7-22-2020

The Role of Plasticity in Bumble Bee Responses to Environmental Variability

Matthew Austin
mwacpb@mail.umsl.edu

The role of plasticity in bumble bee responses to environmental variability

Matthew W. Austin

M.S. in Biology, May, 2018, University of Missouri-St. Louis
B.S. in Biological Sciences, May, 2014, University of Missouri-Columbia

A Dissertation Submitted to The Graduate School at the University of Missouri - St. Louis in partial fulfillment of the requirements for the degree Doctor of Philosophy in Biology with an emphasis in Ecology, Evolution, and Systematics

August
2020

Advisory Committee
Aimee S. Dunlap, Ph.D.
Chairperson

Zuleyma Tang-Martínez, Ph.D.
Robert J. Marquis, Ph.D.
Patricia G. Parker, Ph.D.
Nathan Muchhala, Ph.D.
Edward M. Spevak, Ph.D.

Abstract

An aim of contemporary biology is elucidating the causes and consequences of phenotypic plasticity. Here, I approach this aim by exploring the eco-evolutionary dynamics of phenotypic plasticity and environmental variability in bumble bees (Apidae: *Bombus*), a congeneric clade of eusocial pollinating insects. Throughout their evolution, bumble bees have encountered spatiotemporal variability imposed by dynamic floral environments. Today, bumble bees additionally encounter spatiotemporal variability imposed by anthropogenic environmental change. In this dissertation, I explore how phenotypic plasticity affects how successfully bumble bees respond to environmental variability imposed by anthropogenic global change (Chapters 1 and 2) and their floral resources (Chapters 3 and 4). I focus on two notably plastic traits that have ecologically consequential implications: body size plasticity and behavioral plasticity. Using a combination of phenotypic, molecular, and modeling approaches - with data spanning field populations, biological collections, and laboratory colonies - the results of this work suggest that body size plasticity and behavioral plasticity are integral to the success of bumble bees in variable environments. I find that intraspecific trait variation is key to understanding population responses to environmental variability. Specifically, I find evidence that greater worker body size plasticity enables bumble bees to more successfully contend with anthropogenic environmental change (Chapters 1 and 2) and that behavioral variation is induced by floral variability (Chapters 3 and 4). Overall, this dissertation reveals that bumble bees respond to environmental variability in myriad ways and that these responses manifest at the individual-, colony-, and population-levels of biological organization. In addition to helping elucidate the eco-evolutionary dynamics of

phenotypic plasticity and environmental variability, this work suggests that understanding the relationship between plasticity and bumble bee success in variable environments is integral to conserving these ecologically consequential pollinators.

Keywords*: behavioral plasticity, body size,* Bombus*, conservation, human-induced rapid environmental change, microsatellite, North American Midwest, phenology, pollinator decline, trait variation*

Table of Contents

Acknowledgements

First, I thank my doctoral advisor, Aimee Dunlap, for her unwavering commitment to her students and for cultivating an open intellectual environment in which I have grown as a scientist. I will forever be grateful for Aimee's mentorship. Her ability to inform research by drawing insight from diverse fields has helped my research grow and instilled a sense of curiosity in me. I thank Zuleyma Tang-Martínez for offering limitless guidance and insight, and for helping cultivate an academic environment welcoming of everyone. I thank Nathan Muchhala, Patricia Parker, Robert Marquis, and Ed Spevak for their advice and perspectives, which have helped me become a better scientist.

I thank everyone at the University of Missouri - St. Louis (UMSL) who has supported me throughout my doctoral program. Everyone in Aimee's lab - including, but not limited to, Andreia Figueiredo, Rachel Brant, Michael Austin, Itachi Mills, Jill Lee, Tian Manning, Pablo Iturralde, Lynnsey Street, Ben Abts, Isabel Rojas-Ferrer - has been a remarkable academic family. Additionally, the Animal Behavior Discussion Group - including Dan Lehocky and Sonya Bahar - has been a welcoming intellectual home. And to everyone else in UMSL's Biology Department, who have collectively enriched my life.

Finally, I thank my family, who has been an amazing support network throughout my doctoral program. My partner, Lee Anna Sanders - along with our two dogs, Olive and Louie - has been a consistent source of love and encouragement. I owe my doctorate in many ways to my parents, Kate and Greg Austin, who raised me to pursue my passions. And to Rebecca Austin, David Tilden, and Katherine and Michael Sanders for their perpetual support throughout the years.

Introduction

Ever since bumble bees originated in Asia ~25-40 million years ago (mya) (Hines 2008), environmental variation has been key to their evolution. As the Asia-India collision drove the uplift of the Tibetan plateau (~21 mya), bumble bees began to diversify as populations exploited novel niche space created by modifications to the local landscape (Condamine & Hines 2015; Hines 2008). During a period of climatic cooling in the mid-Miocene (~14.8-14.5 mya), as Antarctic ice-sheets expanded and sea levels dropped (Condamine & Hines 2015), Asia and America were joined by the Bering Land Bridge, facilitating dispersal of bumble bees from the Palearctic to the Nearctic (Hines 2008). Following southward migration through the Nearctic, bumble bees dispersed into South America, where - under analogous circumstances to the conditions that promoted initial bumble bee diversification around the Tibetan Plateau - bumble bees diversified as Andean uplift drove the creation of novel niche space (Hines 2008).

While Andean uplift may have promoted bumble bee speciation by creating conditions for allopatry and the exploitation of novel niches, this uplift additionally created temperate habitats in which herbaceous bumble bee pollinated flora flourished (Hines 2008). Indeed, explosive plant species diversification occurred following Andean uplift, likely aided by the ongoing diversification of sympatric bumble bees (Hughes & Eastwood 2006; Hines 2008). This historic association between bumble bees and their sympatric flora is testament to the role floral variability has played in shaping bumble bee evolution. Floral variability affects bumble bee fitness in myriad ways. Bumble bees form annual colonies, with queens initiating colonies in the late-winter or early-spring, during which they act as foragers for the colony until sufficient foraging worker numbers are

produced to provide for the colony's energetic needs (Goulson 2010). Across the lifespan of a colony, angiosperm phenology and ephemerality cause rapid resource turnover, while stochasticity in the presence of energetic rewards within flowers, caused by resource depletion from competing pollinators and intraspecific variation in the presence of nectar and pollen, lead to changing associations between floral cues and rewards across time and space. How bumble bees respond to this variability has critical impacts on their survival and reproduction (e.g. Woodard et al. 2019).

To cope with the floral variability that bumble bees have encountered throughout their evolution, bumble bees have evolved numerous phenotypic traits to successfully contend with this variation. Among these traits, plasticity in body size and behavioral flexibility are particularly important when encountering floral variation, as body size and cognitive abilities critically impact foraging efficiency. Given allometric scaling between body size and tongue length, and the functional relationship between tongue length and corolla length of exploitable floral species (Miller-Struttmann et al. 2015), high body size variation within colonies can increase the diversity of floral species a colony utilizes (Peat et al. 2005). As floral resource turnover and stochasticity can change associations between floral stimuli and reward, flexible foraging behavior can further increase the diversity of floral species that bumble bees can exploit. In this dissertation, I explore how body size plasticity and behavioral flexibility affect the success of bumble bees in variable environments across different levels of biological organization (e.g. individual, colony, population).

In Chapter 1, I use phylogenetically controlled analyses on 31 North American bumble bee species to test the hypothesis that intraspecific variation in worker body size

and behavioral flexibility, as measured through a brain size proxy, make bumble bee species less susceptible to population declines in response to human-induced environmental changes. In Chapter 2, I build upon this work by testing whether bumble bees can exhibit intraspecific spatial structure in body size across an urban gradient and, if so, whether body size structure coincides with population genetic structure. Through this chapter, I also provide the first population genetic study of five bumble bee species native to the greater Saint Louis area. In Chapter 3, I explore whether bumble bee behavioral flexibility can vary within populations at different points across a reproductive season, by taking direct measurements of worker learning abilities and developing a simulation model of temporal changes to average colony-level cognition. Finally, in Chapter 4, I investigate how three co-occurring traits of floral communities - the number of flower types, reliability that flowers are associated with a reward, and signal complexity of flowers - affect bumble bee foraging behavior. Collectively, these studies suggest that body size plasticity and behavioral flexibility are integral to the success of bumble bees in variable environments.

References

Condamine, F.L. & Hines, H.M.. 2015. Historical species losses in bumblebee evolution. Biology Letters 11: 1-4.

Goulson, D. 2010. Bumblebees: Behaviour, Ecology, and Conservation. Oxford University Press.

Hines, H. M. 2008. Historical biogeography, divergence times, and diversification patterns of bumble bees (Hymenoptera: Apidae: *Bombus*). Systematic Biology 57: 58-75.

Hughes, C. & Eastwood, R.. 2006. Island radiation on a continental scale: exceptional rate of plant diversification after uplift of the Andes. PNAS 103: 10334-10339.

Miller-Struttmann, N.E., Geib, J.C., Franklin, J.D., Kevan, P.G., Holdo, R.M., Ebert-May, D., Lynn, A.M., Kettenbach, J.A., Hedrick, E., & Galen, C. 2015. Functional mismatch in a bumble bee pollination mutualism under climate change. Science 349: 1541-1544.

Peat, J., Tucker, J., & Goulson, D. 2005. Does intraspecific size var- iation in bumblebees allow colonies to efficiently exploit different flowers? Ecological Entomology 30:176–181.

Woodard, H.S., Duennes, M.A., Watrous, K.M., & Jha, S. 2019. Diet and nutritional status during early adult life have immediate and persistent effects on queen bumble bees. Conservation Physiology 7:coz048.

Chapter I: Intraspecific variation in worker body size makes North American bumble bees (*Bombus* spp.) less susceptible to decline

Matthew W. Austin* and Aimee S. Dunlap

Department of Biology, University of Missouri – St. Louis, St. Louis, MO, USA

* Corresponding author

AustinMattW@gmail.com

Published in the *American Naturalist*:

Austin, M.W. & Dunlap, A.S. 2019. Intraspecific variation in worker body size makes North American bumble bees (*Bombus* spp.) less susceptible to decline. The American Naturalist 194:381-394. DOI: 10.1086/704280

Abstract

Population declines have been documented in approximately one-third of bumble bee species. Certain drivers of these declines are known, however less is known about the interspecific trait differences that make certain species more susceptible to decline. Two traits, which have implications for responding to rapidly changed environments, may be particularly consequential for bumble bee populations: intraspecific body size variation and brain size. Bumble bee body size is highly variable and is likely adaptive at the colony level, and brain size correlates with cognitive traits (e.g. behavioral plasticity) in many groups. Trait variation and plasticity may buffer species against negative effects of rapidly changed environments. Using phylogenetically controlled analyses of 31 North American bumble bee species, we find higher intraspecific body size variation is associated with species having increased their relative abundance over time. However, this variation does not significantly interact with tongue length, another trait thought to influence bees' decline susceptibility. Head size, a proxy for brain size, is not correlated with change in relative abundance. Our results support the hypothesis that variation in body size makes species less susceptible to decline in rapidly altered environments and suggests that this variation is important to the success of bumble bee populations.

Keywords*: behavioral plasticity, human-induced rapid environmental change, IUCN, museum data, pollinator decline, trait variation*

Introduction

Bumble bees (Apidae: *Bombus*) are native pollinators throughout much of the Northern Hemisphere and South America, serving essential functional roles in terrestrial ecosystems through their maintenance of biodiversity, wild plant communities, and cultivated crop production (Goulson 2010; Potts et al. 2010). Numerous reports have documented declines of bumble bees (e.g. Williams et al. 2009; Colla et al. 2012; Hatfield et al. 2014; Cameron et al. 2011a), with a recent assessment finding that approximately one-third of extant bumble bees are in decline (Arbetman et al. 2017). While several drivers are attributed to these declines - such as pesticides, parasites/pathogens, and invasive species (Williams et al. 2009; Cameron et al. 2011b; McArt et al. 2017; Siviter et al. 2018) - predominant among them is human-induced rapid environmental change, including climate change (Kerr et al. 2015) and habitat loss (Kosior et al. 2007). Given the ubiquity of such environmental changes throughout the native range of bumble bees, many researchers have investigated why only a fraction of bumble bees are susceptible to decline, while other species are thriving (e.g. Williams et al. 2009; Cameron et al. 2011a; Arbetman et al. 2017). These investigations have revealed numerous traits that may make certain bumble bee species more susceptible to decline (e.g. specialization: Bartomeus et al. 2013), however such traits often receive mixed support among studies (e.g. Williams 2005; Williams et al. 2009; Arbetman et al. 2017). It is clear that the causes of bumble bee declines are multifaceted, and while certain trends have emerged from the literature, there is a need to more fully understand the interspecific variation that has led to disparate population trends between these closely related species.

