(*5), 607-614.

R Core Team. (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rabinowitz, D. (1978). Abundance and diaspore weight in rare and common prairie grasses. *Oecologia, 37*, 213-219.

Rabinowitz, D. (1981). Seven forms of rarity. In Synge, H. [ed.], The biological aspects of rare plant conservation, 205–217. Wiley, Chichester, UK.

Rabinowitz, D., and Rapp, J. K. (1981). Dispersal abilities of seven sparse and common grasses from a Missouri prairie. *American Journal of Botany, 68*, 616-624.

Rebelo, H., and Jones, G. (2010). Ground validation of presence-only modelling with rare species: a case study on barbastelles *Barbastella barbastellus* (Chiroptera: Vespertilionidae). *Journal of Applied Ecology*, *47*(2), 410-420.

Regan, T. J., Master, L. L., & Hammerson, G. A. (2004). Capturing expert knowledge for

threatened species assessments: a case study using NatureServe conservation status ranks. *Acta Oecologica*, *26*(2), 95-107.

Reifenrath, K., Becker, C., and Poethke, H. J. (2012). Diaspore trait preferences of dispersing ants. *Journal of Chemical Ecology, 38*, 1093–1104.

Rickson F. (1976). Ultrastructural differentiation of Mullerian body glycogen plastid of *Cecropia peltata* L. *American Journal of Botany, 63,* 1272–1279.

Rico-Gray, V., and Oliveira, P. (2007). The ecology and evolution of ant–plant interactions. University of Chicago Press, Chicago, IL, USA.

Rios, N. E. and Bart, H. L. (2010). GEOLocate (Version 3.22) [Computer software]. Belle Chasse, LA: Tulane University Museum of Natural History. https://www.geo-locate.org/default.html.

Rodriguez-Cabal, M. A., Stuble, K. L., Guénard, B., Dunn, R. R., and Sanders, N. J. (2012). Disruption of ant–seed dispersal mutualisms by the invasive Asian needle ant (*Pachycondyla chinensis*). *Biological Invasions, 14*, 557-565.

Ruhren, S., and Dudash, M. R. (1996). Consequences of the timing of seed release of *Erythronium americanum* (Liliaceae), a deciduous forest myrmecochore. *American Journal of Botany, 83,* 633-640.

Salazar-Rojas, B., Rico-Gray, V., Canto, A., and Cuautle, M. (2012). Seed fate in the myrmecochorous Neotropical plant *Turnera ulmifolia* L., from plant to germination. *Acta Oecologica, 40*, 1-10.

Schliep, E. M., Lany, N. K., Zarnetske, P. L., Schaeffer, R. N., Orians, C. M., Orwig, D. A., and Preisser, E. L. (2018). Joint species distribution modelling for spatio-

temporal occurrence and ordinal abundance data. *Global Ecology and Biogeography*, *27*(1), 142-155.

Schmidt, K., and Jensen, K. (2000). Genetic structure and AFLP variation of remnant populations in the rare plant *Pedicularis palustris* (Scrophulariaceae) and its relation to population size and reproductive components. *American Journal of Botany*, *87*(5), 678-689.

Schultz, M. (2014). Effects of myrmecochore species abundance, diversity, and fruiting phenology on *Aphaenogaster* (Hymenoptera: Formicidae) nesting and foraging in southern Appalachian rich cove forests. Unpublished thesis.

Schaefer, H. M., Schmidt, V., and Winkler, H. (2003). Testing the defence trade-off hypothesis: how contents of nutrients and secondary compounds affect fruit removal. *Oikos, 102*(2), 318-328.

Schupp, E. W. (1993). Quantity, quality and the effectiveness of seed dispersal by animals. *Vegetatio, 107*(1), 15-29.

Schupp, E., Jordano, P., and Gómez, J. (2010). Seed dispersal effectiveness revisited: a conceptual review. *New Phytologist, 188*, 333–353.

SERNEC Data Portal. (2020). http//:sernecportal.org/portal/index.php. [Aug 1, 26, 30, and Sept. 1 - 4, 2018; July 12, 2019].

Shenoy, M., Radhika, V., Satish, S., and Borges, R. M. (2012). Composition of extrafloral nectar influences interactions between the myrmecophyte *Humboldtia brunonis* and its ant associates. *Journal of Chemical Ecology, 38*, 88–99.

Singh, P., Singh, G., Bhandawat, A., Singh, G., Parmar, R., Seth, R., and Sharma, R. K. (2017). Spatial transcriptome analysis provides insights of key gene (s) involved in steroidal saponin biosynthesis in medicinally important herb *Trillium govanianum. Scientific Reports*, *7*, 45295.

Skidmore, B. A., and Heithaus, E. R. (1988). Lipid cues for seed-carrying by ants in *Hepatica americana. Journal of Chemical Ecology, 14*(12), 2185-2196.

Soberon, J. and Peterson, A. T. (2005). Interpretation of models of fundamental ecological niches and species' distributional areas. *Biodiversity Informatics, 2,* 1–10.

Soltis, P. S., and Gitzendanner, M. A. (1999). Molecular systematics and the conservation of rare species. *Conservation Biology*, *13*(3), 471-483.

Soukup, V. G., and Holman, R. T. (1987). Fatty acids of seeds of North American pedicellate *Trillium* species. *Phytochemistry, 26*(4), 1015-1018.

Speed, J. D., and Austrheim, G. (2017). The importance of herbivore density and management as determinants of the distribution of rare plant species. *Biological Conservation*, *205*, 77-84.

Svenning, J. C., and Skov, F. (2004). Limited filling of the potential range in European tree species. *Ecology Letters, 7*(7), 565-573.

Syphard, A. D., and Franklin, J. (2010). Species traits affect the performance of species distribution models for plants in southern California. *Journal of Vegetation Science*, *21*(1), 177-189.

Takahashi, S., and Itino, T. (2015). Larger seeds are dispersed farther: the long distance seed disperser ant *Aphaenogaster famelica* prefers larger seeds. *Sociobiology, 59,* 1401–1411.

Tanner, E. P., Papeş, M., Elmore, R. D., Fuhlendorf, S. D., and Davis, C. A. (2017). Incorporating abundance information and guiding variable selection for climate-based ensemble forecasting of species' distributional shifts. *PLoS One*, 12, e0184316.

Tewksbury, J. J., and Nabhan, G. P. (2001). Seed dispersal: directed deterrence by capsaicin in chillies. *Nature, 412*(6845), 403.

Tewksbury, J. J., Reagan, K. M., Machnicki, N. J., Carlo, T. A., Haak, D. C., Peñaloza, L. C., and Levey, D. J. (2008). Evolutionary ecology of pungency in wild chilies. *Proceedings of the National Academy of Sciences, 105*(33), 11808-11811.

Thompson, S. N. (1973). A review and comparative characterization of the fatty acid composition of seven insect orders. *Comparative Biochemistry and Physiology, 45B*, 467–482.

Thompson, J. N. (1981). Elaiosomes and fleshy fruits: phenology and selection pressures for ant-dispersed seeds. *American Naturalist, 117*, 104-108.

Thorson, J. T., Ianelli, J. N., Munch, S. B., Ono, K., and Spencer, P. D. (2015). Spatial delay-difference models for estimating spatiotemporal variation in juvenile production and population abundance. *Canadian Journal of Fisheries and Aquatic Sciences*, *72*(12), 1897-1915.

Timmerman-Erskine, M., Dute, R. R., and Boyd, R. S. (2002). Morphometric analysis of the *Trillium pusillum* Michaux complex (Trilliaceae) of the southeastern United States. *Castanea*, 109-119.

Tomimatsu, H., and Ohara, M. (2003). Genetic diversity and local population structure of fragmented populations of *Trillium camschatcense* (Trilliaceae). *Biological Conservation*, *109*(2), 249-258.

Traveset, A., and Richardson, D. M. (2014). Mutualistic interactions and biological invasions. *Annual Review of Ecology, Evolution, and Systematics*, *45*, 89-113.

Tropicos.org. (2020). Missouri Botanical Garden. http://www.tropicos.org [March 28, 2019].

Turner, K., and Frederickson, M. (2013). Signals can trump rewards in attracting seed-dispersing ants. *PLoS ONE, 8,* 1–9.

Tylianakis, J. M., Didham, R. K., Bascompte, J., and Wardle, D. A. (2008). Global change and species interactions in terrestrial ecosystems. *Ecology Letters*, *11*(12), 1351-1363.

Umphrey, G. J. (1996). Morphometric discrimination among sibling species in the *fulva–rudis–texana* complex of the ant genus *Aphaenogaster* (Hymenoptera: Formicidae). *Canadian Journal of Zoology, 74*(3), 528-559.

UT Herbarium – TENN. (2020). The University of Tennessee, Knoxville. https://herbarium.utk.edu/. [Aug. 10 – 15, 2018].

Van der Putten, W. H., Macel, M., and Visser, M. E. (2010). Predicting species distribution and abundance responses to climate change: why it is essential to include biotic interactions across trophic levels. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*(1549), 2025-2034.

VanDerWal, J., Shoo, L. P., Johnson, C. N., & Williams, S. E. (2009). Abundance and the environmental niche: environmental suitability estimated from niche models predicts the upper limit of local abundance. *The American Naturalist*, *174*(2), 282-291.

Vellend, M., Myers, J. A., Gardescu, S., and Marks, P. L. (2003). Dispersal of *Trillium* seeds by deer: implications for long-distance migration of forest herbs. *Ecology, 84,* 1067-1072.

Venables, W. N., and Ripley, B. D. (2002). Modern applied statistics with S, 4th ed. Springer, NY, NY, USA.

Ver Hoef, J. M., and Boveng, P. L. (2007). Quasi-poisson vs. negative binomial regression: how should I model overdispersed count data? *Ecology, 88*, 2766–2772.

von Mueller, F. (1855). Definitions of Rare or Hitherto Undescribed Australian Plants: Chiefly Collected Within the Boundaries of the Colony of Victoria (Vol. 1, No. 1-2). Goodhugh and Tembath.

Walck, J., Baskin, J., and Baskin, C. (1999). Ecology of the endangered species *Solidago shortii*. VI. Effects of habitat type, leaf litter, and soil type on seed germination. *Journal of the Torrey Botanical Society,* 117-123.

Wang, T., Hamann, A. Spittlehouse, D. L., and Carroll, C. (2016). Locally downscaled and spatially customizable climate data for historical and future periods for North America. *PLoS One* 11: e0156720.

Warren, R. J., Bahn, V., and Bradford, M. A. (2011). Temperature cues phenological synchrony in ant-mediated seed dispersal. *Global Change Biology, 17*(7), 2444-2454.

Warren, R. J., and Bradford, M. A. (2014). Mutualism fails when climate response differs between interacting species. *Global Change Biology*, *20*(2), 466-474.

Warren, R. J., Giladi, I., and Bradford, M. A. (2014). Competition as a mechanism structuring mutualisms. *Journal of Ecology, 102*, 486-495.

Warren, R. J., Chick, L. D., DeMarco, B., McMillan, A., De Stefano, V., Gibson, R., and Pinzone, P. (2016). Climate-driven range shift prompts species replacement. *Insectes Sociaux, 63*(4), 593-601.

Washitani, I., Takenaka, A., Kuramoto, N., and Inoue, K. (1997). Aster *Kantoensis kitam*., an endangered flood plain endemic plant in Japan: its ability to form persistent soil seed banks. *Biological Conservation, 82*, 67-72.

Webber, B. L., Abaloz, B. A., and Woodrow, I. E. (2007). Myrmecophilic food body production in the understory tree, *Ryparosa kurrangii* (Achariaceae), a rare Australian rainforest taxon. *New Phytologist, 173*, 250–263.

Wenig, P., and Odermatt, J. (2010). OpenChrom: a cross-platform open source software for the mass spectrometric analysis of chromatographic data. *BMC Bioinformatics, 11*(1), 405.

Whitehead, S. R., Jeffrey, C. S., Leonard, M. D., Dodson, C. D., Dyer, L. A., and Bowers, M. D. (2013). Patterns of secondary metabolite allocation to fruits and seeds in Piper reticulatum. *Journal of Chemical Ecology*, *39*(11-12), 1373-1384.

Wetzel, W. C., and Whitehead, S. R. (2020). The many dimensions of phytochemical diversity: linking theory to practice. *Ecology Letters*, *23*(1), 16-32.

Williams, J. N., Seo, C., Thorne, J., Nelson, J. K., Erwin, S., O'Brien, J. M., and Schwartz, M. W. (2009). Using species distribution models to predict new occurrences for rare plants. *Diversity and Distributions*, *15*(4), 565-576.

Williams, S. E., Williams, Y. M., VanDerWal, J., Isaac, J. L., Shoo, L. P., & Johnson, C. N. (2009). Ecological specialization and population size in a biodiversity hotspot: how rare species avoid extinction. *Proceedings of the National Academy of Sciences*, *106* (Supplement 2), 19737-19741.

Winfree, R., Aguilar, R., Vázquez, D. P., LeBuhn, G., and Aizen, M. A. (2009). A meta-analysis of bees' responses to anthropogenic disturbance. *Ecology*, *90*(8), 2068-2076.

Wisz, M. S., Hijmans, R. J., Li, J., Peterson, A. T., Graham, C. H., Guisan, A., & NCEAS Predicting Species Distributions Working Group. (2008). Effects of sample size on the performance of species distribution models. *Diversity and Distributions*, *14*(5), 763-773.

Wogan, G. O. (2016). Life history traits and niche instability impact accuracy and temporal transferability for historically calibrated distribution models of North American birds. *PLoS One*, *11*(3).

Wu, X. B., and Smeins, F. E. (2000). Multiple-scale habitat modeling approach for rare plant conservation. *Landscape and Urban Planning*, *51*(1), 11-28.

Wu, J., Peng, L., Dong, S., Xia, X., and Zhao, L. (2019). Transcriptome analysis of *Chelidonium majus* elaiosomes and seeds provide insights into fatty acid biosynthesis. PeerJ, 7, e6871.

Zelikova, T. J., Dunn, R., and Sanders, N. J. (2008). Variation in seed dispersal along an elevational gradient in Great Smoky Mountains National Park. *Acta Oecologica, 34*, 155-162.

Zelikova, T. J., Sanders, N. J., and Dunn, R. (2011). The mixed effects of experimental ant removal on seedling distribution, belowground invertebrates, and soil nutrients. *Ecosphere, 2,* 1–14.

Zettler, J. A., Spira, T. P., and Allen, C. R. (2001). Yellow jackets (*Vespula* spp.) disperse *Trillium* (spp.) seeds in eastern North America. *The American Midland Naturalist*, *146*(2), 444-446.

Zhou, H., Chen, J., and Chen, F. (2007). Ant-mediated seed dispersal contributes to the local spatial pattern and genetic structure of *Globba lancangensis* (Zingiberaceae). *Journal of Heredity*, *98*(4), 317-324.

APPENDICES

APPENDIX I: CHAPTER 2

Methods

Liquid chromatography-mass spectrometry

Using sterile techniques, I removed elaiosomes from frozen diaspores of each species using a straight razor. I recorded the fresh mass of single elaiosomes (g) and placed each elaiosome in a 2.0 mL centrifuge tube. Sample contents were suspended in 1.3 mL of extraction solvent (40 : 40 : 20 HPLC grade methanol, acetonitrile, water with 0.1% formic acid), and were kept at 4°C. Extraction proceeded for 20 min at -20°C before samples were centrifuged for 5 min (16.1 relative centrifugal force [rcf]) at 4°C. Supernatants were transferred to new vials. The remaining elaiosome contents were resuspended in 200 µL of cold (4°C) extraction solvent. The extraction again proceeded for 20 min at -20°C before being centrifuged for 5 min (16.1 rcf) at 4°C. Once more, I transferred supernatants to the vials, and added another 200 µL of extraction solvent to the pelleted elaiosomes for a final wash by repeating the extraction once more. The vials containing all of the combined extraction supernatants were then placed in a nitrogen drying apparatus until all the extraction solvent had been evaporated. I resuspended the residue in 300 µL of sterile water to isolate the polar fraction of phytochemicals and transferred this to 300 µL autosampler vials. Samples were immediately placed in a 4°C autosampler for LC-MS analysis.