To understand why certain species are more susceptible than others, one should consider what factors influence how species respond to rapid environmental change. One factor that may be particularly important is the degree of intraspecific trait variation that a population contains. Intraspecific variation can have large ecological consequences; trait heterogeneity can affect demographic variance (Vindenes et al. 2008), genetic variation underlying these traits can promote species coexistence (Imura et al. 2003; Vellend 2006), and intraspecific trait variation is the raw material for adaptation by natural selection and enabling evolutionary change in response to environmental changes (Darwin 1859; Bolnick et al. 2011). Models suggest that intraspecific trait variation may buffer populations against fluctuations in population density (Bolnick et al. 2011) and declines in response to environmental stochasticity (Filin & Ovadia 2007). As the traits enabling invasion success are those traits that allow organisms to successfully contend with a novel environment, the invasion biology literature can provide insight on the traits that should influence the susceptibility of species to rapid environmental changes. Indeed, this literature shows empirical support for higher levels of intraspecific variation increasing establishment success of introduced species (e.g. Forsam 2014; González-Suárez et al. 2015). Studies also suggest that species with low intraspecific variation are more vulnerable to extinction, possibly due to a diminished capability of successfully responding to environmental changes (González-Suárez & Revilla 2013; Kolbe et al. 2011; Liow 2007). Despite this evidence, many studies have overlooked the role that intraspecific trait variation might have in influencing population stability, and have instead focused on mean trait values (e.g. González-Suárez et al. 2015).

In insects, body size is one trait that has particularly notable ecological consequences (Chown & Gaston 2010); for certain insect taxa, large body size makes species more prone to extinction (e.g. Grimbacher et al. 2008). Several studies have revealed this trend in bees (e.g. Bartomeus et al. 2013; Scheper et al. 2014), where species with a larger average body size have an increased chance of decline, perhaps due to a limiting effect of their greater pollen and feeding requirements during development in periods of food scarcity. However, this trend may not be consistently predictive (Williams et al. 2010). In bumble bees, while average body size may be predictive of decline, intraspecific variation in body size may be similarly ecologically relevant. Body size is highly variable within bumble bees; within a colony, workers may exhibit up to a tenfold difference in body size (Couvillon & Dornhaus 2009; Goulson et al. 2010). Investigations into the function of bumble bee body size collectively indicate that this variation may be adaptive at the colony level (but see Jandt & Dornhaus 2014 and Herrmann et al. 2018). Larger workers are more efficient foragers (Spaethe & Weidenmüller 2002), less likely to be predated (Goulson 2010), and are more efficient at nursing brood than smaller bees (Cnaani & Hefetz 1994). Smaller workers, on the other hand, can withstand nectar scarcity for longer periods of time than larger bees (Couvillon & Dornhaus 2010). Additionally, given a positive correlation between body size and tongue length, and that tongue length influences which floral species a bee forages from, high variation in body size may allow a colony to efficiently forage from a range of floral species (Peat et al. 2005). Therefore, high within-colony variation in body size may be adaptive, particularly in environments that experience rapid rates of floral turnover and periods of food scarcity. Despite this evidence that variation in body size may be adaptive

for bumble bees, whether high intraspecific variation in body size makes bumble bees less prone to population declines has not yet been explored.

When environmental change is induced by human activity, it often occurs rapidly and fragments previously continuous habitat (Vitousek et al. 1997). Accordingly, it is often surmised that the rate of such environmental changes exceeds the evolutionary rate of many populations, and that the fragmented habitat creates barriers to dispersal (Snell-Rood 2013; Wong & Candolin 2015). Given this inhibition of adaptation and dispersal for populations encountering a rapidly changed environment, plastic responses of individuals may be particularly important for buffering populations against the negative effects of a rapidly changed environment (Tuomainen & Candolin 2011; Snell-Rood 2013; Wong & Candolin 2015). In particular, behavioral plasticity is thought to play an important role as behavior is sensitive to environmental changes (Snell-Rood 2013) and can alter key demographic parameters (e.g. birth, death, migration) (Tuomainen & Candolin 2011; Wong & Candolin 2015). If an animal is able to plastically match their behavior to environmental novelties, they have an increased chance of survival in environments rapidly altered by human activity (Sih et al. 2011; Snell-Rood 2013; Tello-Ramos et al. 2018). This idea is supported by studies on species invasions that find that species with high phenotypic plasticity are more likely to be successful invaders (e.g. Lodge 1993; Sol et al. 2002; Knop & Reusser 2012; Davidson et al. 2011). Studies on migratory birds further support that behavioral plasticity promotes survival in harsh environments (e.g. Vincze 2016; Roth et al. 2010).

If behavioral plasticity is predictive of a species' success in a human-altered environment, then how can behavioral plasticity be measured for comparative studies?

Several studies have proposed that relative brain size can be used as a proxy for behavioral plasticity in comparative studies, with larger relative brain size conferring greater behavioral plasticity (Sol et al. 2008; Sol 2009). While there is contentious debate as to the function of brain size, and several theories have been proposed as explanations for the evolution of brain size (e.g. social brain hypothesis, ecological problem solving, brain tissue trade-offs; Dunbar & Shultz 2007, Snell-Rood et al. 2011, Kotrschal et al. 2013a, 2013b), known processes of brain function support the idea that an organism's capacity for behavioral plasticity is mediated by relative brain size. Large brains can increase cognitive capacity and produce qualitatively novel behaviors by containing a greater number of neuronal circuits (Chittka & Niven 2009). Additionally, relative brain size is considered more representative of behavioral plasticity than absolute brain size as cognitive processes are partly determined by the amount of energy allocated to neural functioning. Thus, behavioral plasticity is better reflected when the metabolic constraints of body size are considered relative to brain size (Chittka & Niven 2009). Indeed, comparative studies in mammals (Sol et al. 2008), birds (Sol et al. 2002), reptiles and amphibians (Amiel et al. 2011), have supported the idea that large relative brain size confers a fitness benefit in rapidly changed environments, thereby suggesting relative brain size as an appropriate proxy for behavioral plasticity and that behavioral plasticity makes species less susceptible to decline in environments altered by human activity.

Here, we test the hypothesis that intraspecific variation in body size and behavioral plasticity, as measured through a brain size proxy, make bumble bee species less susceptible to population declines in response to human-induced environmental changes. To accomplish this, we study 31 species of bumble bees native to North

America, using specimens from multiple natural history collections and decline assessments from the International Union for Conservation of Nature (IUCN). As bumble bee brain volume positively correlates with head width (Mares et al. 2005; Riveros & Gronenberg 2010), we use relative head size as a proxy for behavioral plasticity in the absence of direct measurements of relative brain size. We predict 1) bumble bee declines will be associated with low intraspecific variation in body size and 2) bumble bee declines will be associated with small relative head size.

Methods

Phenotypic Measurements

We measured 977 worker bumble bees of 31 species from four natural history collections: Smithsonian Institution's National Museum of Natural History, American Museum of Natural History, Field Museum of Natural History, and Illinois Natural History Survey. Prior to inclusion in this study, all specimens were taxonomically identified to the species-level. For each specimen, we recorded full label data and only included specimens that were collected in North America, north of Mexico. We included no more than two conspecific bees if they were collected in the same year and locality as one another. The spatial distributions of these specimens for each species are depicted in maps found in the Supplemental Materials (Figs. S1-S31; mapping methods given in the Appendix).

To obtain body size and relative head width measurements, we first took dorsal photographs of each bee's thorax and head against a known unit of distance. To accomplish this, we pinned each specimen to a foam platform, positioned against a solid

white background, and aligned a ruler to the bee's frontal plane. We took dorsal photographs with a Canon EOS Rebel T5 (Canon EF-S Macro 60 mm lens) mounted approximately 14 cm away from the specimen. We photographed each specimen's thorax and head separately. Subsequently, we measured thorax width and head width from these photographs in ImageJ 1.50i. For these measurements, we set the photograph's scale using a 1 mm segment of the ruler as a known distance and then took width measurements using the 'straight line' tool. To obtain relative head size measurements, we averaged head width and thorax width measurements per species. We then performed a regression of head width and thorax width averages, and took the residuals from this regression as relative head size measurements (Fig. 1). This is standard practice for calculating relative brain size for interspecific comparisons of behavioral plasticity in cognitive ecology (e.g. Sol et al. 2005; Carrete & Tella 2011). Positive residuals indicate that a species has a larger head width than would be expected for their thorax width, on average; negative residuals indicate that a species has a smaller head width than would be expected for their thorax width, on average. To quantify intraspecific variation in body size, we calculated a coefficient of variation (CV) for thorax width per species.

Population Trends

We used measures of change in relative abundance for data on population trend, as calculated by Hatfield et al. (2014) for IUCN assessments of North American bumble bees. These assessments were developed in conformation with the IUCN Red List Criteria, the standard for assessing extinction risk across taxonomic groups (IUCN Standards and Petitions Subcommittee 2017). For these assessments, Hatfield et al.

utilized a database comprised of approximately 300,000 specimen records of bumble bees collected in North America north of Mexico (i.e. from the United States and Canada; database compiled by Williams et al. 2014) obtained from numerous academic, private, research, and citizen science collections. To help reduce bias that may result from using presence-only data from natural history collections, Hatfield et al. employed several quality control measures. First, collections were dropped from analysis if they had not completely digitized their entire *Bombus* collection. Second, specimens were removed if they had not been identified to the species-level, lacked needed label data, and/or were collected from outside of that species' known range. Finally, species were dropped from the dataset if they had a low sample size or only partial coverage throughout their geographic range. Accordingly, the final database comprised 202,198 specimen records of bumble bees collected throughout North America from approximately 150 collections (Hatfield et al. 2014). This quality controlled database helped ensure (i) no species had a biased abundance relative to the other included species and (ii) contained only species with entire geographic coverage.

To calculate change in relative abundance, Hatfield et al. split these records into a historic (1805 - 2001, $N = 128,572$) and a current (2002 - 2012, $N = 73,626$) time period and calculated relative abundance per species for each of these periods. For each relative abundance measurement, they divided the number of observations for a bumble bee species in that time period by the total number of bumble bee observations for that time period [i.e. relative abundance = (number of *Bombus* sp. observations)/(total number of *Bombus* spp. observations)] (R. Hatfield, personal communication). Subsequently, Hatfield et al. calculated change in relative abundance by dividing each species' current

relative abundance by their historic relative abundance. Given as a percentage, values <100% indicate a decrease in relative abundance, values >100% indicate an increase in relative abundance, and a value of 100% indicates no change in relative abundance.

Analyses

To determine whether intraspecific variation in body size and relative head size are correlated with population trend in North American bumble bees, we performed a phylogenetic generalized least squares (PGLS) analysis. In this analysis, we included relative head size residuals and thorax width CVs as predictor variables, and change in relative abundance as the response variable. Subsequently, we asked: if either of these traits significantly correlated with population trend in our first model, do they significantly interact with tongue length, another known correlate of bumble bee population trends? To accomplish this, we obtained tongue length data from Arbetman et al. (2017) and performed an additional PGLS with tongue length as a predictor variable and change in relative abundance as the response variable. We additionally included relative head size residuals and/or thorax width CV as predictor variables, if they were significantly correlated with change in relative abundance in the first model. To obtain tongue length, Arbetman et al. averaged tongue length measurements (i.e. sum of glossa and prementum lengths) per species from a comprehensive literature search. In each of these models, we controlled for phylogenetic relationships between species with the contemporarily most comprehensive *Bombus* phylogeny (Cameron et al. 2007; Hines 2008), pruned to include only the species in each model and forced ultrametric prior to analyses (Fig. 2). We assessed each variable (i.e. change in relative abundance, thorax

width CV, relative head size residuals, tongue length) for phylogenetic signal using Blomberg's *K* (Blomberg et al. 2003; phylogenetic signal methods and results given in the Appendix). For all analyses, we used RStudio (version 0.99.902): the phylogenetic tree was pruned and forced ultrametric using the APE (Paradis & Schliep 2018), GEIGER (Harmon et al. 2008), and PHYTOOLS (Revell 2012) packages and our PGLS analyses were performed with the NLME package (Pinheiro et al. 2018).

Results

Of the 31 species analyzed in this study, 12 (38.7%) increased and 19 (61.3%) decreased their relative abundance from historic (1805 - 2001) to current (2002 - 2012) time periods. These values of change in relative abundance ranged from a minimum of 2.32% (*B. crotchii*) to a maximum of 294.17% (*B. impatiens*). We obtained a sample size of at least 21 bees for each species, for a total of 977 bees. See table 1 for all sample sizes and trait values obtained per species.

Our initial PGLS model shows that thorax width CV is significantly correlated with change in relative abundance among North American bumble bees ($p<0.001$) (Fig. 3). Species with higher intraspecific variation in body size are more likely to have increased their relative abundance from historic to current time periods. Change in relative abundance is not significantly correlated with either relative head size residuals ($p=0.562$) (Fig. 3) or the interaction between thorax width CV and relative head size residuals ($p=0.577$). To ensure that the positive correlation between body size variation and change in relative abundance was not the result of different sample sizes between species, we asked whether sample size was predictive of either body size variation or

change in relative abundance. To answer this question, we performed two linear models, each of which included sample size as the predictor variable and included either thorax width CV or change in relative abundance as the response variable. We find that sample size does not significantly correlate with either of these variables (thorax width CV, p=0.410; change in relative abundance, p=0.282). See Fig. 4 for box plots of thorax widths obtained per species.

As our first PGLS revealed a significant association between thorax width CV and change in relative abundance, we subsequently asked if body size variation might interact with tongue length, another known correlate of bumble bee decline, to affect change in relative abundance. To answer this question, we performed an additional PGLS that included body size variation and tongue length as predictor variables, and change in relative abundance as the response variable. We removed *B. caliginosus*, *B. crotchii*, and *B. sandersoni* from this analysis, due to missing tongue length data for these species. This model shows that body size variation does not significantly interact with tongue length to affect change in relative abundance (CV thorax width:tongue length, p=0.140) (Fig. 5). This model also found a significant univariate effect of body size variation (p<0.05) and a non-significant univariate effect of tongue length (p=0.160) predicting change in relative abundance. See table 2 for full results from these PGLS analyses. Data underlying all analyses can be found in the Dryad Digital Repository: https://doi.org/10.5061/dryad.910105r (Austin & Dunlap 2019).

Discussion

Using specimen data from multiple natural history collections throughout North America, we assessed whether intraspecific variation in body size or relative head size are predictive of changes in relative abundance for North American bumble bees. We found that species with higher intraspecific variation in body size were more likely to have increased their relative abundance from historic (1805 - 2001) to current (2002 - 2012) time periods. Relative head size was not predictive of changes in relative abundance. This study is the first to assess whether population trends of bumble bees may be affected by how variable worker body size is within species. Our results support the hypothesis that variation in body size makes species less susceptible to decline in rapidly altered environments and suggest that greater intraspecific body size variation is associated with increased relative abundance. Given bumble bees' integral pollination services in native ecosystems and agriculture, there is a need to understand the factors underlying their declines.

Intraspecific trait variation is ecologically consequential, as it can affect demographic variance (Vindenes et al. 2008), may buffer against fluctuations in population density (Bolnick et al. 2011), and may allow species to successfully contend with environmental changes (Filin & Ovadia 2007). We find that North American bumble bees with higher intraspecific variation in worker body size were less susceptible to decline and this increased variation is associated with increased abundance. This adds to the growing literature on the ecological importance of intraspecific trait variation and suggests that body size is important for bumble bee population dynamics. However, as we avoided measuring bees from the same colony, we do not know whether our data represent differences between species in within-colony worker size variation or

differences between species in how variable mean worker size is between intraspecific colonies. This is an important consideration when interpreting our results as the benefits of body size variation can manifest differently on the colony-level versus the population-level. At the colony-level, there may be up to a tenfold difference in worker body size despite workers typically being highly related (r = 0.75) (Couvillon & Dornhaus 2009; Goulson et al. 2010). Adult size of bumble bees is positively correlated with the quantity of food that a bee receives during development (Pendrel & Plowright 1981; Sutcliffe & Plowright 1988; Pereboom et al. 2003). Therefore, this within-colony size variation is largely a result of unequal rates of larval feeding, which are partially a function of larval cell location within the colony (Couvillon & Dornhaus 2009). This within-colony size variation may have important fitness consequences. Larger workers are typically more efficient at foraging (Spaethe & Weidenmüller 2002) and raising brood (Cnaani & Hefetz 1994), while smaller workers can withstand starvation for longer periods of time (Couvillon & Dornhaus 2010). Additionally, high variation in body size promotes the exploitation of a greater variety of floral species (Peat et al. 2005). Accordingly, variation in worker body size may be adaptive at the colony level, as this variation can promote colony efficiency while also providing an insurance policy in times of food shortage. The need for both large workers that are efficient foragers and small workers that can withstand periods of food shortage may be particularly important in the wake of human-induced environmental changes, as such changes can promote environmental stochasticity and decrease floral diversity (Jackson & Sax 2010). At the population-level, intraspecific trait variation can promote population stability in changed environments. Founder groups may have increased establishment success if they have greater

intraspecific variation (Forsam et al. 2012; Forsam 2014; González-Suárez et al. 2015) and high intraspecific variation may buffer species against extinction (González-Suárez & Revilla 2013). This has even been supported by the fossil record, which shows that species with low morphological variation went extinct faster than comparable species with greater morphological variation (Liow 2007; Kolbe et al. 2011). An interpretation that can collectively explain these findings is that intraspecific trait variation can provide species with an ability to flexibly respond to a changed environment (González-Suárez et al. 2015). This may be mediated through some individuals being pre-adapted to the changed environmental conditions. Alternatively, the correlation between intraspecific trait variation and population stability may not be causative. The spatial and temporal distributions of specimens we measured for this study are important for the interpretation of our results. Our data collection protocol was designed to help ensure that no time period or location, as represented by specimens in each natural history collection, was biasedly sampled. However, to address our measured specimens' spatial and temporal distributions, we have included a map per species of locations where specimens were collected and analyses of latitudinal and temporal trends in body size in the Appendix and Supplemental Material (Table S2, Figs. S1-S31). Collectively, these suggest that our specimens have broad spatial coverage and our measures of body size variation reflect standing variation throughout each species' range. As our data do not quantify within-colony variation in worker size, additional study on comparative variation in bumble bee body size is needed to more fully resolve how intraspecific variation in body size affects population stability in bumble bees. While body size variation can manifest differently on the colony- and population-levels, the consequences of body size variation at these two

levels are not mutually exclusive and indeed may both have consequences for bumble bee population dynamics.

Declines of bumble bees are multifaceted, with anthropogenic land use being a predominant driver of these declines (Williams et al. 2009; Kerr et al. 2015; Kosior et al. 2007). Numerous studies have investigated potential correlates of bee decline (e.g. Williams et al. 2009; Cameron et al. 2011a; Arbetman et al. 2017), and while several themes have emerged from this literature, mixed support for traits has been produced as well (Williams 2005; Williams et al. 2009; Arbetman et al. 2017). This mixed support is likely due to interrelationships among traits that cannot be captured by single analyses. Nonetheless, among the themes that have emerged from this literature is the relation of bee decline to dietary specialization (Bartomeus et al. 2013). The lack of a significant interaction found from our analysis of body size variation and tongue length (i.e. dietary specialization proxy) may reflect no true functional interaction between these traits, however it may alternatively reflect the difficulty of statistically resolving interrelationships among facets of bee declines. While our results suggest that high intraspecific variation in body size makes bumble bees less susceptible to decline, we emphasize that many traits likely affect this susceptibility (e.g. pesticide tolerance, immunity), which may take a predominant role to body size variation. To illustrate this point, consider the common eastern bumble bee (*B. impatiens*). *B. impatiens* is an extremely successful species - they are the only currently commercially available bumble bee species in the United States (Koppert Biological Systems) and have the greatest increase in relative abundance among our analyzed species (294.17%; Hatfield et al. 2014) - however *B. impatiens* has a thorax width CV that is below the average of the

species included in this study (*B. impatiens* thorax width CV = 9.964; average thorax width CV = 10.220). Clearly, factors other than body size variation must have influenced their success. Nevertheless, to more fully understand the many facets that have led to disparate population trends among bumble bee species, our results suggest that body size variation be considered.

In a rapidly changing environment, an organism's survival may depend on its ability to plastically match its behavior to the changed environmental conditions (Tuomainen & Candolin 2011; Snell-Rood 2013; Wong & Candolin 2015). Behavioral plasticity is often viewed as an indication of cognitive complexity, whereby individuals with greater behavioral plasticity are treated as having enhanced cognitive abilities overall (Mikhalevich et al. 2017). Using relative brain size as a proxy for behavioral plasticity, comparative studies (Sol et al. 2008; Sol et al. 2002; Amiel et al. 2011) have supported the hypothesis that behavioral plasticity buffers populations against the negative effects of rapid environmental change. Here, we invoked these predictions in investigating bumble bee declines by using relative head size as a proxy for behavioral plasticity, however we found a lack of significance for relative head size predicting change in relative abundance. Fitting with the contentious debate about the function of brain size, this result has two primary interpretations. First, one interpretation is that behavioral plasticity as measured by an anatomical proxy is not as important in determining bumble bee population dynamics as it is in other taxa. Indeed, the comparative studies that have found support for a brain size proxy buffering species against environmental change have all been conducted on vertebrates (Sol et al. 2008; Sol et al. 2002; Amiel et al. 2011). A handful of studies in insects suggest that increased

cognitive ability may not be required for success in a novel environment when other traits are present, such as aggression or increased fecundity (e.g. Couvillon et al. 2010; Foucaud et al. 2016). A second interpretation is that relative brain size is not as predictive of behavioral plasticity in social insects as it is in other taxa. Nearly all of the literature on the relationship between relative brain size and behavioral plasticity comes from vertebrate studies. Increases in the size of specific brain components in insects, such as mushroom bodies, which are highly involved in learning and memory, are also known to be associated with an increase in behavioral complexity (Farris & Roberts 2005; Ehmer et al. 2001; Farris & Schulmeister 2011; Julian & Gronenberg 2002) and an increase in brain size overall (Ott & Rogers 2010). However, direct tests of connection between relative brain size and behavioral plasticity across insect species are lacking. One reason for this is that when looking across insect genera, head width may not be a reliable comparative proxy for brain volume due to differences in head morphology across insect taxa (e.g. mandibular structure, eye size and shape). This is less likely to be a factor in female bumble bees because of their similarity in traits likely to affect head width measurements (i.e. eye size and shape where interspecific variation is primarily found in males but not females) (Williams et al. 2014). Importantly, significant positive correlations between head width and brain size have been found for bumble bees (Mares et al. 2005; Riveros & Gronenberg 2010) and across other Hymenopteran species [i.e. paper wasps (Gronenberg et al. 2008; O'Donnell et al. 2018), leaf-cutting ants (Groh et al. 2014), honey bees (Gronenberg & Couvillon 2010)]. These interpretations of our result of a non-significant trend between relative head size and change in relative abundance should be considered in tandem and highlight two areas ripe for future

research: comparative insect neuroanatomy using detailed measurements of brain components and direct tests of whether insect behavioral plasticity is predictive of extinction risk.