A 10 µL injection of each sample was separated through a Synergi 2.5 micron Hydro-RP 100 Å, 100 mm × 2.00 mm LC column (Phenomenex, Torrance, CA, USA) maintained at 25°C. The mass spectrometer and chromatographic separation were

performed similar to the method outlined in Lu et al., 2010. The eluent was introduced into the mass spectrometer via an electrospray ionization source in negative mode before entering an Exactive Plus orbitrap mass spectrometer (Thermo Scientific, Waltham, MA, USA) through a 0.1-mm internal diameter fused silica capillary tube. Samples were run with a spray voltage of 3 kV, a nitrogen sheath gas flow rate of 10 units, a capillary temperature set at 320°C, and an AGC target set to 3e6. Samples were analyzed in full scan mode with a resolution of 140,000 and a scan window of 85 to 800 m/z for from 0 to 9 min and 110 to 1000 m/z from 9 to 25 min. Solvent A consisted of 97:3 HPLC grade water : methanol, 10 mM tributylamine, and 15 mM acetic acid. Solvent B was HPLC grade methanol. The mobile phase gradient from 0 to 5 min was 0% B, from 5 to 13 min was 20% B, from 13 to 15.5 min was 55% B, from 15.5 to 19 min is 95% B, and from 19 to 25 min was 0% B while maintaining a constant flow rate of 200 µL/min.

Raw files obtained from Xcalibur MS software (Thermo Electron Corp., Waltham, MA) were converted into the mzML format using ProteoWizard (Chambers et al., 2012). The converted files were imported into MAVEN (Metabolomic Analysis and Visualization Engine for LC–MS Data), a software package (Clasquin et al. 2012). Peaks for the known metabolites were picked in MAVEN, which automatically performs non-linear retention time correction and calculates peak areas across samples, using a preliminary mass error of ± 20 ppm and retention time window of 5 min. The University of Tennessee, Knoxville Biological and Small Molecule Mass Spectrometry Core (BSMMSC), through which this analysis took place, has replicated and expanded the method of Rabinowitz and coworkers (Lu et al., 2010). Final metabolite annotations were made using a library of 263 retention time-accurate m/z pairs taken from MS1 spectra.

The annotation parameters have been verified previously with pure standards in the course of establishing the method. For a metabolite to be annotated as a known compound, the eluted peak had to be found within 2 min of the expected retention time, and the metabolite mass had to be within ± 5 ppm of the expected value. Metabolite identities were confirmed using the MAVEN software package (Clasquin et al., 2012), and peak areas for each compound were integrated using the Quan Browser function of the Xcalibur MS Software.

Gas-chromatography mass spectrometry

I removed elaiosomes from frozen seeds as described above and placed them in 15-mL glass scintillation vials. Twelve samples were generated, representing the five study species, each from at least two field sites. Each sample contained a minimum of 30 mg of elaiosomes from one individual plant. The number of elaiosomes used for each sample depended on the species but ranged from 14 to 35. To each sample, I added 2.0 mL of the extraction solvent (2 : 1 dichloromethane : methanol). Samples were centrifuged for 30 m (170 rpm), and supernatants were siphoned off and transferred to new glass scintillation vials. I repeated this treatment three times, until 6 mL of supernatants were generated for each sample. The vials containing all of the combined extraction supernatants were placed in a nitrogen drying apparatus until all the extraction solvent had been evaporated. The dry residue in each vial was suspended in 1 mL isooctane : ethyl acetate (10 : 1) and was then applied to an alumina column (4-cm Pasteur pipette filled with flash alumina pre-equilibrated with isooctane : ethyl acetate 10 : 1), as in Boulay et al. (2006). The column was eluted with 4 ml isooctane : ethyl acetate (10 : 1) to yield triglyceride fractions for each sample. Elution of the diglyceride fraction

was realized with 5 mL of isooctane : ethyl acetate (3 : 1). The free fatty acids were captured after elution with 6 mL of isooctane : ethyl acetate : acetic acid (75 : 25 : 2). Fractions were collected in standard glass test tubes.

Resulting fractions of triglycerides, diglycerides, and free fatty acids were dried under nitrogen for 3 h. Residue was dissolved in 1 mL of the extraction solvent, and then 400 µL of this volume was transferred to a 2.0 mL vial containing 20 µg of octadecane in methanol as an internal standard. Solutions were dried again under nitrogen and subsequent steps taken to convert fatty acids to their methyl esters for GC-MS analysis. Transesterification for tri- and diglyceride fractions was accomplished by adding 100 µL of methanolic KOH (0.5 M) to the dry triglyceride and diglyceride fractions. After 30 m, the reaction was stopped with 100 µL of HCl (1 N). Then 20 µL of hexane was added to the mixture. Methylation of the free fatty acid fraction was carried out by treating the solution with 200 µL of toluene, 1.5 mL of methanol, and 300 µL of an 8% (w/v) solution of HCL in methanol/water (85 : 15, w/v), sequentially, to the fraction, as in Ichihara and Fukubayashi (2009). The fraction was then heated to 100 C for 1 h. All the obtained fractions were dried one final time and dissolved in 100 µL of isooctane, of which 1 µL was injected into an Agilent 7820 gas chromatograph coupled with a 5977 mass spectrometer set to 70eV electron ionization and equipped with an HP5-MS column (30 m x 0.25 mm i.d., 0.25 µm film thickness; Agilent Technologies, Santa Clara, CA, USA). Ultra-pure helium was used as a carrier gas at a flow rate of 1 mL min -1, a split flow ratio of 100:1, and a front inlet temperature of 275 °C. The following oven conditions were employed: initial temperature 120 °C, hold time 2 min; ramp 1: 15°C

min^{-1} to 250 °C, hold time 0 min; ramp 2: 5 °C min^{-1} to 300°C, hold time 10 min; total run time of 30.67 min. Data were recorded as TIC chromatograms.

Comparisons of total fatty acids, and fatty acids in three forms (free, di-, and triglycerides)

To enable linear comparisons of concentrations of fatty acids present in total, free form, and di- and triglycerides, I built simple linear models with the averaged probability of seed dispersal for each species as the predictor and percent fresh weight concentrations of total fatty acids, summed free fatty acids, and summed fatty acids in diglycerides and triglycerides as the respective responses. I generated best fit lines for each regression to illustrate the direction of trends.

Results

Full RFCM and RFRM

The full (known + unknown) phytochemical composition of elaiosomes were distinguishable among species of *Trillium* based on the RFCM. The out-of-bag error rate was 0.73, meaning the model only had 27% sample classification accuracy. Boruta variable selection identified 18 of the possible 7,552 variables as important classifiers. A MANOVA showed a significant overall difference among species for these 18 compounds ($F_{36,80}$=8.28, P<0.001). Follow-up ANOVAs revealed that 15 of the 19 compounds were significantly different among the five species (Appendix I: Table 10). The full phytochemical composition of elaiosomes was distinguishable among averaged seed dispersal probabilities based on the RFRM, although the model only explained 0.77% of the variation in the data. The full RFRM identified 14 of the possible 7,552

variables as important in distinguishing among averaged dispersal probabilities. Follow-up linear models revealed that 10 of the 14 variables were significantly different among the species. Of the 14 compounds identified by the full RFRM, 5 compounds were also identified by the full RFCM (methylhistidine, creatinol O-phosphate, citrulline, d-glucarate, and indole 3-carboxylate; Appendix I: Table 10).

Comparisons of total fatty acids, and fatty acids in three forms (free, di-, and triglycerides)

Concentrations of oleic, linoleic, hexadecenoic, stearic, and total fatty acids differed significantly among the three fatty acid forms ($X^2(2)=15.43$, $P<0.001$; $X^2(2)=29.34$, $P<0.001$; $X^2(2)=30.93$, $P<0.001$; $X^2(2)=211.57$, $P<0.001$; $X^2(2)=19.51$, $P<0.001$, respectively). Post hoc tests revealed that concentrations of oleic, linoleic, hexadecenoic, and total fatty acids were significantly greater in triglyceride forms than in diglyceride forms, with free fatty acid forms intermediate; concentrations of stearic acid were significantly higher in free form than either di- or triglycerides (Appendix I: Table 9). The concentration of palmitoleic acid was not significantly different among the three forms ($X^2(2)=5.26$, $P=0.07$). Concentrations of summed total fatty acids and summed free fatty acids were significantly positively related to seed dispersal probability (t=2.9, $P=0.05$; t=2.8, $P=0.02$, respectively). Concentrations of summed fatty acids in diglycerides and triglycerides were not significantly related to seed dispersal probability, although both trends are positive ($P > 0.05$).

Concentrations of all fatty acids except palmitoleic acid were significantly different in free, diglyceride, and triglyceride forms, and post hoc tests revealed that concentrations tended to be lowest in diglycerides and highest in free fatty acids and

triglycerides. Linear trends shoId that only summed total fatty acids and summed free fatty acids were significantly, positively related to the probability of seed dispersal. It is interesting that only summed free fatty acids, not summed triglycerides, are significantly related to dispersal probability, given that specific fatty acids were present in high concentrations in triglycerides. Of the three forms, free fatty acids may be most important to ants in this study system, corroborating the findings of Boulay, et al. (2006) and Pfeiffer, et al. (2010), but disagreeing with Marshall, et al. (1979) and Gammans, et al. (2006).

Tables

Table 9: Tukey post hoc comparisons following linear mixed-effects models for diaspore morphological metrics and elaiosome fatty acids that were found to be significantly different among the species of *Trillium.* Significant *P*-values are in bold.

Post Hoc comparisons	Linear Hypothesis			Est.	S.E.	Z-value	P-value
Diaspore mass	cuneatum	-	catesbaei	-0.0071225	0.001843	-3.865	**<0.001**
	decumbens	-	catesbaei	-0.0116465	0.0020864	-5.582	**<0.001**
	discolor	-	catesbaei	-0.0014638	0.0009875	-1.482	0.69
	lancifolium	-	catesbaei	-0.0054543	0.0093268	-0.585	1
	decumbens	-	cuneatum	-0.004524	0.0010062	-4.496	**<0.001**
	discolor	-	cuneatum	0.0056587	0.0015719	3.6	**0.002**
	lancifolium	-	cuneatum	0.0016682	0.0092622	0.18	1
	discolor	-	decumbens	0.0101827	0.0018536	5.493	**<0.002**
	lancifolium	-	decumbens	0.0061922	0.009286	0.667	1
	lancifolium	-	discolor	-0.0039905	0.0092948	-0.429	1
Diaspore width	cuneatum	-	catesbaei	-0.3591	0.1825	-1.967	0.34
	decumbens	-	catesbaei	-0.8886	0.206	-4.314	**<0.001**
	discolor	-	catesbaei	-0.1336	0.1015	-1.316	0.94
	lancifolium	-	catesbaei	-0.4643	0.5788	-0.802	1
	decumbens	-	cuneatum	-0.5295	0.1032	-5.132	**<0.001**
	discolor	-	cuneatum	0.2255	0.1562	1.444	0.89
	lancifolium	-	cuneatum	-0.1053	0.5688	-0.185	1
	discolor	-	decumbens	0.755	0.1836	4.111	**<0.001**
	lancifolium	-	decumbens	0.4242	0.5726	0.741	1
	lancifolium	-	discolor	-0.3308	0.5739	-0.576	1
Diaspore length	cuneatum	-	catesbaei	-0.04863	0.16532	-0.294	1
	decumbens	-	catesbaei	-0.40048	0.18681	-2.144	0.29
	discolor	-	catesbaei	-0.19048	0.09057	-2.103	0.29
	lancifolium	-	catesbaei	-0.37139	0.59833	-0.621	1
	decumbens	-	cuneatum	-0.35185	0.09218	-3.817	**0.001**
	discolor	-	cuneatum	-0.14186	0.14126	-1.004	1
	lancifolium	-	cuneatum	-0.32276	0.59032	-0.547	1
	discolor	-	decumbens	0.20999	0.16631	1.263	1
	lancifolium	-	decumbens	0.02909	0.59331	0.049	1
	lancifolium	-	discolor	-0.1809	0.59436	-0.304	1

(Table 9 continued.)

Elaiosome Length	cuneatum	-	catesbaei	-0.78961	0.2178	-3.625	**0.002**
	decumbens	-	catesbaei	-0.90123	0.2462	-3.661	**0.002**
	discolor	-	catesbaei	-0.51484	0.11878	-4.334	**<0.001**
	lancifolium	-	catesbaei	-0.69509	0.82936	-0.838	1
	decumbens	-	cuneatum	-0.11162	0.12092	-0.923	1
	discolor	-	cuneatum	0.27477	0.18603	1.477	0.84
	lancifolium	-	cuneatum	0.09452	0.8193	0.115	1
	discolor	-	decumbens	0.38639	0.21909	1.764	0.54
	lancifolium	-	decumbens	0.20614	0.82305	0.25	1
	lancifolium	-	discolor	-0.18025	0.82438	-0.219	1
Total FAs	cuneatum	-	catesbaei	0.006376	0.555193	0.011	1
	decumbens	-	catesbaei	-0.131835	0.574143	-0.23	1
	discolor	-	catesbaei	-0.593352	0.216414	-2.742	0.05
	lancifolium	-	catesbaei	-1.189319	0.578966	-2.054	0.28
	decumbens	-	cuneatum	-0.138211	0.194939	-0.709	1
	discolor	-	cuneatum	-0.599728	0.555193	-1.08	1
	lancifolium	-	cuneatum	-1.195695	0.260869	-4.584	**<0.001**
	discolor	-	decumbens	-0.461517	0.574143	-0.804	1
	lancifolium	-	decumbens	-1.057484	0.319823	-3.306	**0.009**
	lancifolium	-	discolor	-0.595967	0.578966	-1.029	1
Hexadecanoic acid	cuneatum	-	catesbaei	-0.09618	0.10993	-0.875	1
	decumbens	-	catesbaei	-0.1689	0.11524	-1.466	0.72
	discolor	-	catesbaei	-0.21142	0.05224	-4.047	**<0.001**
	lancifolium	-	catesbaei	-0.34022	0.11641	-2.923	**0.03**
	decumbens	-	cuneatum	-0.07272	0.04675	-1.555	0.72
	discolor	-	cuneatum	-0.11524	0.10993	-1.048	1
	lancifolium	-	cuneatum	-0.24405	0.06101	-4	**<0.001**
	discolor	-	decumbens	-0.04252	0.11524	-0.369	1
	lancifolium	-	decumbens	-0.17132	0.07481	-2.29	0.15
	lancifolium	-	discolor	-0.1288	0.11641	-1.106	1

(Table 9 continued.)

Oleic acid	cuneatum	-	catesbaei	-0.10517	0.3148	-0.334	1
	decumbens	-	catesbaei	-0.18314	0.32868	-0.557	1
	discolor	-	catesbaei	-0.40261	0.14187	-2.838	0.05
	lancifolium	-	catesbaei	-0.62874	0.33189	-1.894	0.41
	decumbens	-	cuneatum	-0.07797	0.12722	-0.613	1
	discolor	-	cuneatum	-0.29744	0.3148	-0.945	1
	lancifolium	-	cuneatum	-0.52356	0.16728	-3.13	0.02
	discolor	-	decumbens	-0.21948	0.32868	-0.668	1
	lancifolium	-	decumbens	-0.4456	0.20511	-2.172	0.24
	lancifolium	-	discolor	-0.22612	0.33189	-0.681	1
Stearic acid	cuneatum	-	catesbaei	0.1092	0.6234	0.175	1
	decumbens	-	catesbaei	-0.1722	0.6386	-0.27	1
	discolor	-	catesbaei	-0.7965	0.202	-3.943	**<0.001**
	lancifolium	-	catesbaei	-0.9746	0.6428	-1.516	0.86
	decumbens	-	cuneatum	-0.2814	0.1827	-1.54	0.86
	discolor	-	cuneatum	-0.9058	0.6234	-1.453	0.86
	lancifolium	-	cuneatum	-1.0838	0.2486	-4.36	**<0.001**
	discolor	-	decumbens	-0.6243	0.6386	-0.978	1
	lancifolium	-	decumbens	-0.8024	0.3047	-2.634	0.07
	lancifolium	-	discolor	-0.1781	0.6428	-0.277	1
Linoleic acid	cuneatum	-	catesbaei	-0.01398	0.38576	-0.036	1
	decumbens	-	catesbaei	-0.18767	0.39847	-0.471	1
	discolor	-	catesbaei	-0.45184	0.14748	-3.064	**0.02**
	lancifolium	-	catesbaei	-0.89245	0.40175	-2.221	0.18
	decumbens	-	cuneatum	-0.17369	0.13291	-1.307	1
	discolor	-	cuneatum	-0.43786	0.38576	-1.135	1
	lancifolium	-	cuneatum	-0.87847	0.17822	-4.929	**<0.001**
	discolor	-	decumbens	-0.26417	0.39847	-0.663	1
	lancifolium	-	decumbens	-0.70478	0.21849	-3.226	**0.01**
	lancifolium	-	discolor	-0.44061	0.40175	-1.097	1

(Table 9 continued.)

Palmitoleic acid	cuneatum	-	catesbaei	-0.91227	0.24477	-3.727	**0.002**
	decumbens	-	catesbaei	-1.10379	0.27825	-3.967	**<0.001**
	discolor	-	catesbaei	-1.02766	0.2796	-3.675	**<0.001**
	lancifolium	-	catesbaei	-1.13922	0.27829	-4.094	**<0.001**
	decumbens	-	cuneatum	-0.19152	0.22374	-0.856	1
	discolor	-	cuneatum	-0.11539	0.24477	-0.471	1
	lancifolium	-	cuneatum	-0.22695	0.22669	-1.001	1
	discolor	-	decumbens	0.07613	0.27825	0.274	1
	lancifolium	-	decumbens	-0.03543	0.26479	-0.134	1
	lancifolium	-	discolor	-0.11156	0.27829	-0.401	1

Table 10: Follow-up ANOVAs to RFCMs and follow-up linear models to RFRMs. Single asterisks indicate compounds selected by both the partial RFCM and RFRM or both the full RFCM and RFRM; double asterisks indicate compounds selected by all four RF models. Significant *P*-values are in bold.

ANOVAs (RFCMs)

Model	Compounds	Sum Sq.	Df	F-value	P-value
Partial RFCM	lactate	4.53E+14	4	7.00	**<0.001**
	thymine	1.13E+10	4	6.55	**<0.001**
	pyroglutamic acid	3.93E+14	4	9.44	**<0.001**
	citraconate	7.21E+12	4	5.60	**0.002**
	adenine	1.50E+15	4	9.93	**<0.001**
	salicylate*	8.77E+11	4	6.98	**<0.001**
	xanthine*	1353239135	4	8.61	**<0.001**
	hydroxyphenylacetate	1.55E+14	4	13.34	**<0.001**
	histidine*	6.41E+13	4	6.36	**0.001**
	indole 3-carboxylate	8615674602	4	11.78	**<0.001**
	3, 4-dihydroxyphenylacetate*	1.52E+11	4	13.10	**<0.001**
	methylhistidine**	7.92E+10	4	9.63	**<0.001**
	aconitate	2.28E+14	4	5.41	**0.003**
	citrulline**	5.69E+12	4	7.93	**<0.001**
	glucosamine*	9.07E+11	4	10.15	**<0.001**
	hydroxyphenylpyruvate	2.34E+12	4	527.87	**<0.001**
	homocysteic acid	1.41E+11	4	17.44	**<0.001**
	citrate isocitrate	2.79E+16	4	4.68	**0.006**
	creatinol o-phosphate**	9.31E+09	4	2.46	0.07
	d-glucarate**	1.57E+13	4	18.23	**<0.001**
	pantothenate*	3.08E+14	4	4.38	**0.008**
	inosine	2.75E+12	4	1.85	0.15
	IMP	9.11E+11	4	1.15	0.35

(Table 10 continued.)

Model	Compounds	Sum Sq.	Df	F-value	P-value
Full RFCM	lactate	4.53E+14	4	7.00	**<0.001**
	pyroglutamic acid	3.93E+14	4	9.44	**<0.001**
	hydroxy 2-methylsuccinate	6.42E+14	4	3.96	**0.01**
	hydroxyphenylacetate	1.55E+14	4	13.34	**<0.001**
	histidine	6.41E+13	4	6.36	**0.001**
	indole 3-carboxylate*	8615674602	4	11.78	**<0.001**
	methylhistidine**	7.92E+10	4	9.63	**<0.001**
	citrulline**	5.69E+12	4	7.93	**<0.001**
	hydroxyphenylpyruvate	2.34E+12	4	527.87	**<0.001**
	homocysteic acid	1.41E+11	4	17.44	**<0.001**
	creatinol o-phosphate**	9.31E+09	4	2.46	0.07
	d-glucarate**	1.57E+13	4	18.23	**<0.001**
	pantothenate	3.08E+14	4	4.38	**0.008**
	X846	2.42E+13	4	3.42	**0.02**
	X1417	3.74E+14	4	3.50	**0.02**
	X3731	1.10E+14	4	7.60	**<0.001**
	X4861	3.38E+12	4	2.57	0.06
	X5318	9.44E+09	4	5.40	0.003

(Table 10 continued).

Linear Models (RFRMs)

Model	Compounds	Sum Sq.	Df	F-value	P-value
Partial RFRM	salicylate *	-2.10E+06	7E+05	-3.15	**<0.001**
	guanine	3.02E+05	1E+05	2.78	**0.009**
	xanthine *	-96641	22624	-4.27	**<0.001**
	histidine *	-4.43E+06	7E+06	-0.66	0.52
	uric acid	-1.57E+06	1E+06	-1.52	0.14
	3, 4-dihydroxyphenylacetate *	4.58E+05	3E+05	1.69	0.1
	methylhistidine **	6.06E+05	2E+05	3.28	**0.003**
	citrulline **	-5.75E+06	2E+06	-3.64	**0.001**
	glucosamine *	-2.36E+06	6E+05	-4.10	**<0.001**
	creatinol o-phosphate **	-1.58E+05	1E+05	-1.51	0.14
	d-glucarate **	8.45E+06	2E+06	3.74	**<0.001**
	pantothenate *	-3.65E+07	1E+07	-2.45	**0.02**
Full RFRM	salicylate	-6.23E+05	6E+05	-1.04	0.31
	xanthine	-96641	22624	-4.27	**<0.001**
	indole 3-carboxylate *	-5.75E+06	2E+06	-3.64	**0.001**
	citrulline **	-5.75E+06	2E+06	-3.64	**0.001**
	d-glucarate **	8.45E+06	2E+06	3.74	**<0.001**
	methylhistidine **	3.65E+07	1E+07	2.45	**0.02**
	glucosamine	-2.36E+06	6E+05	-4.10	**<0.001**
	creatinol o-phosphate **	-1.87E+07	2E+07	-1.15	0.26
	Unknown X1216	-2.17E+06	5E+05	-3.95	**<0.001**
	Unknown X2570	4.63E+04	1E+06	0.05	0.96
	Unknown X4619	-2.24E+06	8E+05	-2.78	**0.01**
	Unknown X4870	2.30E+05	2E+05	1.32	0.2
	Unknown X4906	-2.18E+05	70176	-3.10	**0.004**
	Unknown X6186	3.70E+06	2E+06	2.28	**0.03**

Figures

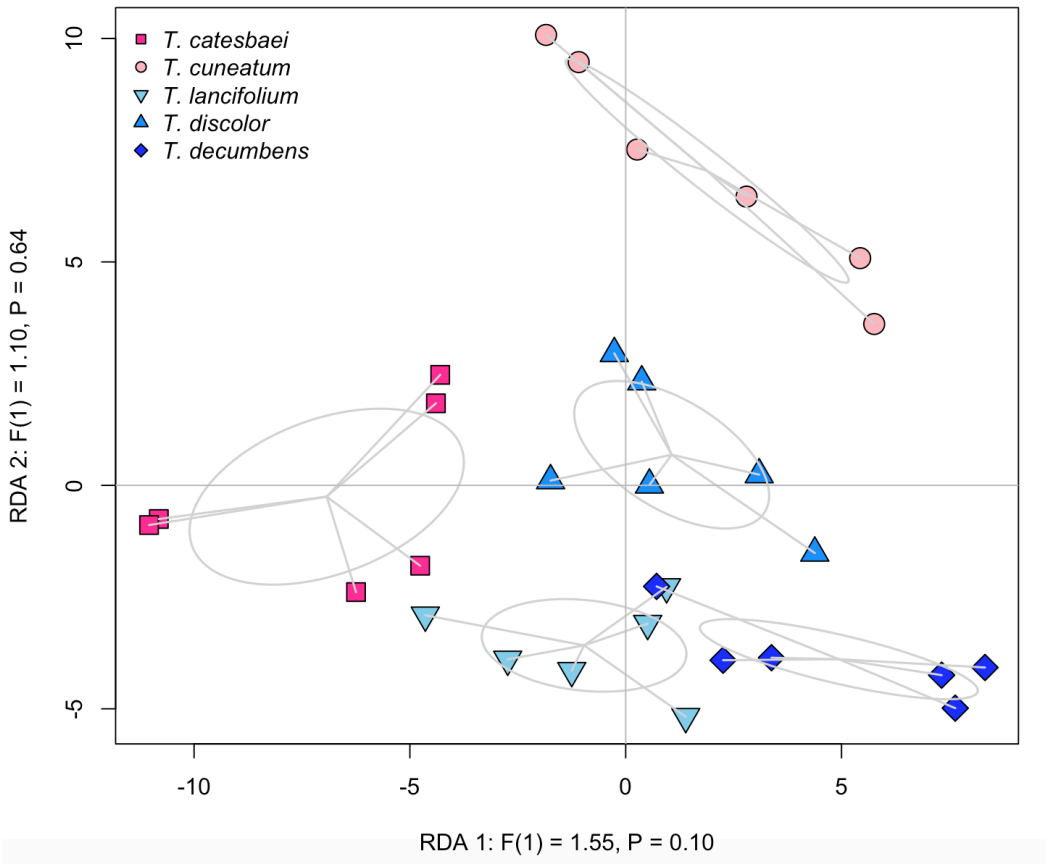


Figure 15: Constrained RDA of general elaiosome phytochemical profiles (n = 6 samples per species; n = 30), with relative abundance of 122 known phytochemical compounds and 7,431 unknown features as the response variables and species as the fixed effect (n = 7,552). Ellipses represent the standard deviation of the points around the centroid for each species. Preferred species are indicated by pink, and less-preferred are indicated by blue. The overall model is significant ($P = 0.03$), and explains 15.52% of the variation in the data. Species is a significant predictor of the total phytochemical composition of elaiosomes, including both known and unknown constituents.

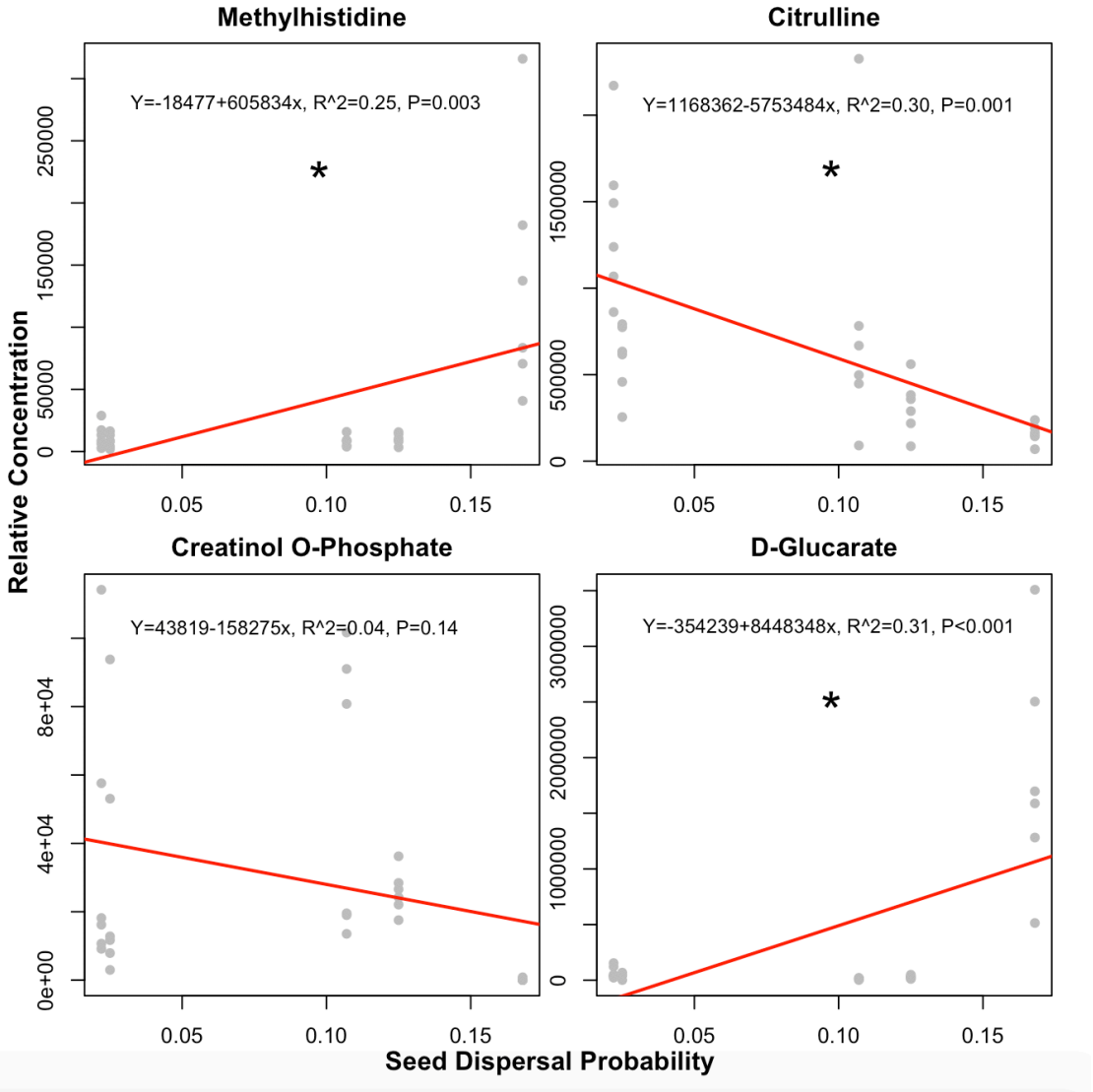


Figure 16: Linear relationships between relative concentrations of four elaiosome phytochemicals selected by all random forest models as important in distinguishing among species of *Trillium,* and their corresponding averaged seed dispersal probabilities (0.022 - *T. decumbens,* 0.025 - *T. discolor,* 0.107 - *T. lancifolium*, 0.125 - *T. cuneatum,* 0.168 - *T. catesbaei)*. Equation of the line of best fit is included at the top of each plot. Asterisks indicate significant relationships.

APPENDIX II: CHAPTER 3

Tables

Table 11: Multicollinearity diagnostics for reproductive life history traits data matrix containing biomass, ovule number, seed mass, and flower type as predictors. Detection: "1" indicates collinearity was detected by the test. Detection: "0" indicates collinearity was not detected by the test. No collinearity was detected among the predictor variables.

	MC Results	Detection
Determinant |X'X|	0.478	0
Farrar Chi-Square	11.6884	0
Red Indicator	0.3296	0
Sum of Lambda Inverse	5.7623	0
Theil's Method	-0.2367	0
Condition Number	10.7622	0

Table 12: Results of GLMs testing the effects of four reproductive life history traits on two model fit statistics, standardized testing AUC and omission error. LR Chisq contains the likelihood ratio chi-square test statistics (Wald tests). d.f. contains the degrees of freedom. *P* contains the p-values. Flower type and ovule number were significant predictors of standardized testing AUC. None of the traits were significant predictors of omission error.