In a notable study, Sol et al. (2008) found that mammalian species with larger relative brain sizes had an increased likelihood of establishment success. However, when using the same data with a model fitted to include both adult body mass variation and relative brain size, González-Suárez et al. (2015) found that the significant association found by Sol et al. disappeared while a significant positive association between intraspecific variation in adult body mass and establishment success appeared. Consequently, González-Suárez et al. concluded that intraspecific variation in body mass better captures the flexibility of mammalian populations to successfully respond to environmental changes than does the plasticity of a population's individual constituents. Our results may similarly suggest that intraspecific variation in body size better captures the flexibility of bumble bees to respond to environmental changes than does individual behavioral plasticity. If our data reflect differences between species in within-colony worker size variation, this flexibility conferred by size variation might be mediated by resource partitioning within colonies, even if the behavioral plasticity of individual workers is relatively low. Body size influences which floral species a bee forages from (Peat et al. 2005) and bumble bees are known to show high floral constancy (i.e. make consecutive visits to one floral species; Chittka et al. 1999). Thus, a colony with high worker size variation may be able to decrease competition for floral resources while simultaneously increasing the variety of floral species the colony has access to. The positive correlation we found between intraspecific body size and change in relative

abundance may reflect the importance that this allocation of workers to different floral species has to population stability.

Determining the correlates of bumble bee decline is needed to conserve these ecologically and economically important species. This study adds to the growing base of knowledge on the traits that may influence the susceptibility of bumble bee populations to decline. While this base of knowledge has revealed dominant drivers of decline (Williams et al. 2009; Cameron et al. 2011b; McArt et al. 2017; Siviter et al. 2018; Kerr et al. 2015; Kosior et al. 2007), how these drivers interact with one another and differentially affect species is needed for the successful development of bumble bee conservation programs. It is clear from numerous studies that interspecific trait differences between bumble bees significantly influence how susceptible species are to decline (Williams et al. 2009; Cameron et al. 2011a; Arbetman et al. 2017), however these studies have often focused on mean trait values while overlooking the potential influence of variation within these traits. A next step for the field of bumble bee conservation is to address the potential role that intraspecific trait variation has in influencing population dynamics. Efforts must also be made to move beyond generalizing traits of one bumble bee species as representative of all bumble bee taxa. Interspecific comparisons of behavior would be particularly valuable as behavior has considerable ecological consequences (Sih et al. 2011; Snell-Rood 2013) and critically affects colony function in eusocial insects (Jandt et al. 2013; Jandt & Gordon 2016). Such comparisons should be made among a broad number of species, while analyzing how such behavioral differences may be consequential at the colony- and population-levels. The conservation of biodiversity is among the greatest challenges faced by modern-day biologists. Given

the integral functional role bumble bee pollination plays in a variety of communities, the effective conservation of bumble bees will significantly aid the promotion of this biodiversity.

Acknowledgements

We thank the editors and reviewers at *The American Naturalist* for thoughtful comments that improved this manuscript's quality. We thank the Smithsonian Institution's National Museum of Natural History, American Museum of Natural History, Field Museum of Natural History, and Illinois Natural History Survey for access to their entomological collections. We acknowledge various members of Dr. Aimee Dunlap's lab for assistance throughout this study's conduction. Finally, our gratitude is extended to the many citizen scientists and professionals who have collectively contributed specimens to natural history collections, which made this work possible. This work was supported by a Whitney R. Harris World Ecology Center grant to M.W.A.

References

Amiel, J. J., R. Tingley & R. Shine. 2011. Smart moves: effects of relative brain size on establishment success of invasive amphibians and reptiles. PLoS ONE 6: e18277.

Arbetman, M. P., G. Gleiser, C. L. Morales, P. Williams & M. A. Aizen. 2017. Global decline of bumblebees is phylogenetically structured and inversely related to species range size and pathogen incidence. Proceedings of the Royal Society B: Biological Sciences, DOI: 10.1098/rspb.2017.0204.

Austin, M. W. & A. S. Dunlap. 2019. Data from: Intraspecific variation in worker body size makes North American bumble bees (*Bombus* spp.) less susceptible to decline. American Naturalist, Dryad Digital Repository, https://doi.org/10.5061/dryad.910105r.

Bartomeus, I., J. S. Ascher, J. Gibbs, B. N. Danforth, D. L. Wagner, S. M. Hedtke & R. Winfree. 2013. Historical changes in northeastern US bee pollinators related to shared ecological traits. Proceedings of the National Academy of Sciences of the United States of America 110: 4656-4660.

Blomberg, S. P., T. Garland & A. R. Ives. 2003. Testing for phylogenetic signal in comparative data: behavioral traits are more labile. Evolution 57: 717-745.

Bolnick, D. I., P. Amarasekare, M. S. Araújo, R. Bürger, J. M. Levine, M. Novak, V. H. W. Rudolf, S. J. Schreiber, M. C. Urban & D. Vasseur. 2011. Why intraspecific trait variation matters in community ecology. Trends in Ecology and Evolution 26: 183-192.

Cameron, S. A., H. M. Hines & P. H. Williams. 2007. A comprehensive phylogeny of the bumble bees (*Bombus*). Biological Journal of the Linnean Society 91: 161-188.

Cameron, S. A., J. D. Lozier, J. P. Strange, J. B. Koch, N. Cordes, L. F. Solter & T. L. Griswold. 2011a. Patterns of widespread decline in North American bumble bees. Proceedings of the National Academy of Sciences of the United States of America 108: 662-667.

Cameron, S., S. Jepsen, E. Spevak, J. Strange, M. Vaughan, J. Engler & O. Byers, editors. 2011b. North American Bumble Bee Species Conservation Planning

Workshop Final Report. IUCN/SSC Conservation Breeding Specialist Group: Apple Valley, Minnesota, USA.

Carrete, M. & J. L. Tella. 2011. Inter-individual variability in fear of humans and relative brain size of the species are related to contemporary urban invasion in birds. PLoS ONE 6: e18859.

Chittka, L. & J. Niven. 2009. Are bigger brains better? Current Biology 19: R995-R1008.

Chittka, L., Thomson, J. D. & Waser, N. M. 1999. Flower constancy, insect psychology, and plant evolution. Naturwissenschaften 86: 361-377.

Chown, S. L. & K. J. Gaston. 2010. Body size variation in insects: a macroecological perspective. Biological Reviews 85: 139-169.

Cnaani, J. & A. Hefetz. 1994. The effect of workers size frequency distribution on colony development in *Bombus terrestris*. Insectes Sociaux 41: 301-307.

Colla, S. R., F. Gadallah, L. Richardson, D. Wagner & L. Gall. 2012. Assessing declines of North American bumble bees (*Bombus* spp.) using museum specimens. Biodiversity and Conservation 21: 3585-3595.

Couvillon, M. J. & A. Dornhaus. 2009. Location, location, location: larvae position inside the nest is correlated with adult body size in worker bumble-bees (*Bombus impatiens*). Proceedings of the Royal Society B: Biological Sciences DOI: 10.1098/rspb.2009.0172.

Couvillon, M. J. & A. Dornhaus. 2010. Small worker bumble bees (*Bombus impatiens*) are hardier against starvation than their larger sisters. Insectes Sociaux 57: 193-197.

Couvillon, M. J., G. DeGrandi-Hoffman & W. Gronenberg. 2010. Africanized honeybees are slower learners than their European counterparts. Naturwissenschaften 97: 153-160.

Darwin, C. 1859. On the origin of species by means of natural selection, or preservation of favoured races in the struggle for life.

Davidson, A. M., M. Jennions & A. B. Nicotra. 2011. Do invasive species show higher phenotypic plasticity than native species and, if so, is it adaptive? A meta-analysis. Ecology Letters 14: 419-431.

Dunbar, R. I. M. & S. Shultz. 2007. Evolution in the social brain. Science 317: 1344-1347.

Ehmer, B., H. K. Reeve & R. R. Hoy. 2001. Comparison of brain volumes between single and multiple foundresses in the paper wasp *Polistes dominulus*. Brain, Behavior and Evolution 57: 161-168.

Farris, S. M. & N. S. Roberts. 2005. Coevolution of generalist feeding ecologies and gyrencephalic mushroom bodies in insects. Proceedings of the National Academy of Sciences of the United States of America 102: 17394-17399.

Farris, S. M. & S. Schulmeister. 2011. Parasitoidism, not sociality, is associated with the evolution of elaborate mushroom bodies in the brains of hymenopteran insects. Proceedings of the Royal Society B: Biological Sciences 278: 940-951.

Filin, I. & O. Ovadia. 2007. Individual size variation and population stability in a seasonal environment: a discrete-time model and its calibration using grasshoppers. The American Naturalist 170: 719-733.

Forsam, A. 2014. Effects of genotypic and phenotypic variation on establishment are important for conservation, invasion, and infection biology. Proceedings of the National Academy of Sciences of the United States of America 111: 302-307.

Forsam, A., L. Wennersten, M. Karlsson & S. Caesar. 2012. Variation in founder groups promotes establishment success in the wild. Proceedings of the Royal Society B: Biological Sciences 279: 2800-2806.

Foucaud, J., C. Morena, M. Pascual, E. L. Rezende, L. E. Castañeda, P. Gibert & F. Mery. 2016. Introduced *Drosophila subobscura* populations perform better than native populations during an oviposition choice task due to increased fecundity but similar learning ability. Ecology and Evolution 6: 1725-1736.

González-Suárez, M. & E. Revilla. 2013. Variability in life‐history and ecological traits is a buffer against extinction in mammals. Ecology Letters 16: 242-251.

González-Suárez, M., S. Bacher & J. M. Jeschke. 2015. Intraspecific trait variation is correlated with establishment success of alien mammals. The American Naturalist 185: 737-746.

Goulson, D. 2010. Bumble bees: Behaviour, Ecology, and Conservation. Oxford University Press Inc., New York, New York, USA.

Grimbacher, P. S., C. P. Catterall & R. L. Kitching. 2008. Detecting the effects of environmental change above the species level with beetles in a fragmented tropical rainforest landscape. Ecological Entomology 33: 66-79.

Groh, C., C. Kelber, K. Grübel & W. Rössler. 2014. Density of mushroom body synaptic complexes limits intraspecies brain miniaturization in highly polymorphic leaf-

cutting ant workers. Proceedings of the Royal Society B: Biological Sciences 281: 20140432.

Gronenberg, W. & M. J. Couvillon. 2010. Brain composition and olfactory learning in honey bees. Neurobiology of Learning and Memory 93: 435-443.

Gronenberg, W., L. E. Ash & E. A. Tibbetts. 2008. Correlation between facial patter recognition and brain composition in paper wasps. Brain, Behavior and Evolution 71: 1-14.

Harmon, L. J., J. T. Weir, C. D. Brock, R. E. Glor & W. Challenger. 2008. GEIGER: investigating evolutionary radiations. Bioinformatics 24: 129-131.

Hatfield, R., S. Colla, S. Jepsen, L. Richardson, R. Thorp & S. F. Jordan. 2014. IUCN assessments for North American *Bombus* spp. http://www.xerces.org/

Herrmann, J. D., N. M. Haddad & D. J. Levey. 2018. Mean body size predicts colony performance in the common eastern bumble bee (*Bombus impatiens*). Ecological Entomology DOI: 10.1111/een.12517.

Hines, H. M. 2008. Historical biogeography, divergence times, and diversification patterns of bumble bees (Hymenoptera: Apidae: *Bombus*). Systematic Biology 57: 58-75.

Imura, D., Y. Toquenaga & K. Fujii. 2003. Genetic variation can promote system persistence in an experimental host-parasitoid system. Population Ecology 45: 205-212.

IUCN Standard and Petitions Subcommittee. 2017. Guidelines for using the IUCN Red List Categories and Criteria. Version 13. Prepared by the Standards and Petitions Subcommittee.

Jackson, S. T. & D. F. Sax. 2010. Balancing biodiversity in a changing environment: extinction debt, immigration credit and species turnover. Trends in Ecology and Evolution 25: 153-160.

Jandt, J. M. & A. Dornhaus. 2014. Bumblebee response thresholds and body size: does worker diversity increase colony performance? Animal Behaviour 87: 97-106.

Jandt, J. M. & D. M. Gordon. 2016. The behavioral ecology of variation in social insects. Current Opinion in Insect Science 15: 40-44.