Response: Standardized testing AUC	**LR Chisq**	**d.f.**	***P***
Flower type	38.464	1	<0.001*
Biomass (g)	1.710	1	0.1910
No. ovules per flower	23.132	1	<0.001*
Seed mass (mg)	0.334	1	0.5631
Response: Omission Error	**LR Chisq**	**d.f.**	***P***
Flower type	2.08912	1	0.14835
Biomass (g)	0.27158	1	0.60227
No. ovules per flower	2.92973	1	0.08696
Seed mass (mg)	0.15438	1	0.69438

Figures

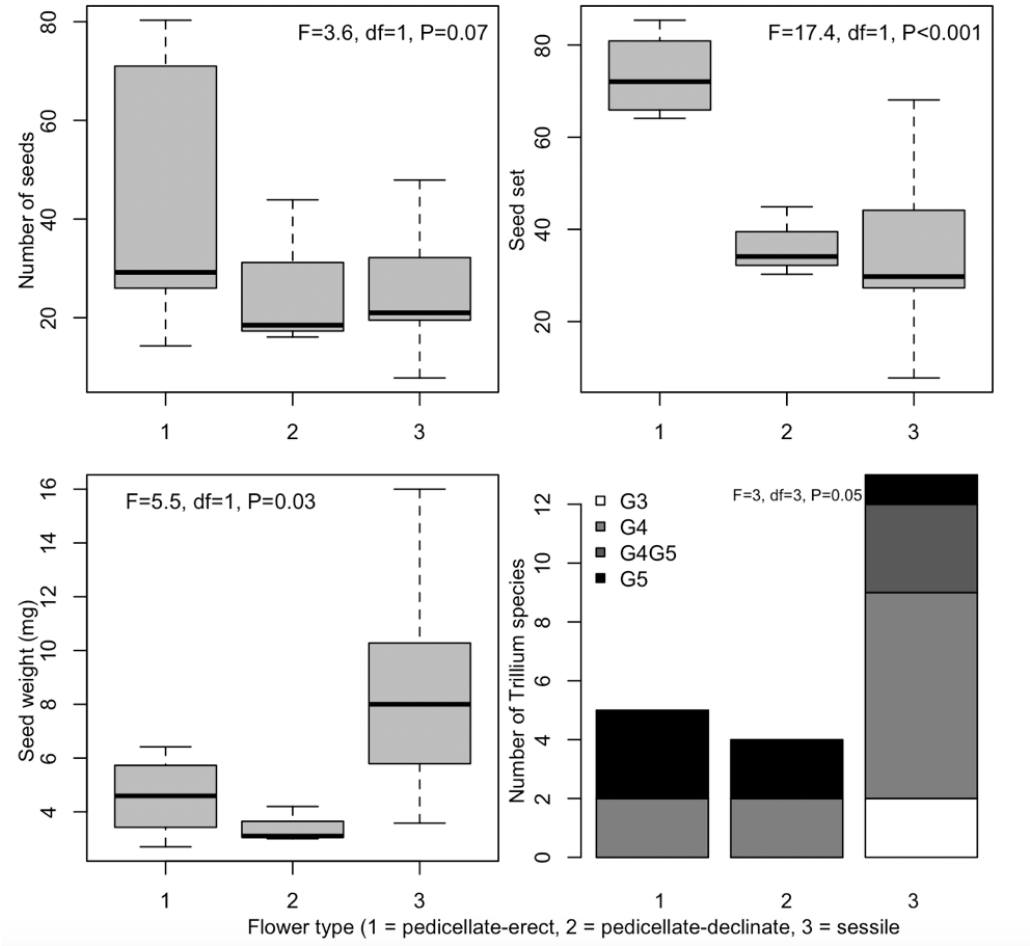


Figure 17: Boxplots illustrating median, interquartile range, and outliers for number of seeds (upper left), seed set (upper right), seed weight (lower left), and number of *Trillium* species grouped as G3, G4, G4G5, or G5 NatureServe conservation status (lower right) by flower type (1 = pedicellate-erect, 2 = pedicellate-declinate, 3=sessile). ANOVA test statistics are included at the top of each plot. Pedicellate-declinate flower types have significantly higher seed set than pedicellate-declinate and sessile species. Sessile species produce significantly heavier seeds. The sessile group is represented by species from all four NatureServe conservation ranks; this is significantly different than the pedicellate-flowered groups, which only contain species with G4 or G5 rankings. While the number of seeds produced is not significantly different among the flower types, pedicellate-erect species produce a much wider range of numbers than either pedicellate-declinate or sessile species.

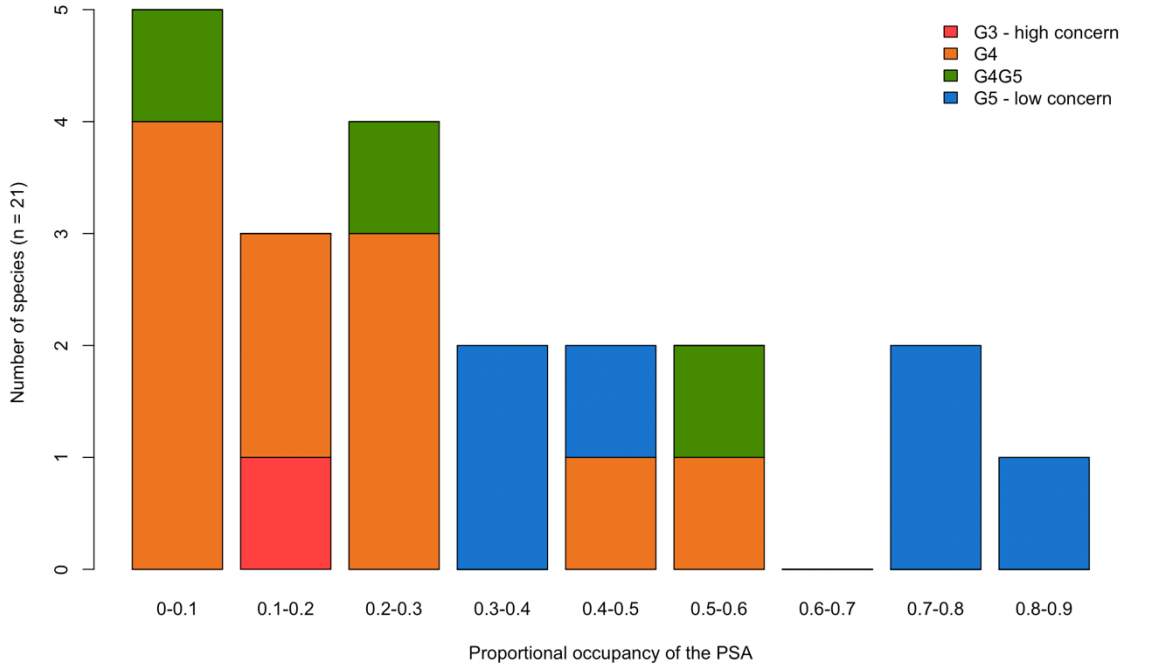


Figure 18: Frequency histogram illustrating the proportional occupancy of the PSA for 21 species of *Trillium*. Slightly more than half of the species (12) have known ranges that intersect with 30% or less of their PSA. Six species have known ranges that intersect with between 40 and 60% of their PSA, and three species have known ranges that intersect with more than 60% of their PSA.

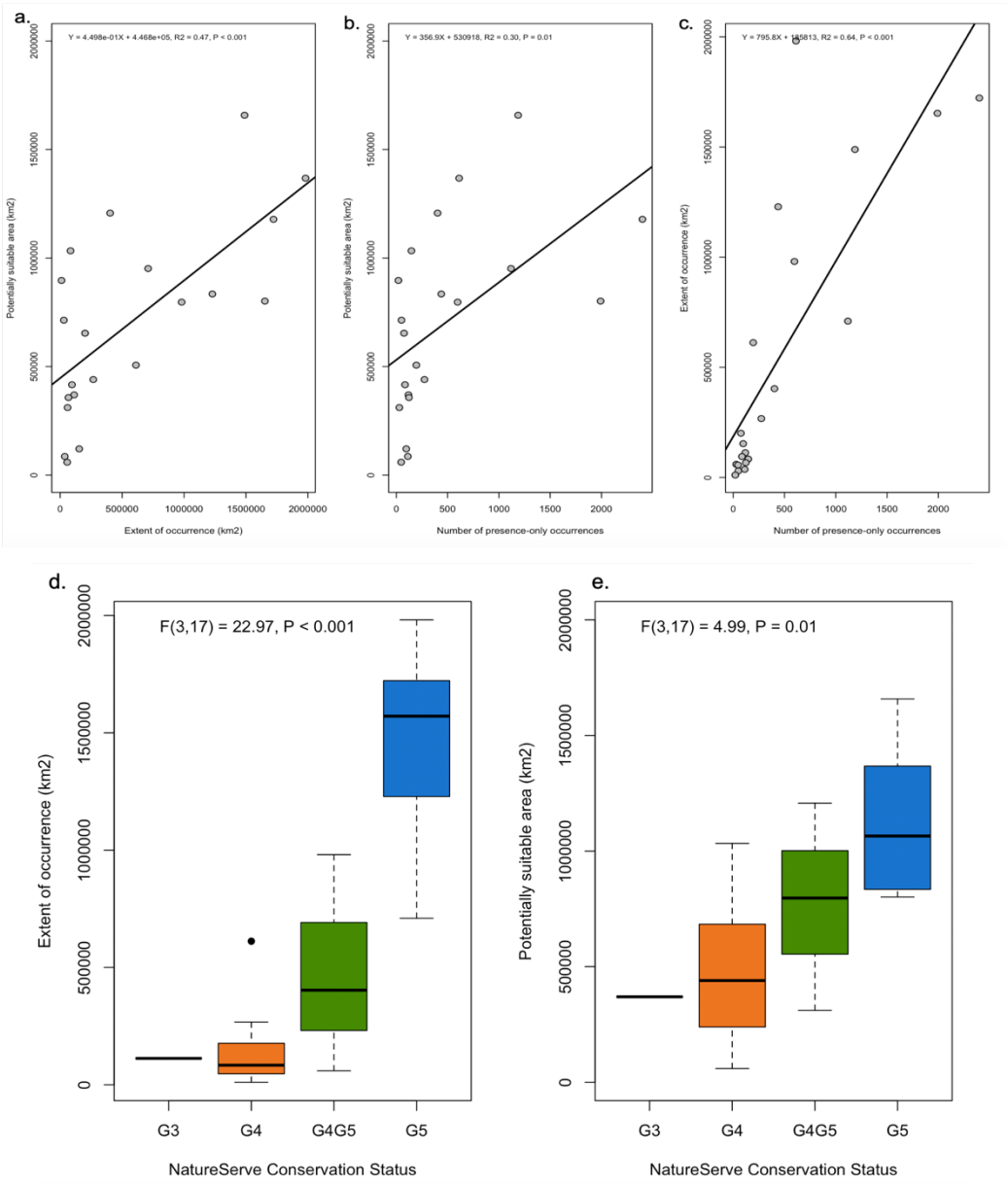


Figure 19: Scatterplots illustrating the significant, positive relationships between a. EOO (km^2) and PSA (km^2), b. the number of presence-only occurrences and PSA (km^2), and c. the number of presence-only occurrences and the EOO (km^2) for 21 species of *Trillium.* The equation of the line, R^2 values, and *P*-values are included at the top left of each plot. Boxplots illustrating the significant relationships between d. EOO and NatureServe conservation status, and e. PSA and NatureServe conservation status. F-statistics, degrees of freedom, and p-values are included at the top left of each plot.

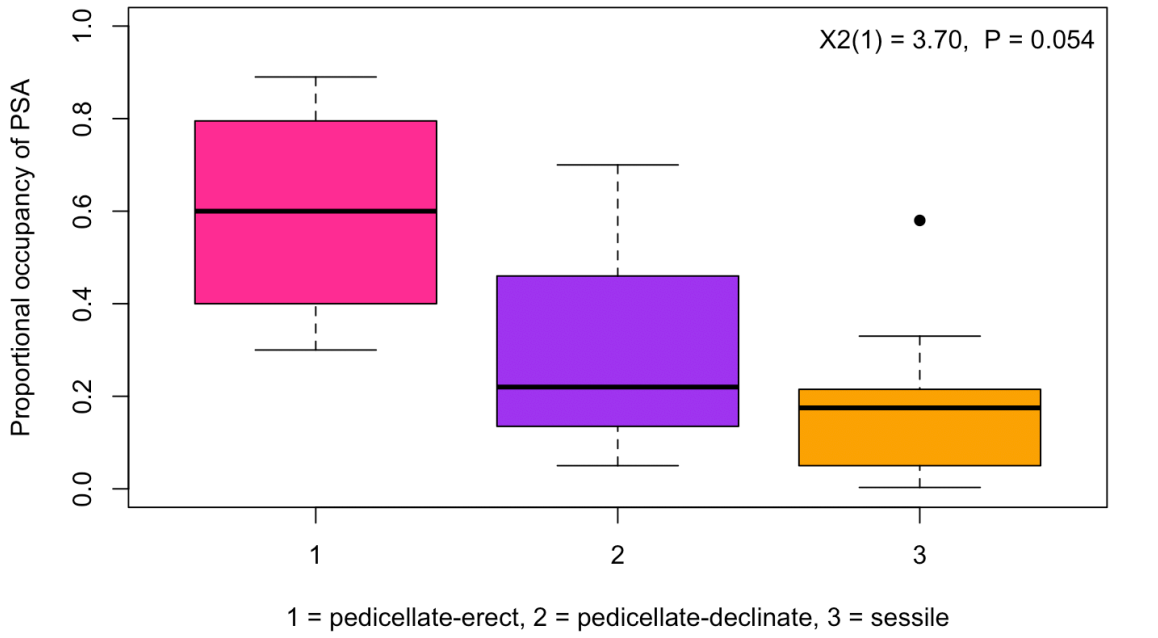


Figure 20: Boxplot illustrating median, interquartile range, and outliers for the proportional occupancy of the PSA for *Trillium* species belonging to one of three flower types: pedicellate-erect ("1"), pedicellate-declinate ("2"), and sessile ("3"). Although flower type was only marginally significant, species belonging to the sessile group tended to have lower proportional occupancy of the PSA than species belonging to either pedicellate group.

(1.)

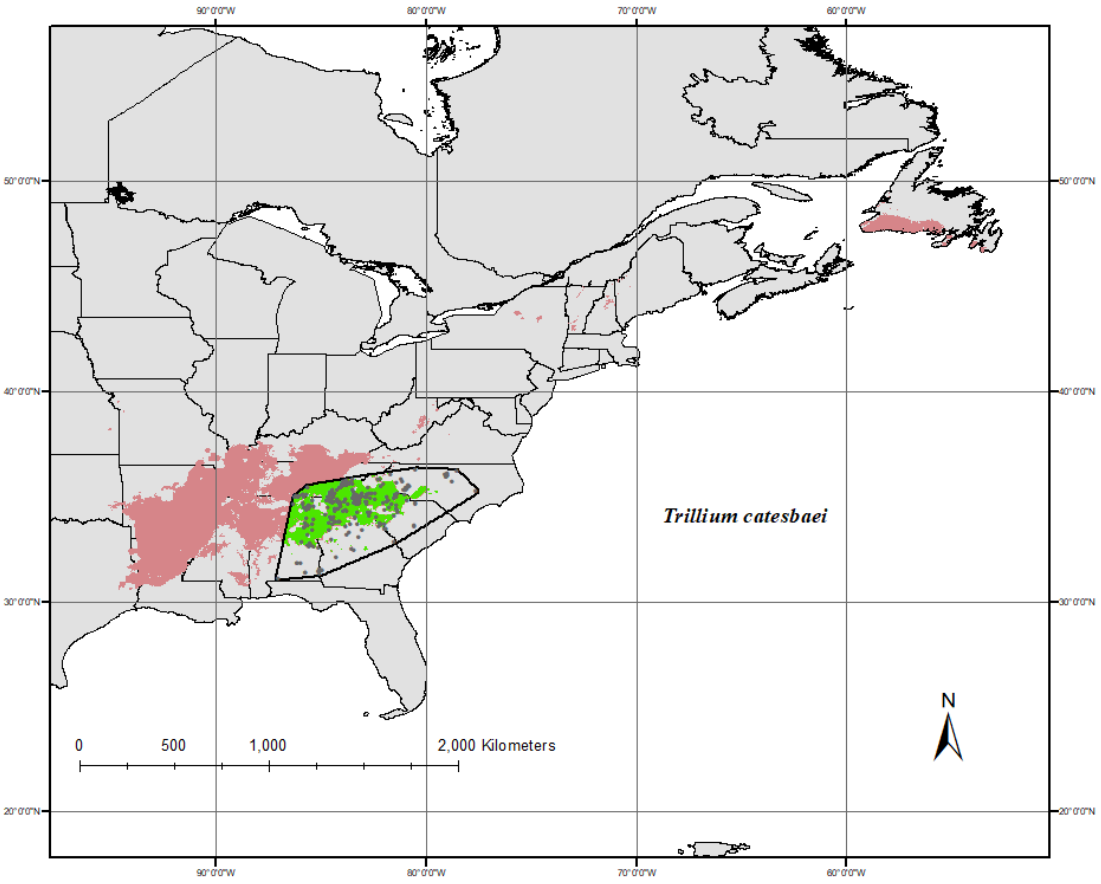
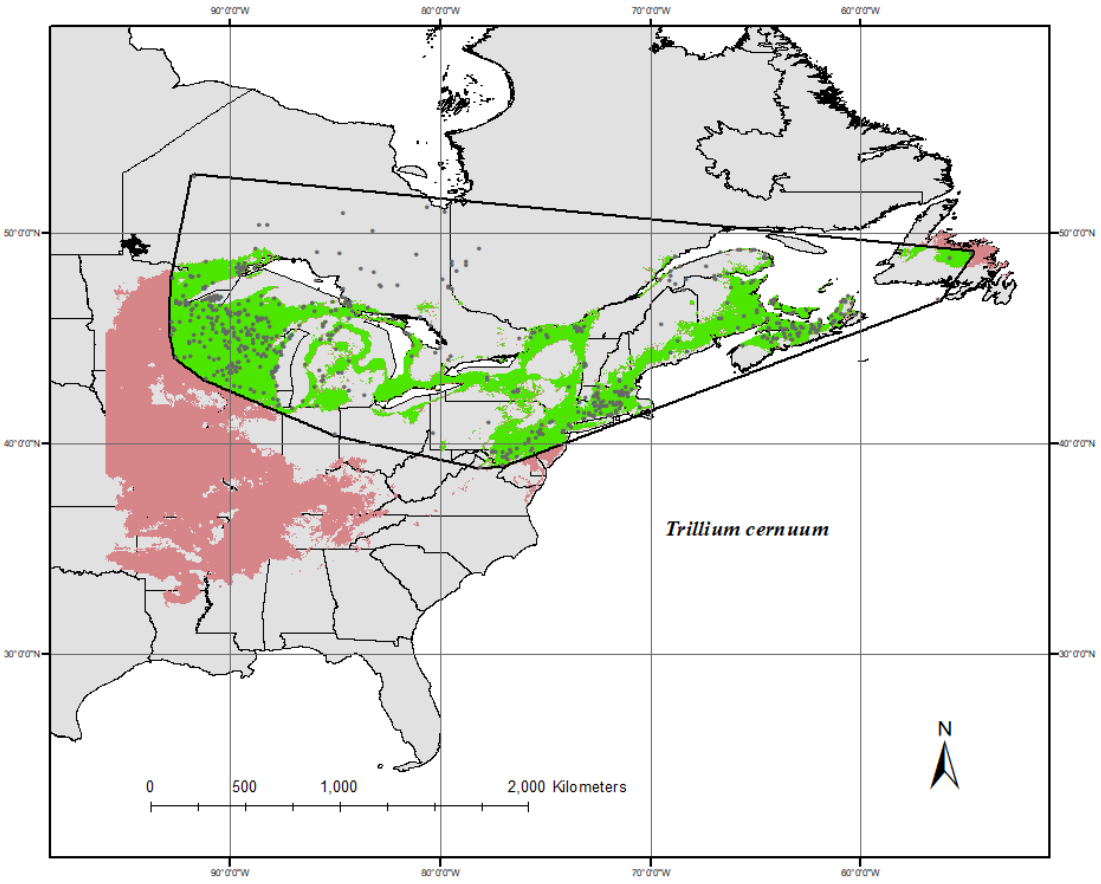


Figure 21: Maps of the predicted suitable area (PSA), depicted in pink, estimated by Maxent SDMs for (1.) *Trillium catesbaei,* (2.) *T. cernuum,* (3.) *T. cuneatum,* (4.) *T. decipiens,* (5.) *T. decumbens,* (6.) *T. discolor,* (7.) *T. erectum,* (8.) *T. flexipes,* (9.) *T. foetidissimum,* (10.) *T. grandiflorum,* (11.) *T. ludovicianum,* (12.) *T. lutem,* (13.) *T. maculatum,* (14). *T. nivale,* (15.) *T. recurvatum,* (16.) *T. sessile,* (17.) *T. stamenium,* (18.) *T. sulcatum,* (19.) *T. underwoodii,* (20.) *T. undulatum,* and (21.) *T. vaseyi.* Convex hulls depicting the extent of occurrence (EOO) for each species are overlaid. Areas in green represent the overlap between the PSA and the EOO (i.e. the proportional occupancy of the PSA).

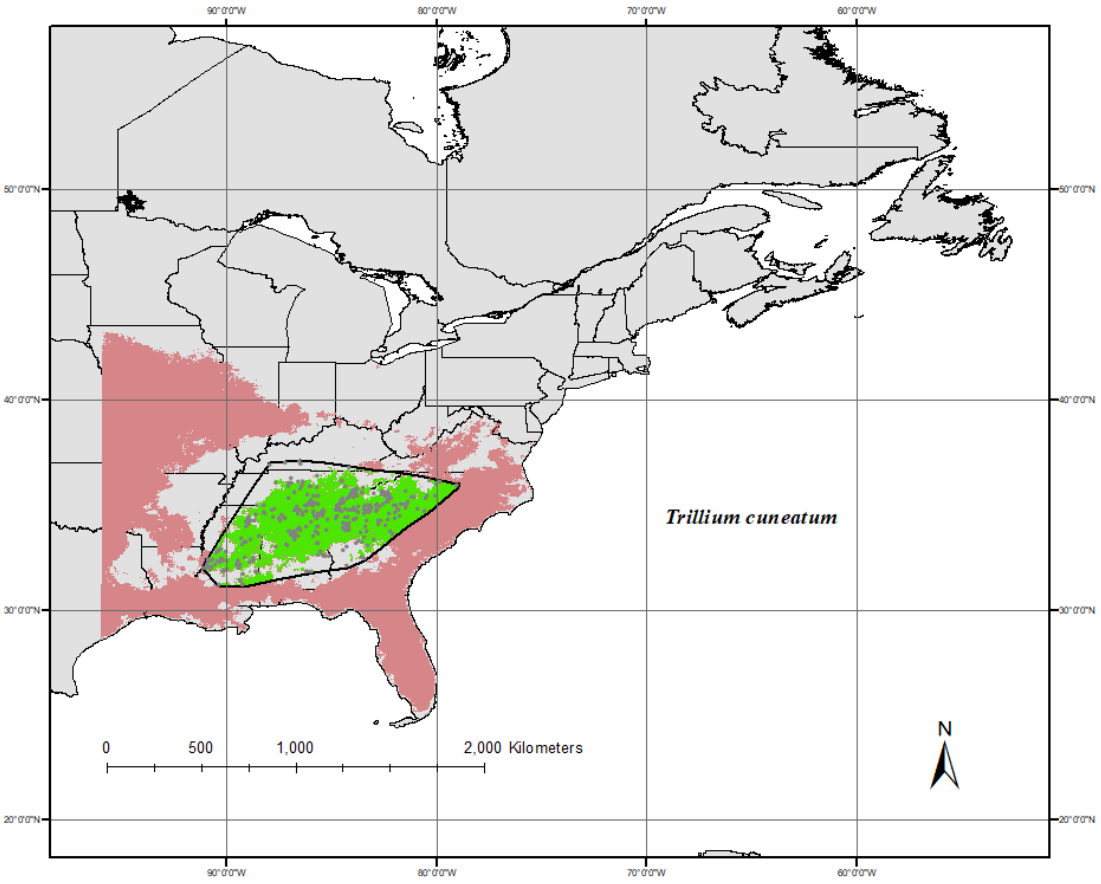
(Figure 21 continued.)

(2.)



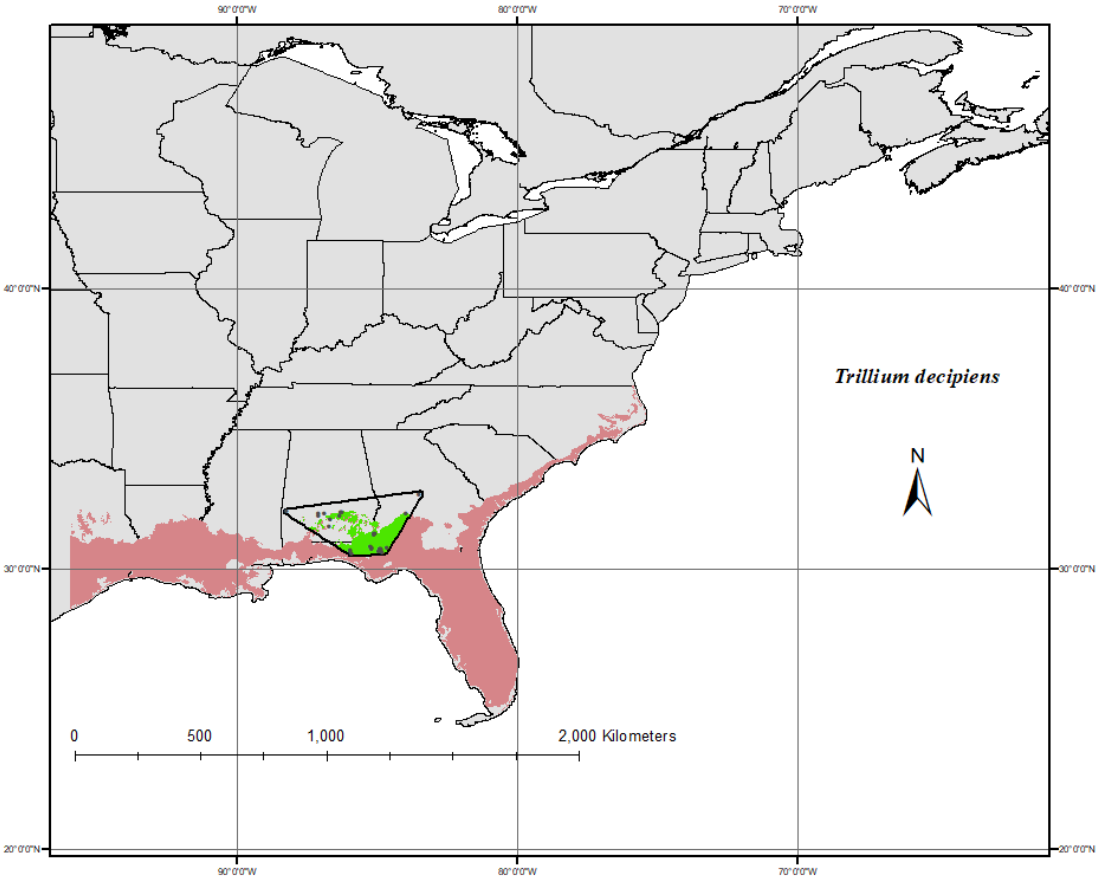
(Figure 21 continued.)

(3.)



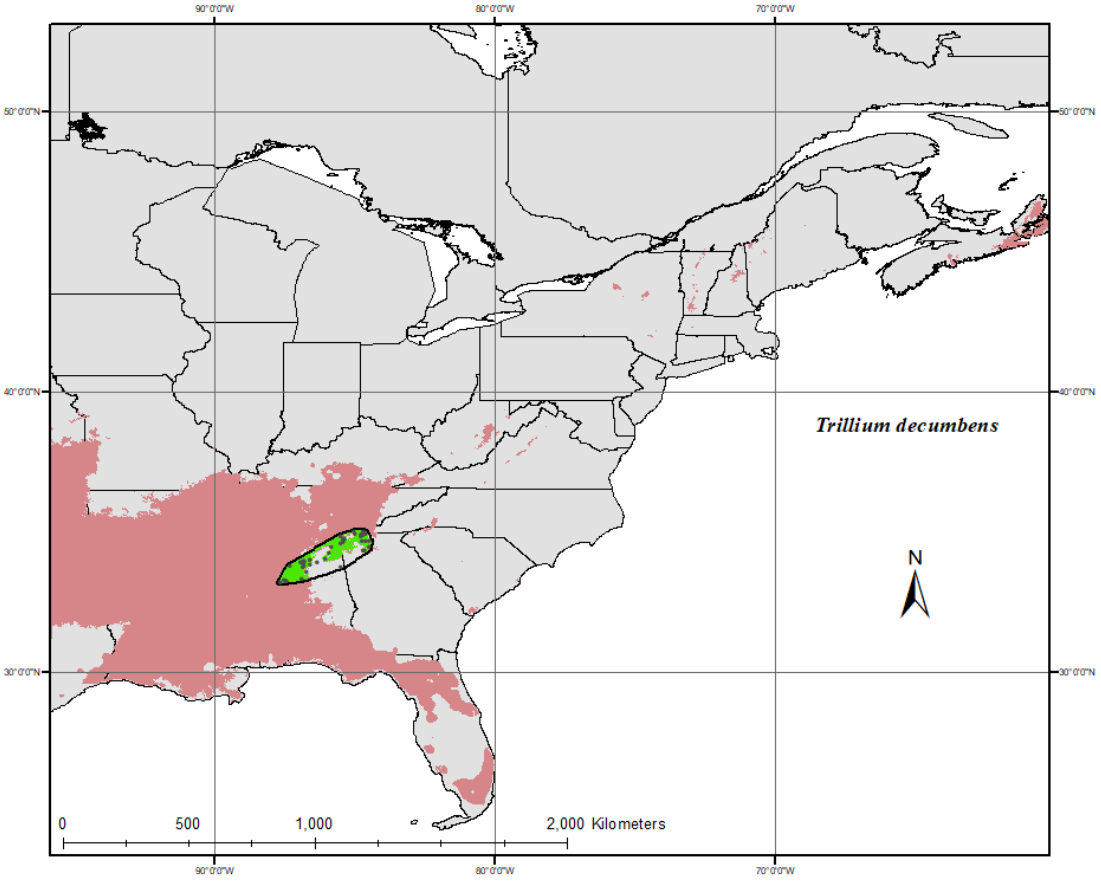
(Figure 21 continued.)

(4.)



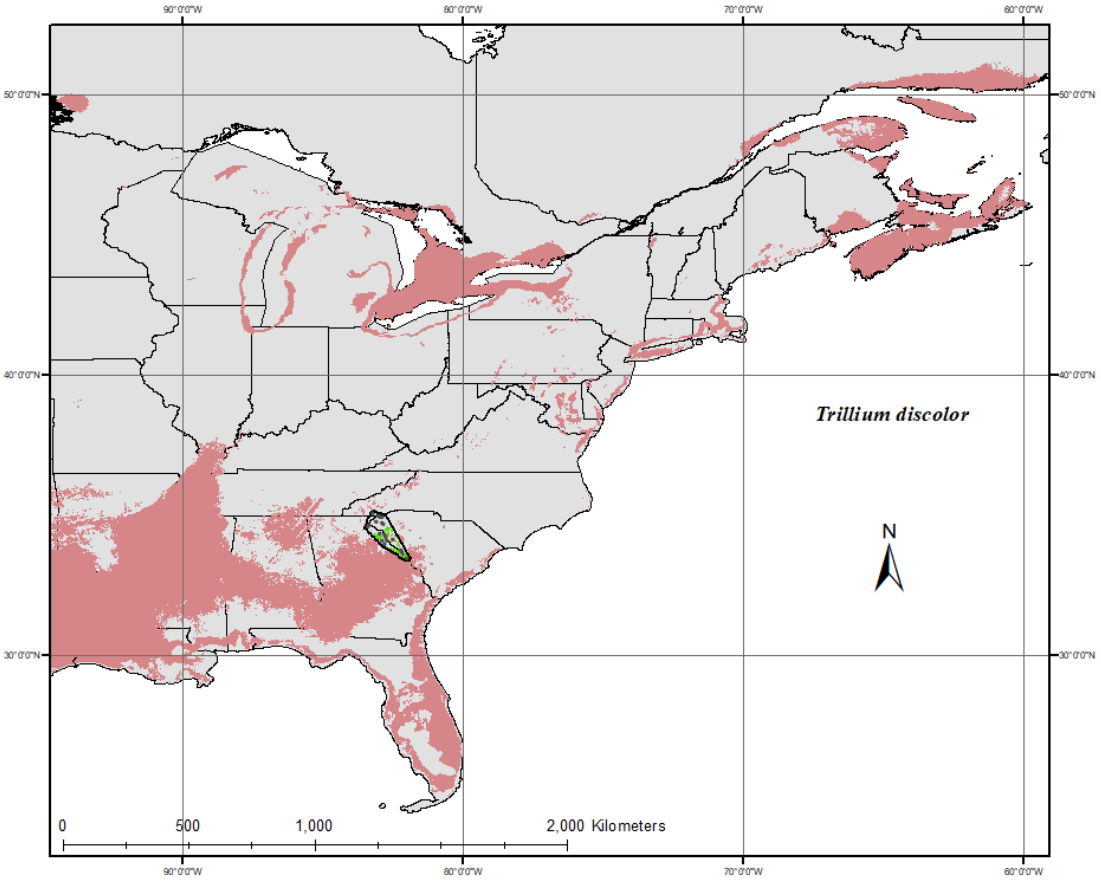
(Figure 21 continued.)