Jandt, J. M., S. Bengston, N. Pinter-Wollman, J. N. Pruitt, N. E. Raine, A. Dornhaus & A. Sih. 2013. Behavioural syndromes and social insects: personality at multiple levels. Biological Reviews 89: 48-67.

Julian, G. E. & W. Gronenberg. 2002. Reduction of brain volume correlates with behavioral changes in queen ants. Brain, Behavior and Evolution 60: 152-164.

Kerr, J. T., A. Pindar, R. Galpern, L. Packer, S. G. Potts, S. M. Roberts, P. Rasmont, O. Schweiger, S. R. Colla, L. L. Richardson, D. L. Wagner, L. F. Gall, D. S. Sikes & A. Pantoja. 2015. Climate change impacts on bumblebees converge across continents. Science 349: 177-180.

Knop, E. & N. Reusser. 2012. Jack-of-all-trades: phenotypic plasticity facilitates the invasion of an alien slug species. Proceedings of the Royal Society B: Biological Sciences DOI:10.1098/rspb.2012.1564.

Kolbe, S. E., R. Lockwood & G. Hunt. 2011. Does morphological variation buffer against extinction? A test using veneroid bivalves from the Plio-Pleistocene of Florida. Paleobiology 37: 355-368.

Koppert Biological Systems. 2018. Accessed 24 July 2018. <https://www.koppertus.com>

Kosior, A., W. Celary, P. Olenjniczak, J. Fijal, W. Król, W. Solarz & P. Płonka. 2007. The decline of the bumble bees and cuckoo bees (Hymenoptera: Apidae: Bombini) of Western and Central Europe. Oryx 41: 79-88.

Kotrschal, A., B. Rogell, A. Bundsen, B. Svensson, S. Zajitschek, I. Brännström, S. Immler, A. A. Maklakov & N. Kolm. 2013a. Artificial selection on relative brain size in the guppy reveals costs and benefits of evolving a larger brain. Current Biology 23: 168-171.

Kotrschal, A., B. Rogell, A. Bundsen, B. Svensson, S. Zajitschek, I. Brännström, S. Immler, A. A. Maklakov & N. Kolm. 2013b. The benefit of evolving a larger brain: big-brained guppies perform better in a cognitive task. Animal Behaviour 86: e4-e6.

Liow, L. H. 2007. Does versatility as measured by geographic range, bathymetric range and morphological variability contribute to taxon longevity? Global Ecology and Biogeography 16: 117-128.

Lodge, D. M. 1993. Biological invasions: lessons for ecology. Trends in Ecology and Evolution 8: 133-136.

Mares, S., L. Ash & W. Gronenberg. 2005. Brain allometry in bumblebee and honey bee workers. Brain, Behavior and Evolution 66: 50-61.

McArt, S. H., C. Urbanowicz, S. McCoshum, R. E. Irwin & L. S. Adler. 2017. Landscape predictors of pathogen prevalence and range contractions in US bumblebees.

Proceedings of the Royal Society B: Biological Sciences DOI: 10.1098/rspb.2017.2181.

Mikhalevich, I., R. Powell & C. Logan. 2017. Is behavioural flexibility evidence of cognitive complexity? How evolution can inform comparative cognition. Interface Focus DOI: 10.1098/rsfs.2016.0121.

O'Donnell, S., S. J. Bulova, M. Barrett & K. Fiocca. 2018. Size constraints and sensory adaptations affect mosaic brain evolution in paper wasps (Vespidae: Epiponini). Biological Journal of the Linnean Society 123: 302-310.

Ott, S. R. & S. M. Rogers. 2010. Gregarius desert locusts have substantially larger brains with altered proportions compared with the solitarious phase. Proceedings of the Royal Society B: Biological Sciences 277: 3087-3096.

Paradis, E. & K. Schliep. 2018. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics DOI: 10.1093/bioinformatics/bty633.

Peat, J., J. Tucker & D. Goulson. 2005. Does intraspecific size variation in bumblebees allow colonies to efficiently exploit different flowers? Ecological Entomology 30: 176-181.

Pendrel, B. A. & R. C. Plowright. 1981. Larval feeding by adult bumble bee workers (Hymenoptera: Apidae). Behavioral Ecology and Sociobiology 8: 71-76.

Pereboom, J. J. M., H. H. W. Velthuis & M. J. Duchateau. 2003. The organisation of larval feeding in bumblebees (Hymenoptera, Apidae) and its significance to caste differentiation. Insectes Sociaux 50: 127-133.

Pinheiro, J., D. Bates, S. DebRoy, D. Sarkar & R Core Team. 2018. nlme: Linear and Nonlinear Mixed Effects Models. R package version 3.1-137, https://CRAN.R-project.org/package=nlme.

Potts, S. G., J. C. Biesmeijer, C. Kremen, P. Neumann, O. Schweiger & W. E. Kunin. 2010. Global pollinator declines: trends, impacts, and drivers. Trends in Ecology and Evolution 25: 345-353.

Revell, L. J. 2012. phytools: An R package for phylogenetic comparative biology (and other things). Methods in Ecology and Evolution 3: 217-223

Riveros, A. J. & W. Gronenberg. 2010. Brain allometry and neural plasticity in the bumblebee Bombus occidentalis. Brain, Behavior and Evolution 75: 138-148.

Roth, T. C., L. D. LaDage & V. Pravosudov. 2010. Learning capabilities enhanced in harsh environments: a common garden approach. Proceedings of the Royal Society B: Biological Sciences 277: 3187-3193.

Scheper, J., M. Reemer, R. van Kats, W. A. Ozinga, G. T. J. van der Linden, J. H. J. Schaminée, H. Siepel & D. Kleijn. 2014. Museum specimens reveal loss of pollen host plants as key factor driving wild bee decline in The Netherlands. Proceedings of the National Academy of Sciences of the United States of America 111: 17552-17557.

Sih, A., M. C. O. Ferrari & D. J. Harris. 2011. Evolution and behavioural responses to human-induced rapid environmental change. Evolutionary Applications 4: 367-387.

Siviter, H., J. Koricheva, M. J. F. Brown & E. Leadbetter. 2018. Quantifying the impact of pesticides on learning and memory in bees. Journal of Applied Ecology DOI: 10.1111/1365-2664.13193.

Snell-Rood, E. C. 2013. An overview of the evolutionary causes and consequences of behavioural plasticity. Animal Behaviour 85: 1004-1011.

Snell-Rood, E. C., G. Davidowitz & D. R. Papaj. 2011. Reproductive tradeoffs of learning in a butterfly. Behavioral Ecology 22: 291-302.

Sol, D. 2009. The cognitive-buffer hypothesis for the evolution of large brains. Cognitive Ecology II. The University of Chicago Press, Chicago, Illinois, USA.

Sol, D., R. P. Duncan, T. M. Blackburn, P. Cassey & L. Lefebvre. 2005. Big brains, enhanced cognition, and response of birds to novel environments. Proceedings of the National Academy of Sciences of the United States of America 102: 5460-5465.

Sol, D., S. Bacher, S. M. Reader & L. Lefebvre. 2008. Brain size predicts the success of mammal species introduced into novel environments. The American Naturalist 172: S63-S71.

Sol, D., S. Timmermans & L. Lefebvre. 2002. Behavioural flexibility and invasion success in birds. Animal Behaviour 63: 495-502.

Spaethe, J. & A. Weidenmüller. 2002. Size variation and foraging rate in bumblebees (*Bombus terrestris*). Insectes Sociaux 49: 142-146.

Sutcliffe, G. H. & R. C. Plowright. 1988. The effects of food supply on adult size in the bumble bee *Bombus terricola* Kirby. The Canadian Entomologist 120: 1051-1058.

Tello-Ramos, M. C., C. L. Branch, D. Y. Kozlovsky, A. M. Pitera & V. V. Pravosudov. (in press). Spatial memory and cognitive flexibility trade-offs: to be or not to be flexible, that is the question. Animal Behaviour DOI: 10.1016/j.anbehav.2018.02.019.

Tuomainen, U. & U. Candolin. 2011. Behavioural responses to human-induced environmental change. Biological Reviews 86: 640–657.

Vellend, M. 2006. The consequences of genetic diversity in competitive communities. Ecology 87: 304-311.

Vincze, O. 2016. Light enough to travel or wise enough to stay? Brain size evolution and migratory behavior in birds. Evolution 70: 2123-2133.

Vindenes, Y., S. Engen & B. Sæther. 2008. Individual heterogeneity in vital parameters and demographic stochasticity. The American Naturalist 171: 455-467.

Vitousek, P. M., H. A. Mooney, J. Lubchenco & J. M. Melillo. 1997. Human's domination of Earth's ecosystems. Science 227: 494-499.

Williams, N. M., E. E. Crone, T. H. Roulston, R. L. Minckley, L. Packer & S. G. Potts. 2010. Ecological and life-history traits predict bee species responses to environmental disturbances. Biological Conservation 143: 2280-2291.

Williams, P. 2005. Does specialization explain rarity and decline among British bumblebees? A response to Goulson et al. Biological Conservation 122: 33-43.

Williams, P. H., R. Thorp, L. Richardson & S. R. Colla. 2014. Bumble Bees of North America. Princeton University Press, Princeton, New Jersey, USA.

Williams, P. H., S. A. Cameron, H. Hines, B. Cederberg & P. Rasmont. 2008. A simplified subgeneric classification of the bumblebees (genus *Bombus*). Apidologie 39: 1–29.

Williams, P., S. Colla & Z. Xie. 2009. Bumblebee vulnerability: common correlates of winners and losers across continents. Conservation Biology 23: 931-940.

Wong, B. B. M. & U. Candolin. 2015. Behavioral responses to changing environments. Behavioral Ecology 26: 665–673.

Figure 1 *Bombus* spp. picture attributions:

B. nevadensis is by Sesamehoneytart [CC BY-SA 4.0 (https://creativecommons.org/licenses/by-sa/4.0)], from Wikimedia Commons. https://commons.wikimedia.org/wiki/File:Bombus_nevadensis_080115.jpg

B. appositus is by JerryFriedman [CC BY-SA 3.0 (https://creativecommons.org/licenses/by-sa/3.0)], from Wikimedia Commons. https://commons.wikimedia.org/wiki/File:Bombus_appositus_dorsal.jpg

B. fervidus is by John Baker [CC BY 2.0 (https://creativecommons.org/licenses/by/2.0)], via Wikimedia Commons. https://commons.wikimedia.org/wiki/File:Bombus_fervidus_(_Golden_Northern_Bumble_Bee).jpg

B. griseocollis is by USFWS Mountain-Prairie [CC BY 2.0 (https://creativecommons.org/licenses/by/2.0) or Public domain], via Wikimedia Commons. https://commons.wikimedia.org/wiki/File:Bombus_griseocollis_Female_brown-

belted_bumble_bee_on_purple_prairie_clover_Sand_Lake_WMD_(12842806815).jpg

B. occidenatlis is by Stephen Ausmus, USDA ARS [Public domain], via Wikimedia Commons. https://commons.wikimedia.org/wiki/File:Bombus_occidentalis.jpg

B. impatiens is by Andrew C (Common Eastern Bumble bee (Bombus impatiens)) [CC BY 2.0 (https://creativecommons.org/licenses/by/2.0)], via Wikimedia Commons. https://commons.wikimedia.org/wiki/File:Common_Eastern_Bumble_bee_(Bombus_impatiens)_(14668688779).jpg

References Cited Only in the Appendix

Blomberg, S. P., T. Garland & A. R. Ives. 2003. Testing for phylogenetic signal in comparative data: behavioral traits are more labile. Evolution 57: 717-745.

Gérard, M., M. Vanderplanck, M. Frazen, M. Kuhlmann, S. G. Potts, P. Rasmont, O. Schweiger & D. Michez. 2018. Patterns of size variation in bees at a continental scale: does Bergmann's rule apply? Oikos 127: 1095-1103.

Ramírez-Delgado, V. H., S. Sanabria Urbán, M. A. Serrano-Meneses & R. C. del Castillo. 2016. The converse to Bergmann's rule in bumblebees, a phylogenetic approach. Ecology and Evolution 6: 6160-6169.