(5.)



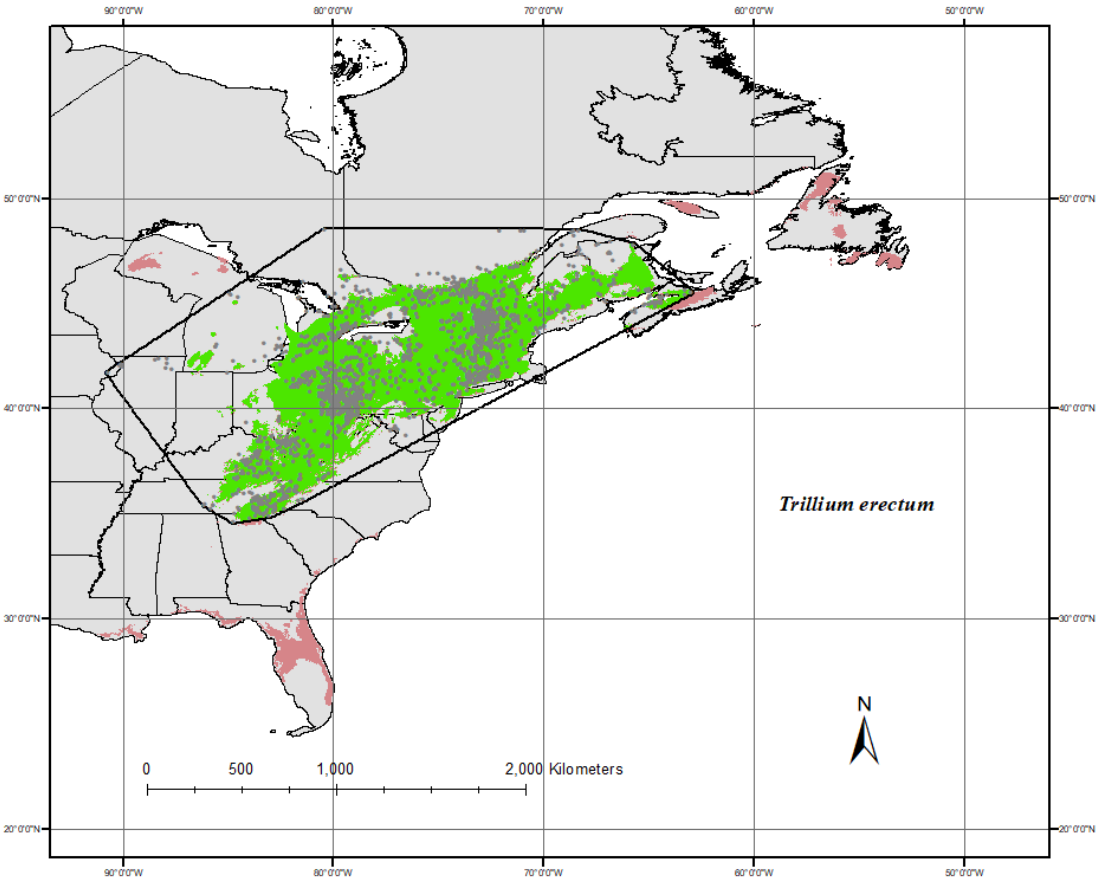
(Figure 21 continued.)

(6.)



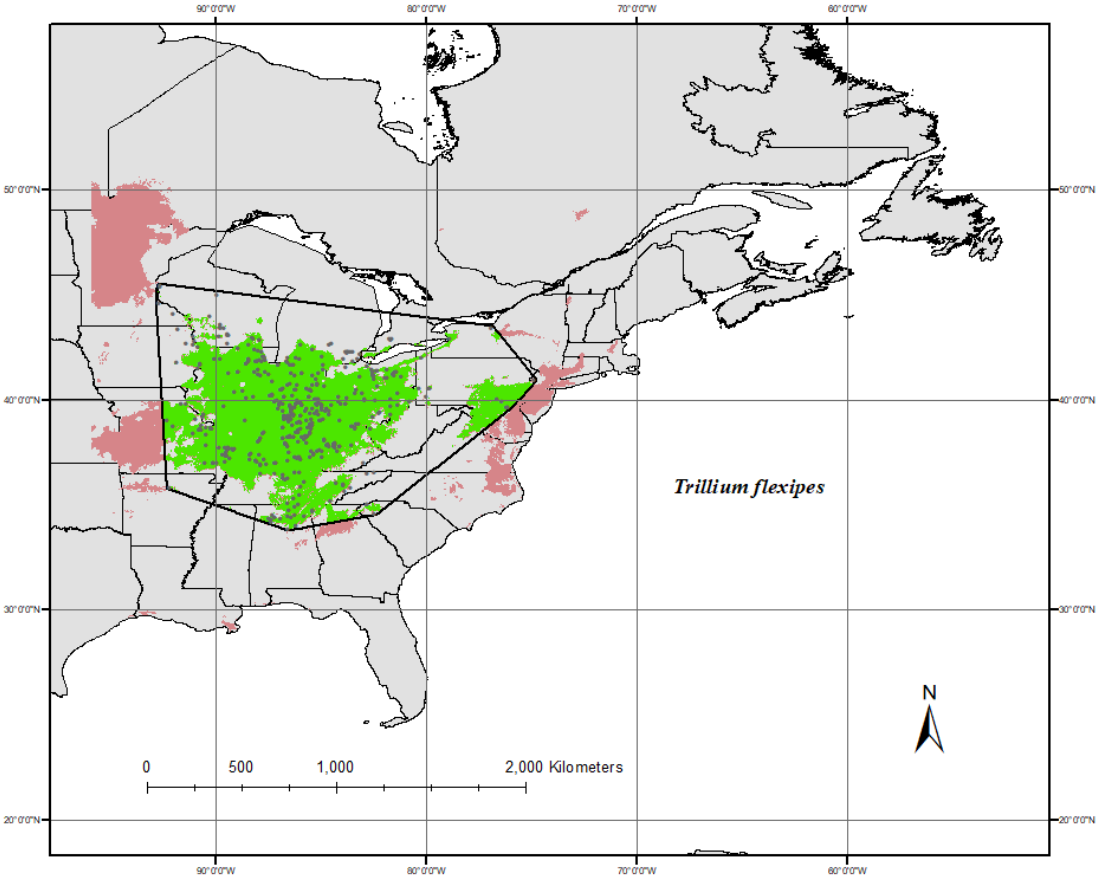
(Figure 21 continued.)

(7.)



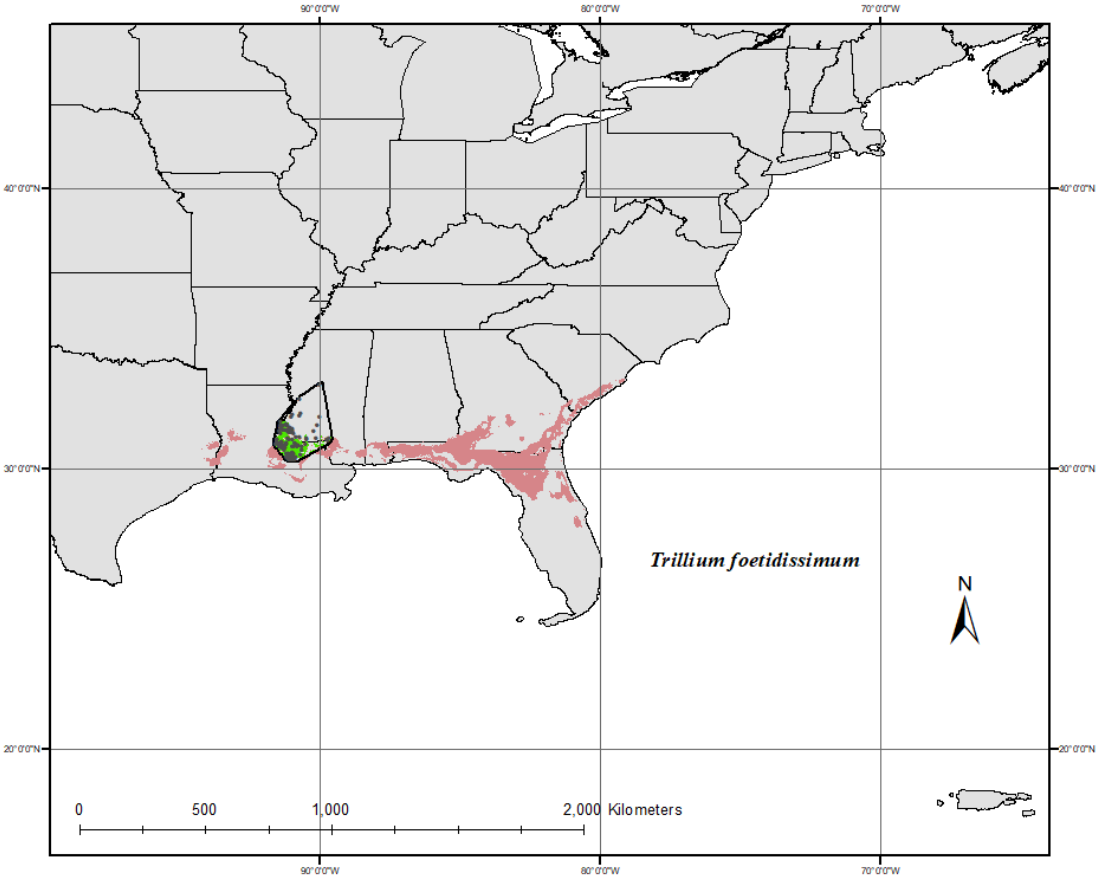
(Figure 21 continued.)

(8.)



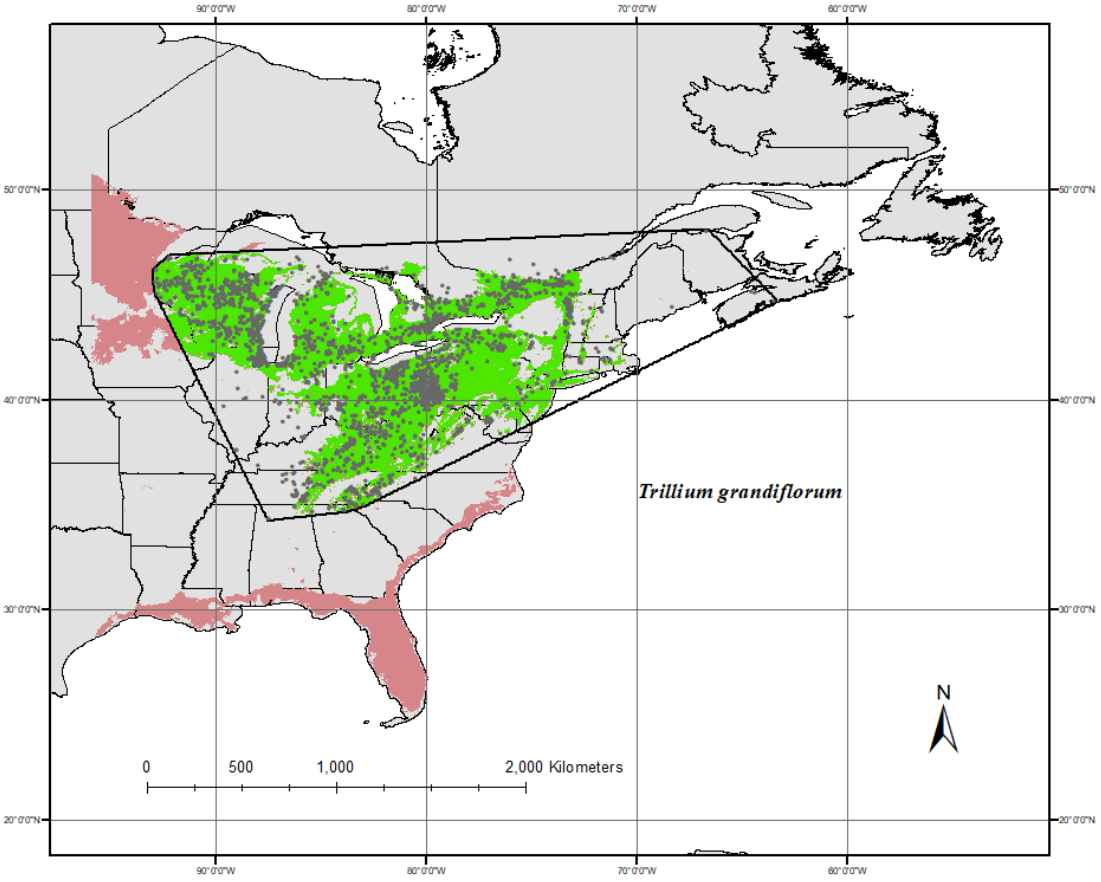
(Figure 21 continued.)

(9.)



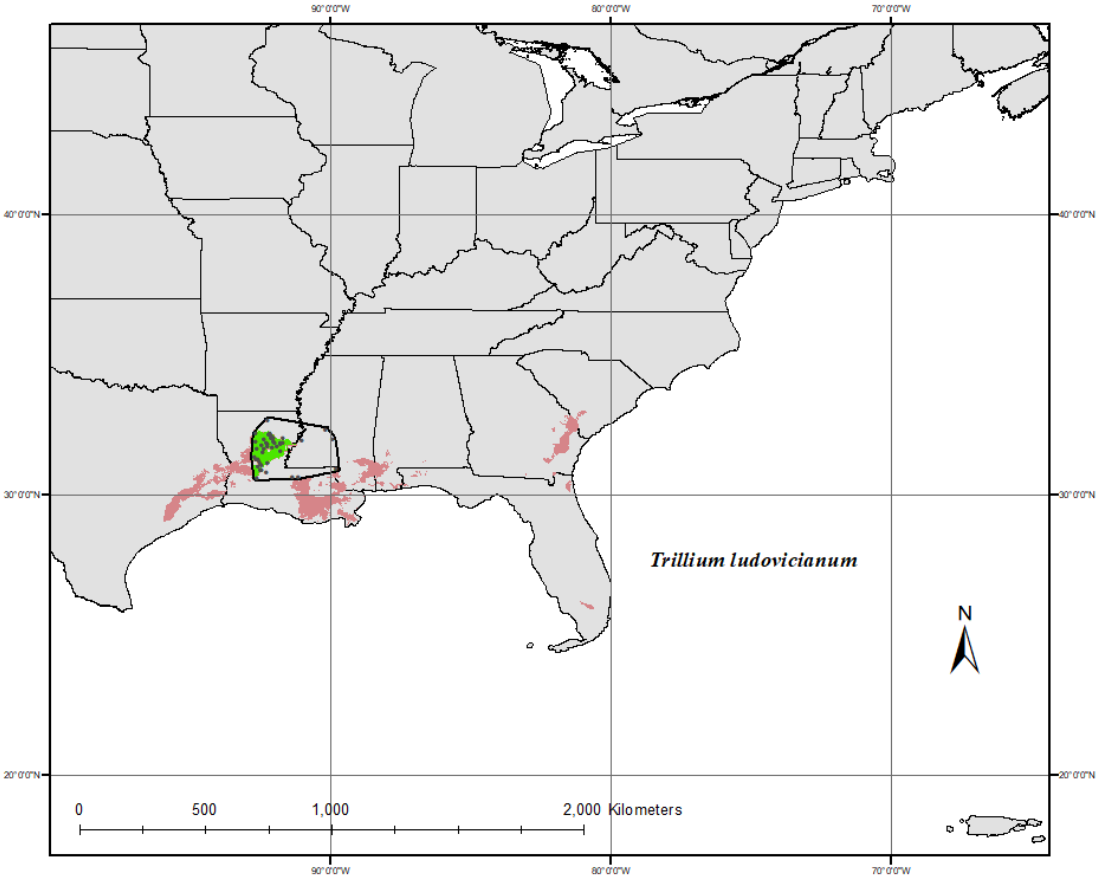
(Figure 21 continued.)