Revell, L. J. 2012. phytools: An R package for phylogenetic comparative biology (and other things). Methods in Ecology and Evolution 3: 217-223.

Appendix

Temporal Distribution of Body Size

We tested whether certain species have exhibited temporal trends in body size by running a linear model for each species, which regressed specimen collection year, the predictor variable, against thorax width, the response variable. The collection year was absent from the label data of 65 specimens, which were accordingly dropped from these analyses. Three of our 31 species showed a statistically significant temporal trend in body size: *Bombus bifarius* (positive correlation, increase in body size with time; $p<0.05$), *B. flavifrons* (positive correlation, increase in body size with time; $p<0.0005$), and *B. fraternus* (negative correlation, decrease in body size with time; $p<0.05$). The remaining 28 species did not show a statistically significant temporal trend in body size ($p>0.05$). When the three species that showed a significant temporal trend in body size (i.e. *B. bifarius*, *B. flavifrons*, *B. fraternus*) are removed from our phylogenetic generalized least squares (PGLS) analyses, the same PGLS results are obtained [first PGLS: thorax width CV ($p<0.005$), relative head size residuals ($p=0.575$), thorax width CV:relative head size residuals ($p=0.567$); second PGLS: thorax width CV ($p<0.05$), tongue length ($p=0.125$), thorax width CV:tongue length ($p=0.112$)]. Figures depicting these temporal trends in body size and a table of these models' full results can be found in the Supplemental Material (Table S2; Figs. S1-S31).

Spatial Distribution of Body Size

To assess the spatial distribution of the specimens we measured from natural history collections, we developed a map for each species that depicts the locations each specimen

was collected (Figs. S1-S31). Each map was made in ArcGIS 10.2.1 by plotting each specimen's coordinates against a world map using the World Geodetic System 1984 as the reference coordinate system. Coordinates in decimal degrees were obtained from specimens' label data or, when a specimen's label data lacked coordinates, from Google Earth using the approximate mid-point of the county that specimen was collected in. The coordinates of 38 specimens could not be determined using these methods and were accordingly dropped from the maps.

We tested whether certain species exhibit spatial trends in body size by running a linear model for each species, which regressed latitude at which the specimen was collected, the predictor variable, against thorax width, the response variable. The 38 specimens for which coordinates could not be determined were dropped from these analyses. Previous studies have suggested that bumble bee species follow the converse trend of Bergmann's rule (i.e. bumble bee species are larger at warmer latitudes, which are lower latitudes in the Northern Hemisphere) (Gérard et al. 2018; Ramírez-Delgado et al. 2016). Our analyses differ from these previous studies in that our analyses assess latitudinal trends in intraspecific variation, while these previous studies assessed latitudinal trends in interspecific variation. Two of our 31 species showed a statistically significant latitudinal trend in body size: *B. occidentalis* (positive correlation, following Bergmann's rule; $p<0.005$) and *B. vagans* (negative correlation, following the converse of Bergmann's rule; $p<0.05$). The remaining 29 species did not show a statistically significant latitudinal trend in body size ($p>0.05$). Figures depicting these latitudinal trends in body size and a table of these models' full results can be found in the Supplemental Material (Table S2; Figs. S1-S31).

Phylogenetic Signal

We assessed change in relative abundance, thorax width coefficients of variation, relative head size residuals, and tongue length for phylogenetic signal (i.e. the tendency for interspecific trait differences to depend on phylogeny) using Blomberg's *K* (Blomberg et al. 2003). Blomberg's *K* assesses phylogenetic signal for continuous traits using a Brownian motion model of character evolution; $K = 1$ for traits that show a statistical dependency on phylogenetic relationships, $K = 0$ for traits that are not statistically dependent on phylogeny. To assess the statistical significance of *K*, the observed *K* value is compared to simulated *K* values generated from randomized data. We used 1,000 simulations for each of these randomization tests. These phylogenetic signal analyses were performed with the PHYTOOLS package (Revell 2012) in RStudio (version 0.99.902).

Our analyses of Blomberg's *K* reveal that relative head size and tongue length each show phylogenetic signal among North American bumble bees (relative head size residuals, K=0.316, p<0.05; tongue length, K=0.711, p=0.001). Hence, each of these traits shows statistical dependency on phylogenetic relationships according to a Brownian motion model of character evolution (Blomberg et al. 2003). Blomberg's *K* did not detect phylogenetic signal for thorax width CV or change in relative abundance among these species (thorax width CV, K=0.243, p=0.156; change in relative abundance, K=0.271, p=0.081). All traits were assessed for phylogenetic signal using the tree topology pruned to all 31 species included in this study, with the exception of tongue length, for which *B.*

caliginosus, *B. crotchii*, and *B. sandersoni* were also removed. A table summarizing these phylogenetic signal results can be found in the Supplemental Material (Table S1).

Tables

Table 1. Taxonomic depiction of *Bombus* spp. and traits included in analyses.

Subgenus	Species	Sample Size	Thorax Width Coefficient of Variation	Relative Head Size Residual	Tongue Length (mm)	Change in Relative Abundance
Bombias	*nevadensis*	28	14.252	-0.159	8.995	64.08%
	auricomus	26	8.574	-0.014	10.805	50.08%
Subterraneobombus	*borealis*	32	10.770	-0.094	8.585	86.91%
	appositus	36	11.094	-0.179	10.507	46.65%
Fervidobombus	*fervidus*	35	11.792	0.055	9.679	38.04%
	pensylvanicus	41	11.981	0.187	9.679	11.44%
Cullumanobombus	*rufocinctus*	31	12.693	-0.122	5.529	154.88%
	morrisoni	31	8.956	-0.029	8.248	17.43%
	crotchii	24	8.624	0.103	-	2.32%
	griseocollis	42	9.208	-0.033	7.614	215.25%
	fraternus	34	6.193	0.188	7.434	14.40%
Bombus (sensu stricto)	*affinis*	40	10.716	0.210	6.89	7.46%
	terricola	35	8.847	0.068	6.297	19.17%
	occidentalis	33	8.003	-0.039	5.966	28.51%
Pyrobombus	*vagans*	37	14.488	0.151	8.004	108.97%
	caliginosus	28	8.201	0.035	-	15.60%
	centralis	25	7.257	-0.021	7.096	81.27%
	vandykei	26	11.099	-0.018	8.101	163.71%
	flavifrons	33	12.485	-0.045	7.396	161.79%
	melanopygus	30	11.985	-0.011	6.488	81.85%
	bimaculatus	35	10.944	-0.037	8.415	188.19%
	sylvicola	28	9.667	-0.284	5.789	96.41%
	impatiens	45	9.964	0.115	7.243	294.17%
	vosnesenskii	29	9.448	0.030	7.714	122.30%
	huntii	32	6.263	0.041	6.896	70.51%
	ternarius	27	8.720	0.020	5.9	162.21%
	bifarius	28	5.821	-0.008	5.495	126.53%
	perplexus	27	9.161	-0.010	7.463	92.19%
	mixtus	32	16.613	-0.006	5.495	263.61%
	sandersoni	21	11.533	-0.060	-	87.37%
	frigidus	26	11.470	-0.035	5.732	116.34%

Note: Species are arranged by phylogeny (Fig. 2) and grouped by subgenera according to the most recent taxonomic revisions in Williams et al. (2008).

Table 2. Phylogenetic generalized least squares (PGLS) results.

PGLS 1				
Effect	Value	Standard Error	*t*-value	*p*-value
Intercept	-1.108	0.870	-1.273	0.214
Thorax Width CV	0.168	0.045	3.761	*<0.001**
Relative Head Size Residual	-3.235	5.505	-0.588	0.562
Thorax Width CV x Relative Head Size Residual	0.286	0.507	0.565	0.577
PGLS 2				
Effect	Value	Standard Error	*t*-value	*p*-value
Intercept	-3.488	2.196	-1.589	0.125
Thorax Width CV	0.430	0.207	2.073	*<0.05**
Tongue Length	0.411	0.283	1.450	0.160
Thorax Width CV x Tongue Length	-0.043	0.028	-1.527	0.140

Note: PGLS 1 included thorax width coefficients of variation (CV) and relative head size residuals as predictor variables. PGLS 2 included thorax width CV and tongue length as predictor variables. Both models included change in relative abundance as the response variable. Significant *p*-values indicated in italic with an asterisk (*).

Figures

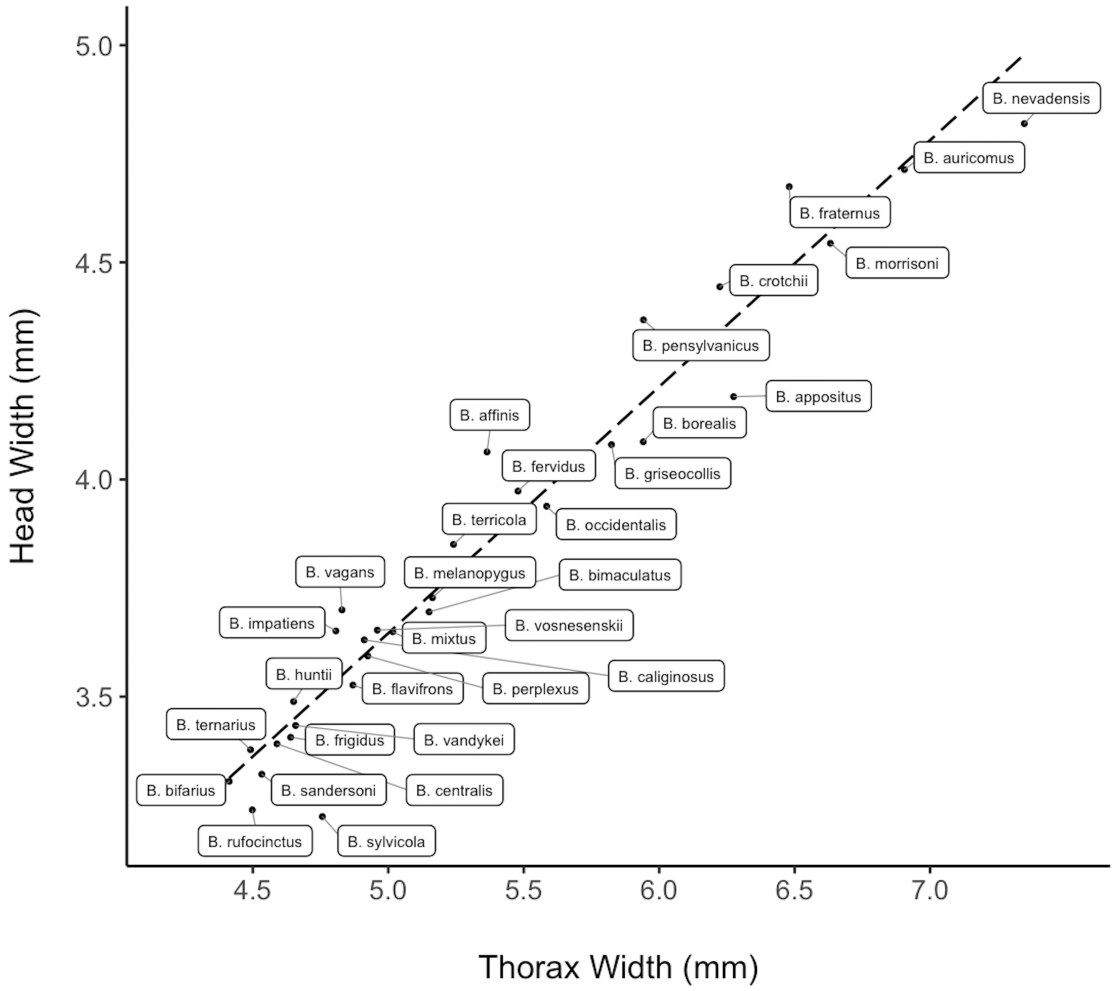

Fig 1. Correlation between average head width and average thorax width per species, used for calculating relative head size residuals. When compared to all *Bombus* species included in this regression, species above the best-fit line have a larger head width than expected for their thorax width, and species below the best-fit line have a smaller head width than expected for their thorax width.