(10.)



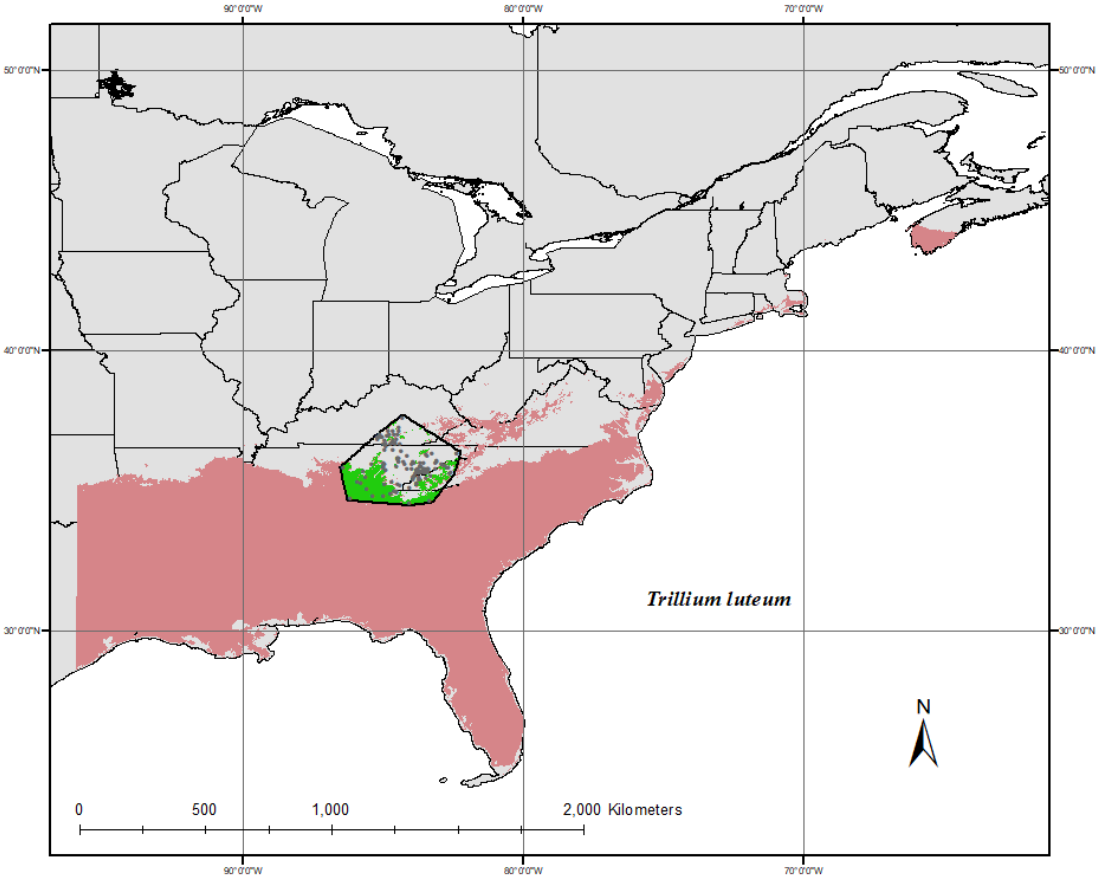
(Figure 21 continued.)

(11.)



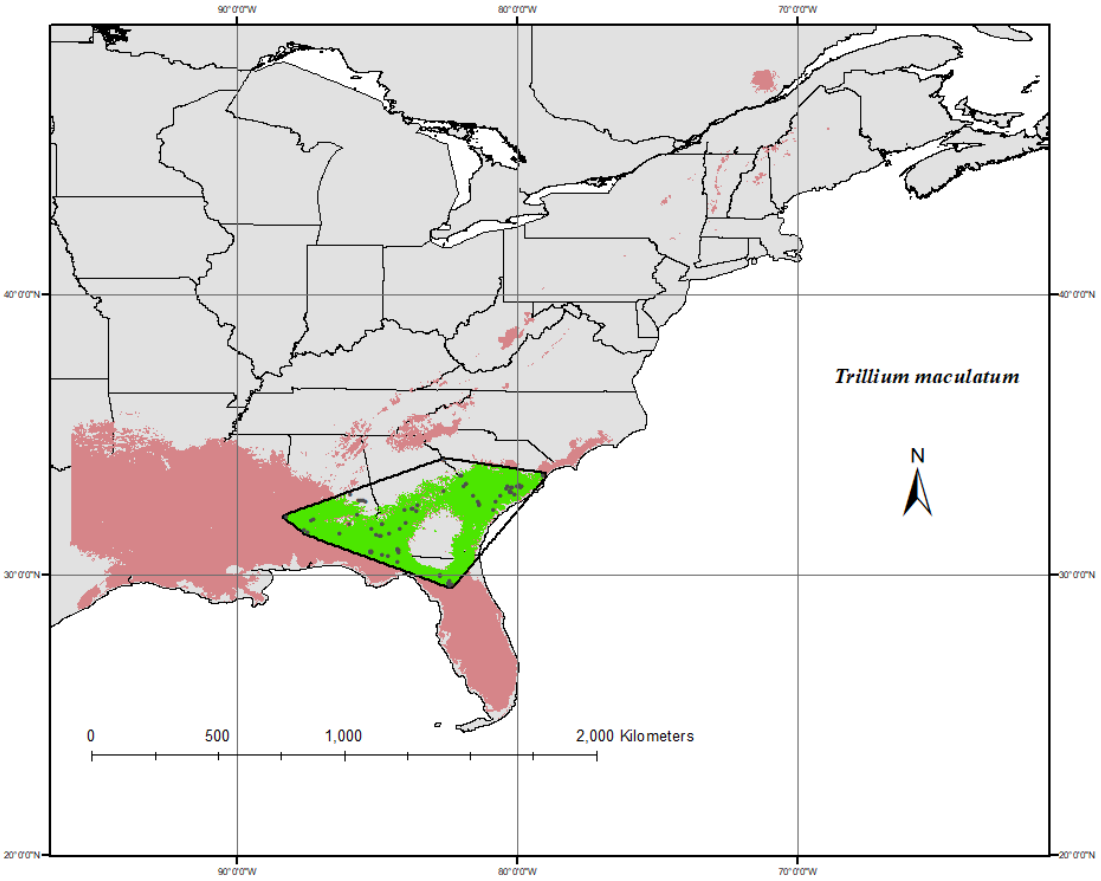
(Figure 21 continued.)

(12.)



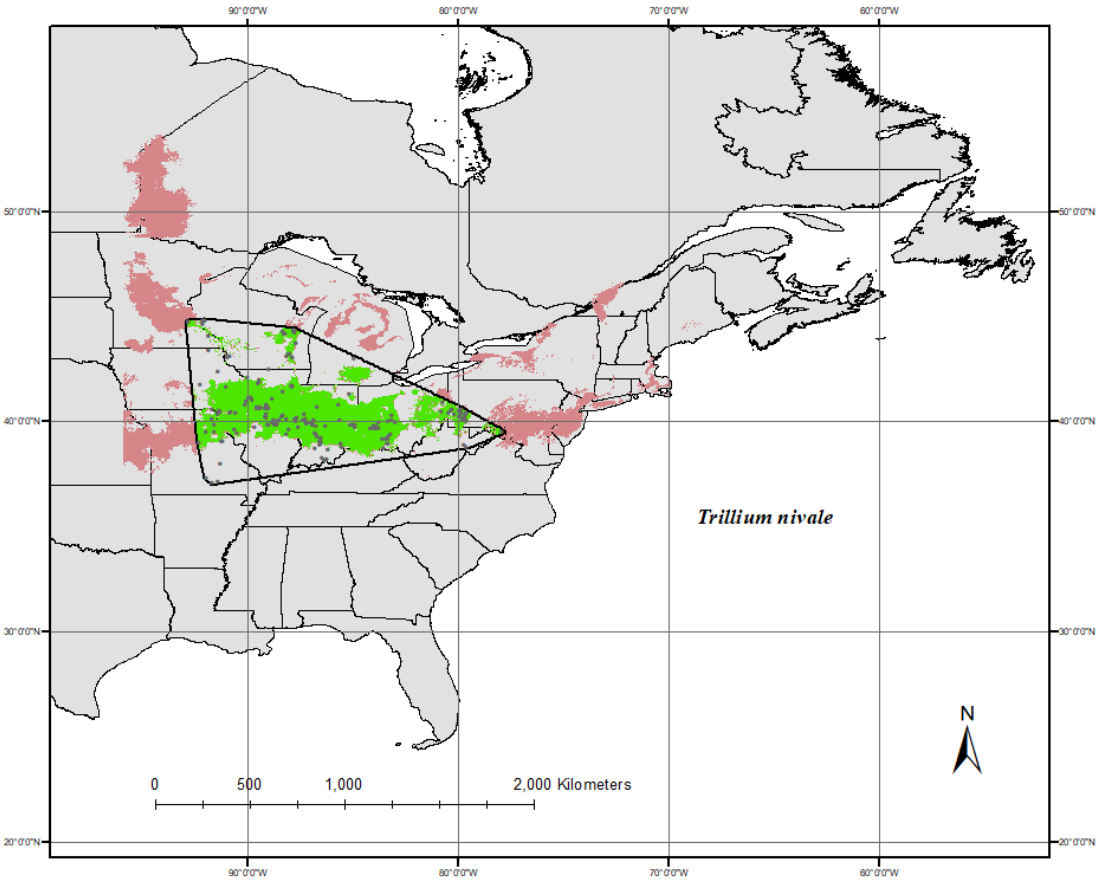
(Figure 21 continued.)

(13.)



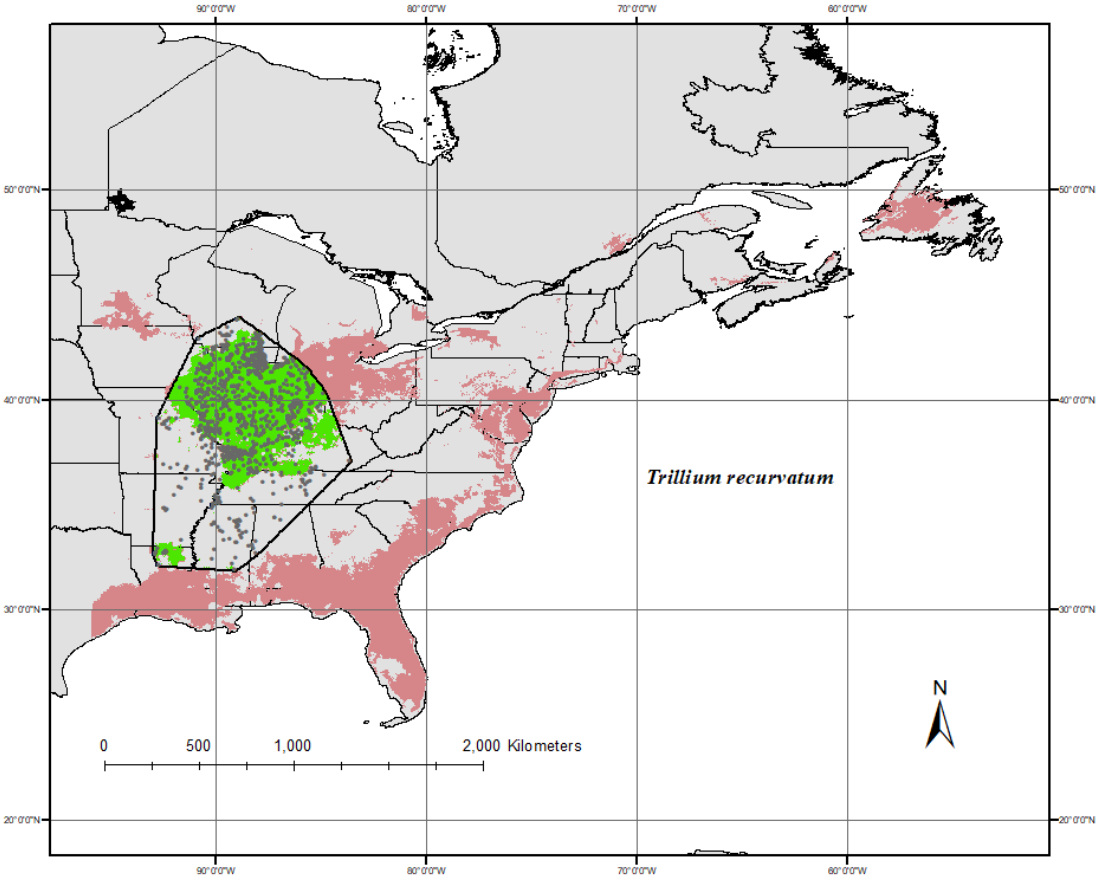
(Figure 21 continued.)

(14.)



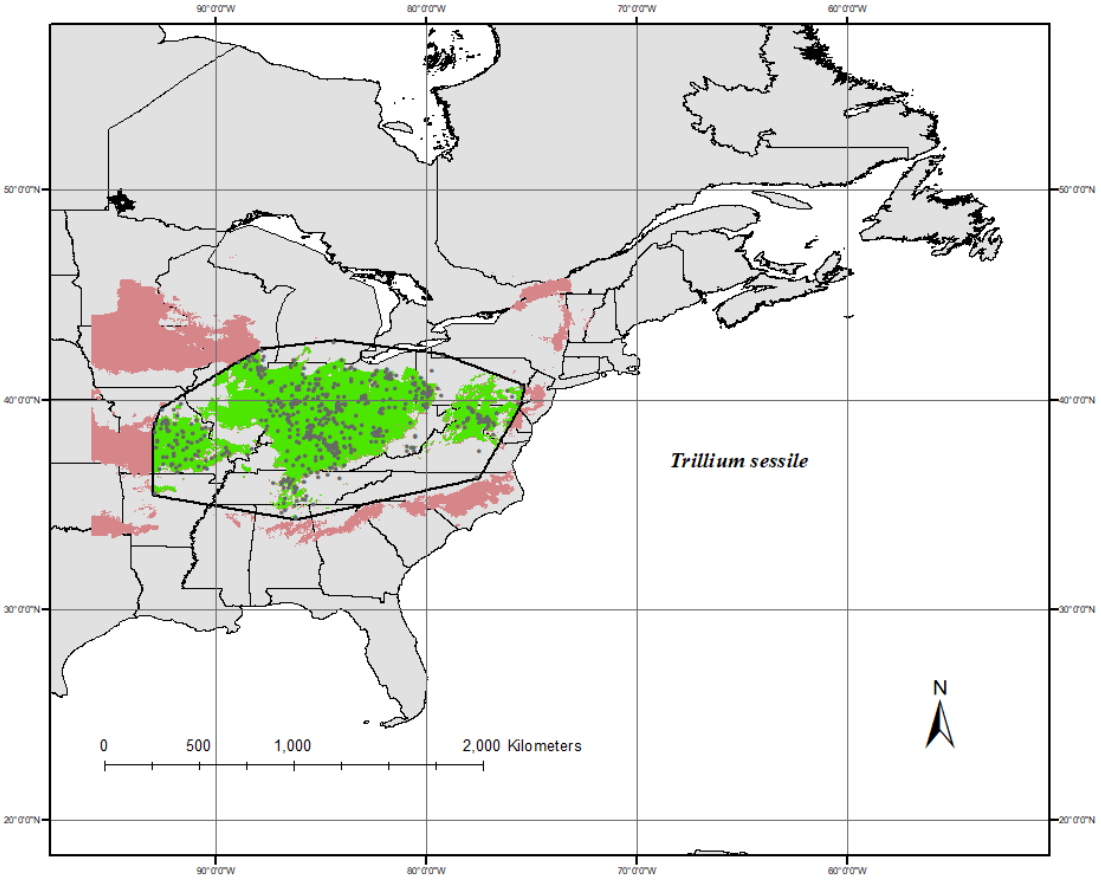
(Figure 21 continued.)

(15.)