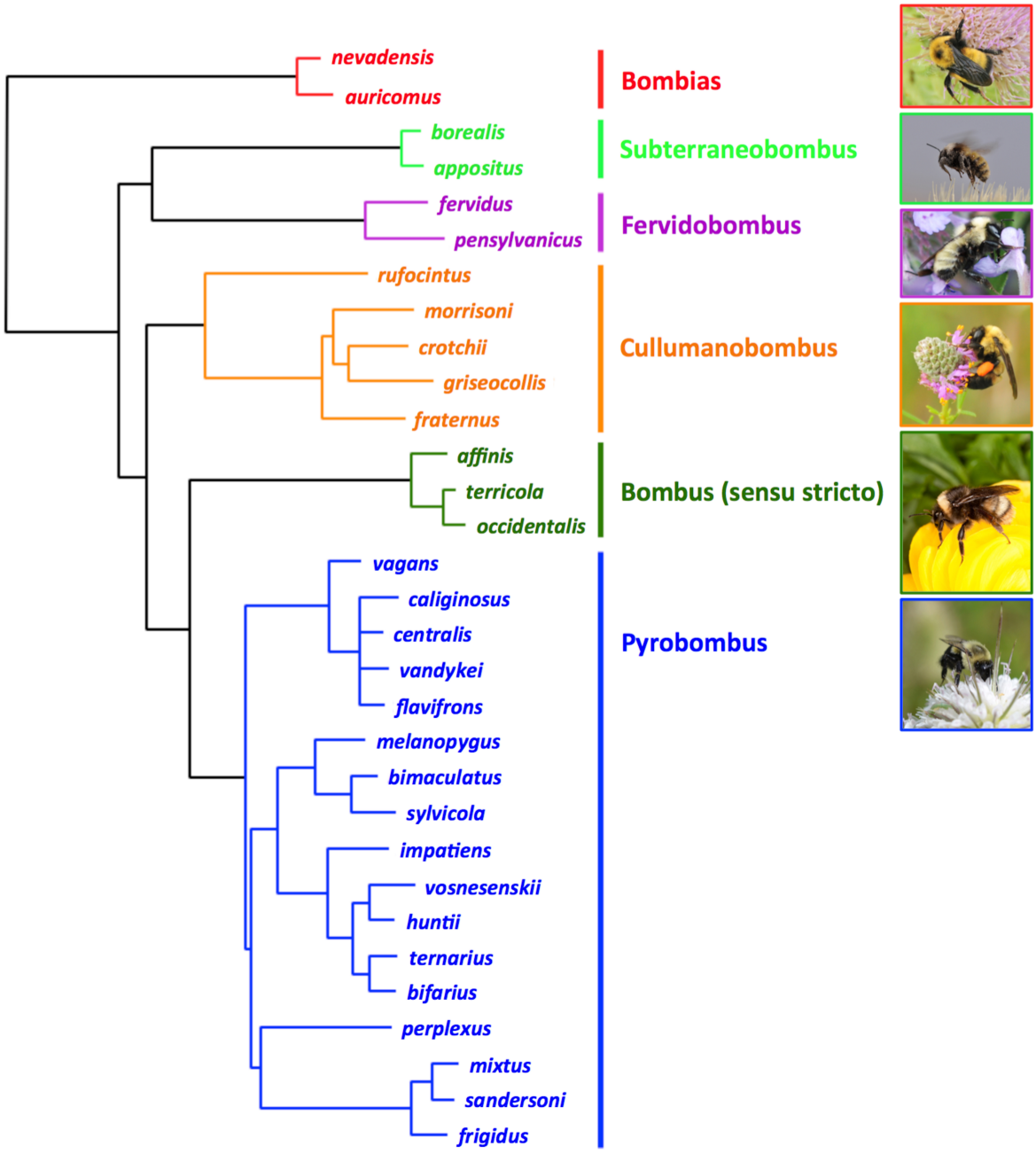

Fig 2. Pruned bumble bee phylogeny used for analyses, adapted from Cameron et al. (2007) and Hines (2008). Subgenera are denoted by color according to the most recent taxonomic revisions in Williams et al. (2008). *Bombus* pictures are from Wikimedia Commons and attributed under the references section. From top to bottom, the species

depicted in each picture are: *B. nevadensis*, *B. appositus*, *B. fervidus*, *B. griseocollis*, *B. occidentalis*, and *B. impatiens*.

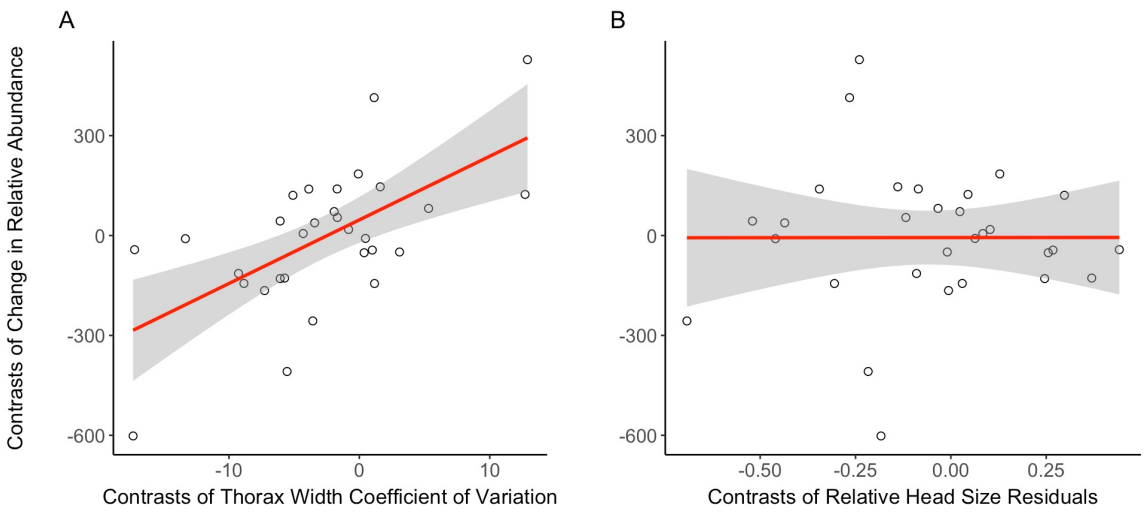

Fig 3. Correlations between change in relative abundance with (A) body size variation (p<0.001) and (B) relative head size (p=0.562). Gray areas are 95% confidence intervals. These correlations depict the phylogenetically controlled relationships between these traits with phylogenetically independent contrasts.

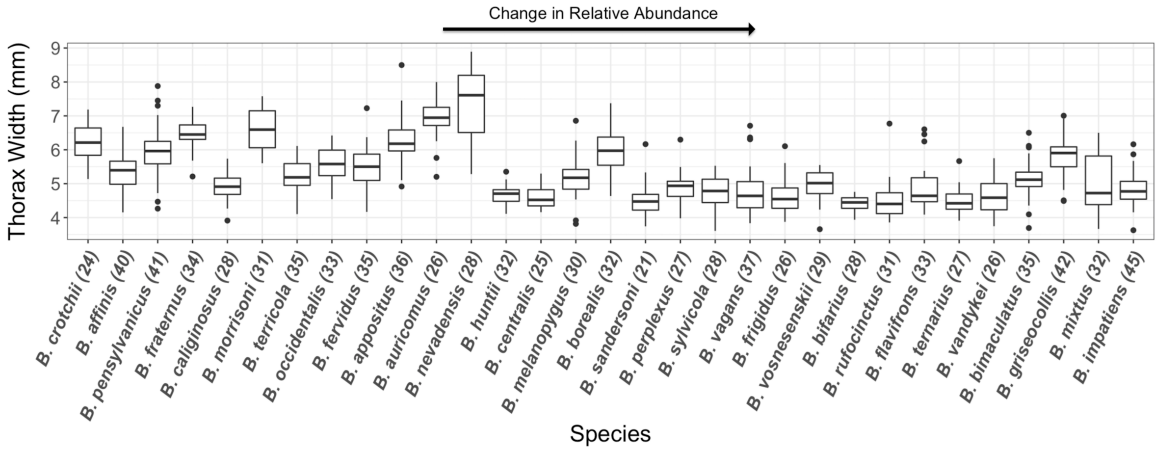

Fig 4. Box plots of thorax width measurements obtained per species. Sample sizes are listed in parentheses. Species are arranged from left to right in order of increasing change in relative abundance.

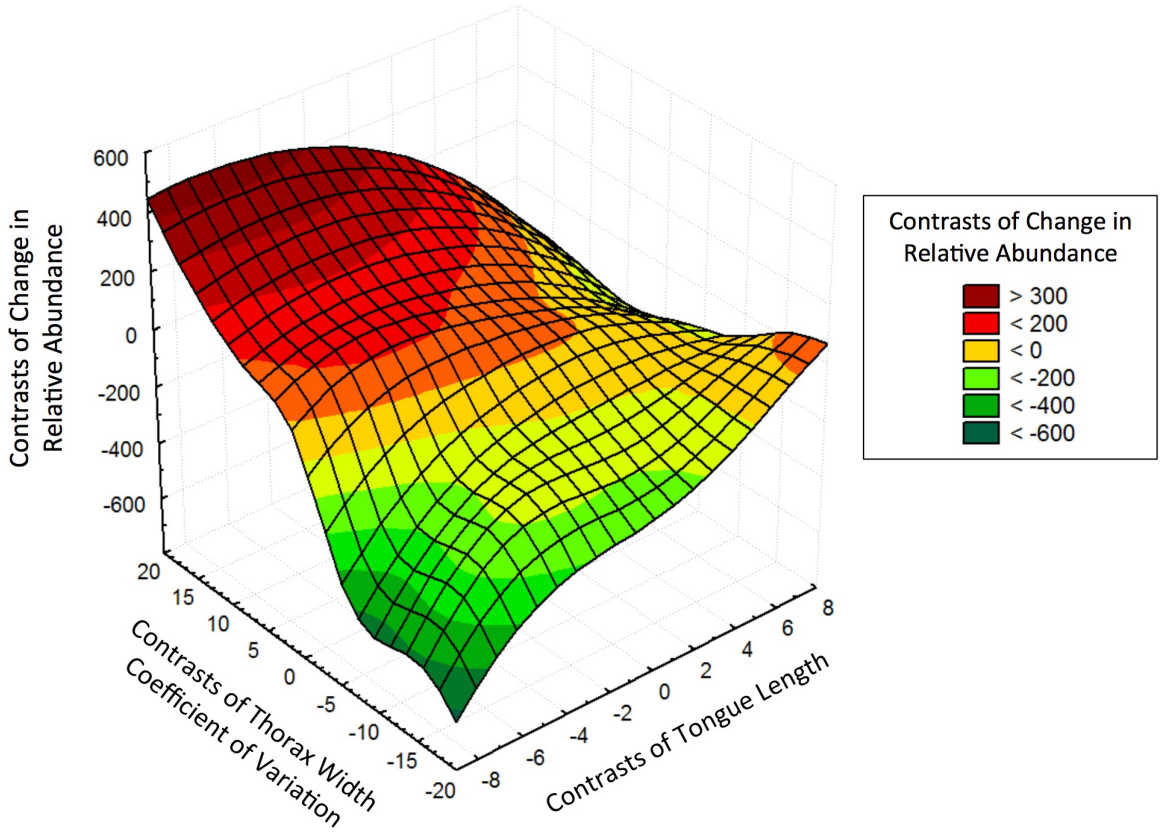

Fig 5. Three-dimensional surface plot (distance-weighted least squares fitting) depicting how thorax width coefficient of variation and tongue length interact to affect change in relative abundance (p=0.140). Surface color indicates values of contrasts of change in relative abundance. This surface plot depicts the phylogenetically controlled relationships between these traits with phylogenetically independent contrasts.

Supplemental Tables and Figures

Table of contents:

Table S1. Phylogenetic signal (Blomberg's *K*) results. As described in the appendix, we assessed change in relative abundance, thorax width coefficients of variation, relative head size residuals, and tongue length for phylogenetic signal using Blomberg's *K*. Significant *p*-values indicated in italic with an asterisk (*).