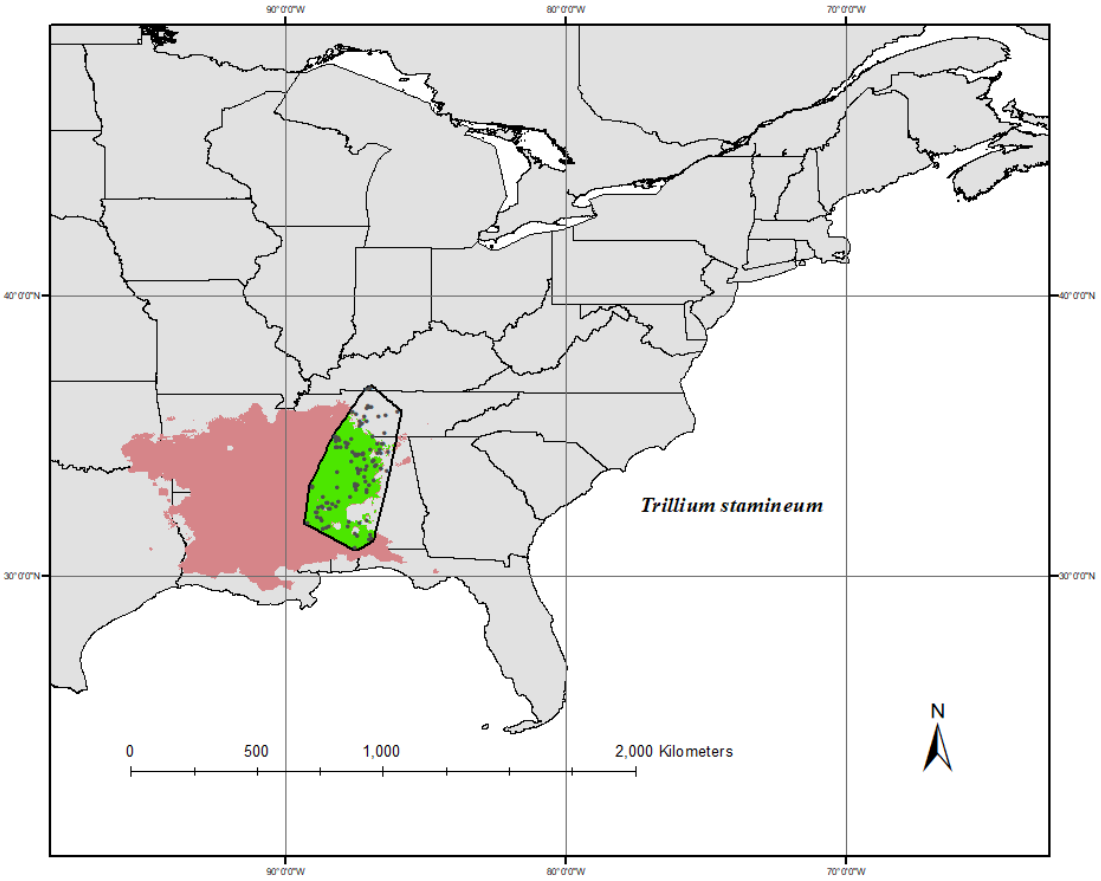
(Figure 21 continued.)

(16.)



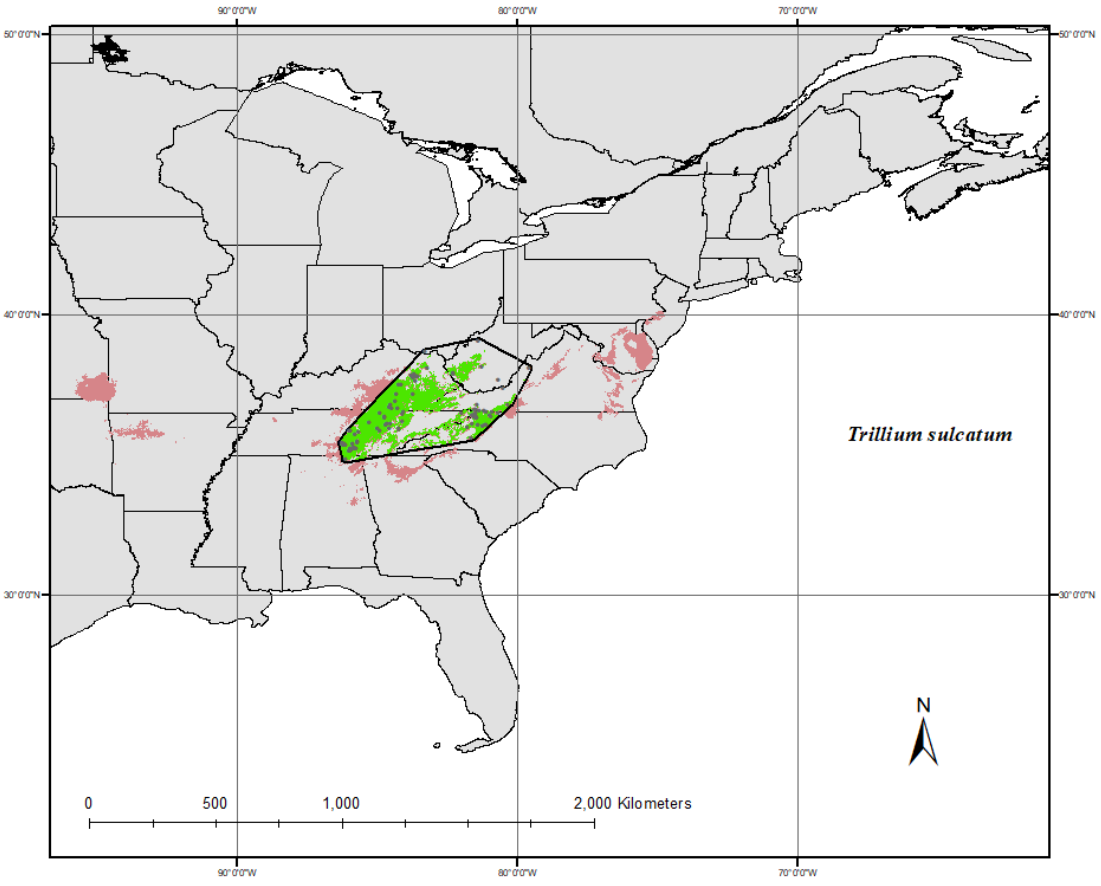
(Figure 21 continued.)

(17.)



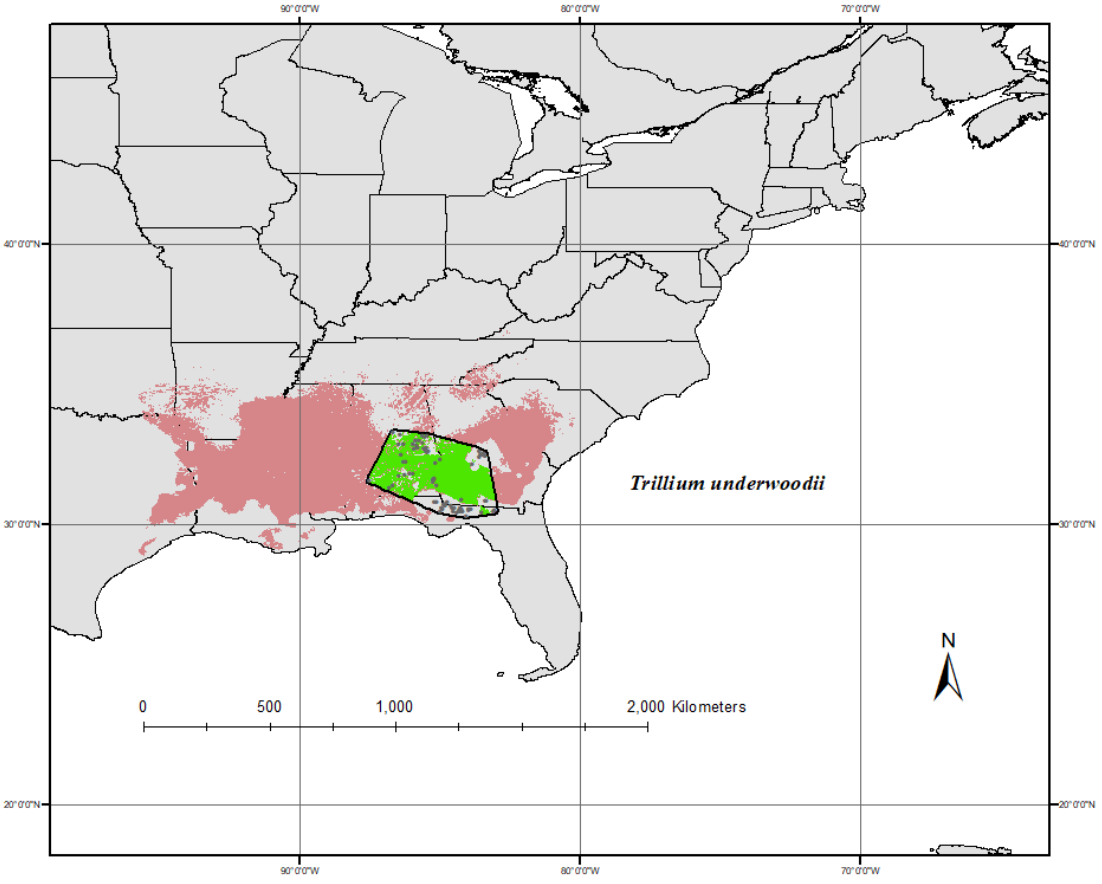
(Figure 21 continued.)

(18.)



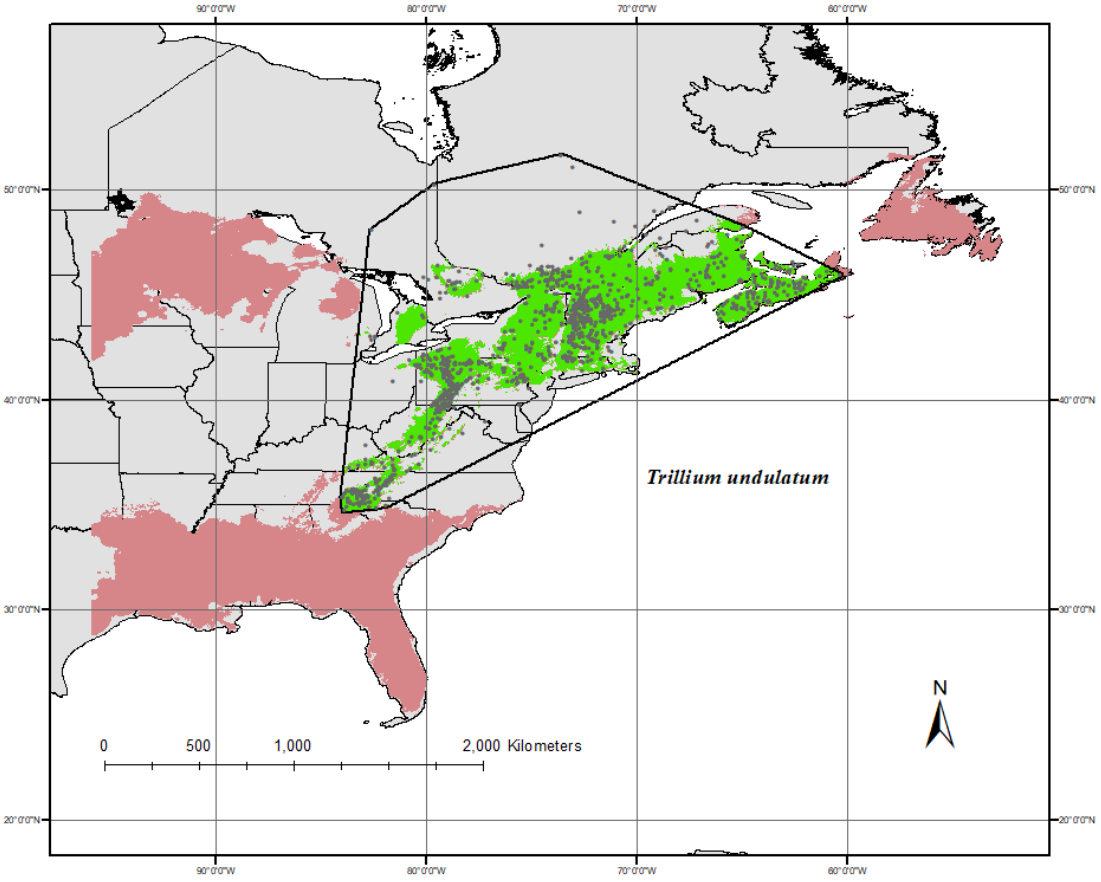
(Figure 21 continued.)

(19.)



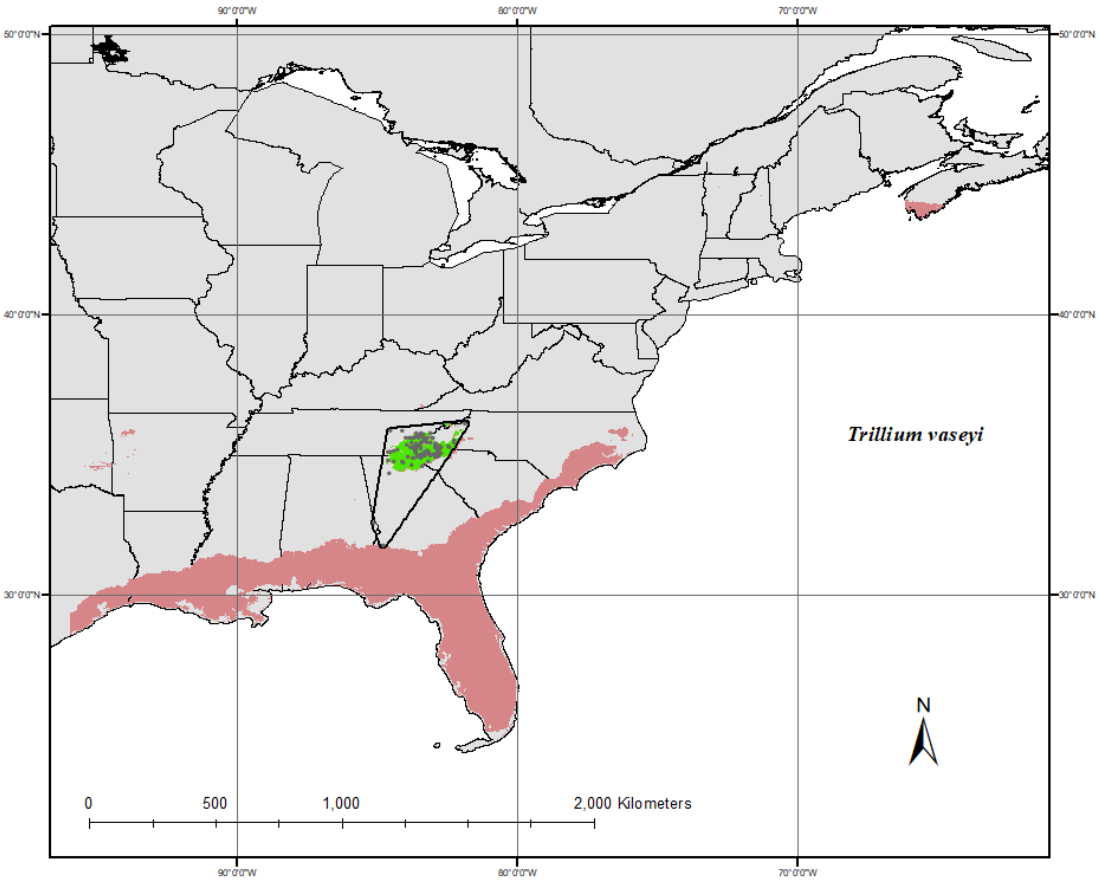
(Figure 21 continued.)

(20.)



(Figure 21 continued.)

(21.)



VITA

Chelsea Nicole Miller was born to Paul Richard Miller and Rita Maher Miller and raised in northern Illinois. She attended the University of Central Arkansas on the Norbert O. Schedler Honors College Scholarship, where she obtained a B.S. in Biology with a minor in Interdisciplinary Studies. In 2014, she was accepted to graduate school and offered a Graduate Teaching Assistantship in the Department of Ecology and Evolutionary Biology at the University of Tennessee, Knoxville (UTK), where she studied under Dr. Charles Kwit. In 2020, she obtained her M.S. in Statistics through the Intercollegiate Graduate Statistics Program and her Ph.D in Ecology & Evolutionary Biology from UTK. She hopes to pursue a career as a research scientist, and has interests in understanding the impacts of climate and anthropogenic change on species interactions.