Trait	K	*p*-value
Change in Relative Abundance	0.271	0.081
Thorax Width Coefficient of Variation	0.243	0.156
Relative Head Size Residual	0.316	*<0.05**
Tongue Length	0.711	*0.001**

Table S2. Linear model results of temporal and latitudinal trends in body size. Temporal trends are regressions of collecting date year against thorax width. Latitudinal trends are regressions of latitude against thorax width. Significant *p*-values indicated in italic with an asterisk (*).

	Temporal Trend				Latitudinal Trend			
Species (*Bombus* spp.)	**Estimate**	**Standard Error**	***t*-value**	***p*-value**	**Estimate**	**Standard Error**	***t*-value**	***p*-value**
affinis	-0.002	0.006	-0.426	0.67	-0.119	0.068	-1.741	0.09
appositus	0.017	0.012	1.464	0.15	0.016	0.037	0.420	0.68
auricomus	0.005	0.003	1.960	0.06	0.018	0.084	0.208	0.84
bifarius	0.005	0.002	2.333	*<0.05**	-0.012	0.012	-1.034	0.31
bimaculatus	0.004	0.004	1.009	0.32	-0.084	0.042	-1.980	0.06
borealis	0.008	0.008	0.990	0.33	0.070	0.048	1.456	0.16
caliginosus	0.004	0.003	1.338	0.19	-0.007	0.026	-0.285	0.78
centralis	0.000	0.004	-0.039	0.97	-0.012	0.021	-0.571	0.57
crotchii	-0.007	0.004	-1.707	0.10	-0.074	0.043	-1.741	0.10
fervidus	-0.002	0.005	-0.434	0.67	0.002	0.036	0.056	0.96
flavifrons	0.029	0.007	3.969	*<0.0005**	0.002	0.016	0.101	0.92
fraternus	-0.007	0.003	-2.358	*<0.05**	0.017	0.018	0.962	0.34
frigidus	0.002	0.006	0.420	0.68	-0.017	0.014	-1.273	0.22
griseocollis	0.006	0.004	1.587	0.12	0.006	0.029	0.189	0.85
huntii	-0.004	0.003	-1.571	0.13	0.012	0.016	0.731	0.47
impatiens	0.005	0.004	1.322	0.19	-0.018	0.019	-0.951	0.35
melanopygus	0.006	0.004	1.472	0.15	0.006	0.022	0.283	0.78
mixtus	-0.018	0.015	-1.172	0.25	0.074	0.039	1.901	0.07
morrisoni	-0.006	0.008	-0.666	0.51	-0.027	0.034	-0.796	0.43
nevadensis	0.015	0.015	0.959	0.35	0.039	0.055	0.701	0.49
occidentalis	-0.009	0.007	-1.250	0.22	0.054	0.016	3.483	*<0.005**
pensylvanicus	0.004	0.005	0.771	0.45	-0.037	0.019	-1.962	0.06
perplexus	0.003	0.003	0.907	0.37	0.012	0.015	0.782	0.44
rufocinctus	0.005	0.004	1.085	0.29	-0.027	0.029	-0.914	0.37
sandersoni	-0.002	0.004	-0.466	0.65	-0.034	0.020	-1.641	0.12
sylvicola	-0.002	0.003	-0.759	0.45	-0.001	0.008	-0.116	0.91
ternarius	0.002	0.003	0.598	0.56	-0.012	0.032	-0.363	0.72
terricola	-0.006	0.004	-1.412	0.17	0.003	0.034	0.089	0.93
vagans	0.001	0.006	0.251	0.80	-0.085	0.041	-2.037	*<0.05**
vandykei	0.006	0.004	1.530	0.14	0.009	0.023	0.396	0.70
vosnesenskii	0.002	0.004	0.508	0.62	-0.002	0.022	-0.073	0.94

Species: *Bombus affinis*

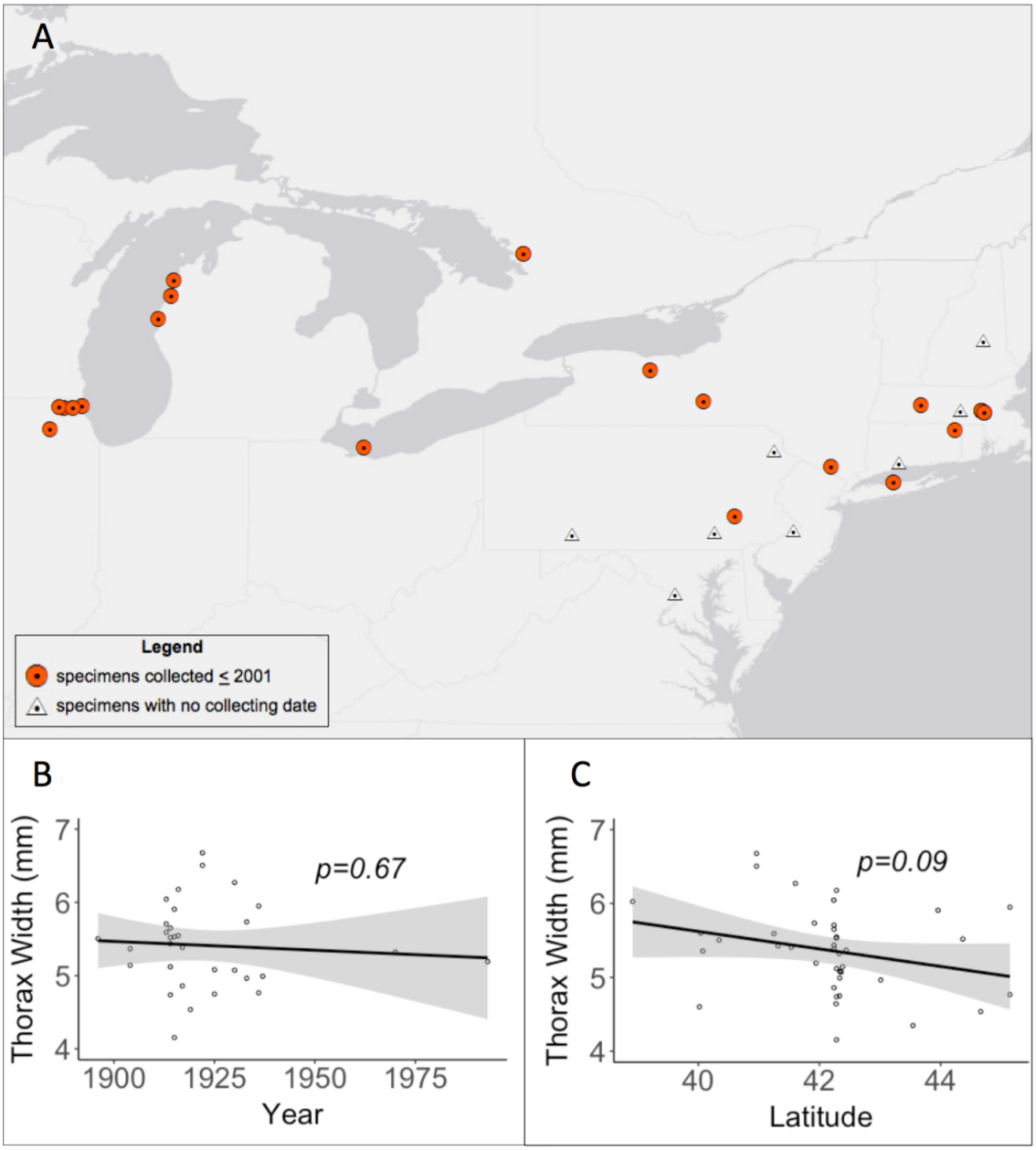

Fig S1. Spatial and temporal distribution of *Bombus affinis* specimens used in analyses. A) Map of *Bombus affinis* specimens' collecting locations. B) Temporal trend of *Bombus affinis* body size. C) Latitudinal trend of *Bombus affinis* body size.

Species: *Bombus appositus*

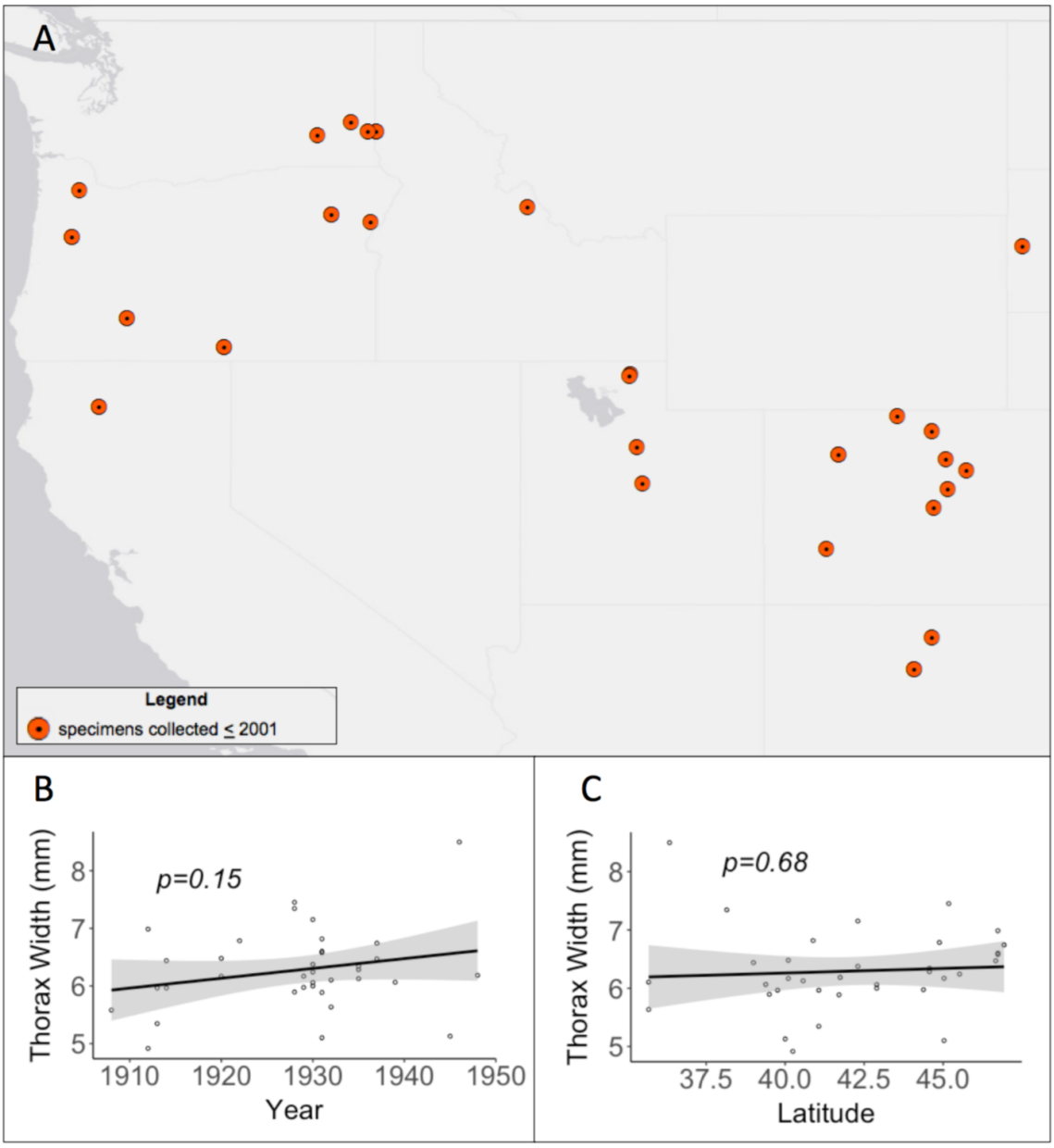

Fig S2. Spatial and temporal distribution of *Bombus appositus* specimens used in analyses. A) Map of *Bombus appositus* specimens' collecting locations. B) Temporal trend of *Bombus appositus* body size. C) Latitudinal trend of *Bombus appositus* body size.

Species: *Bombus auricomus*

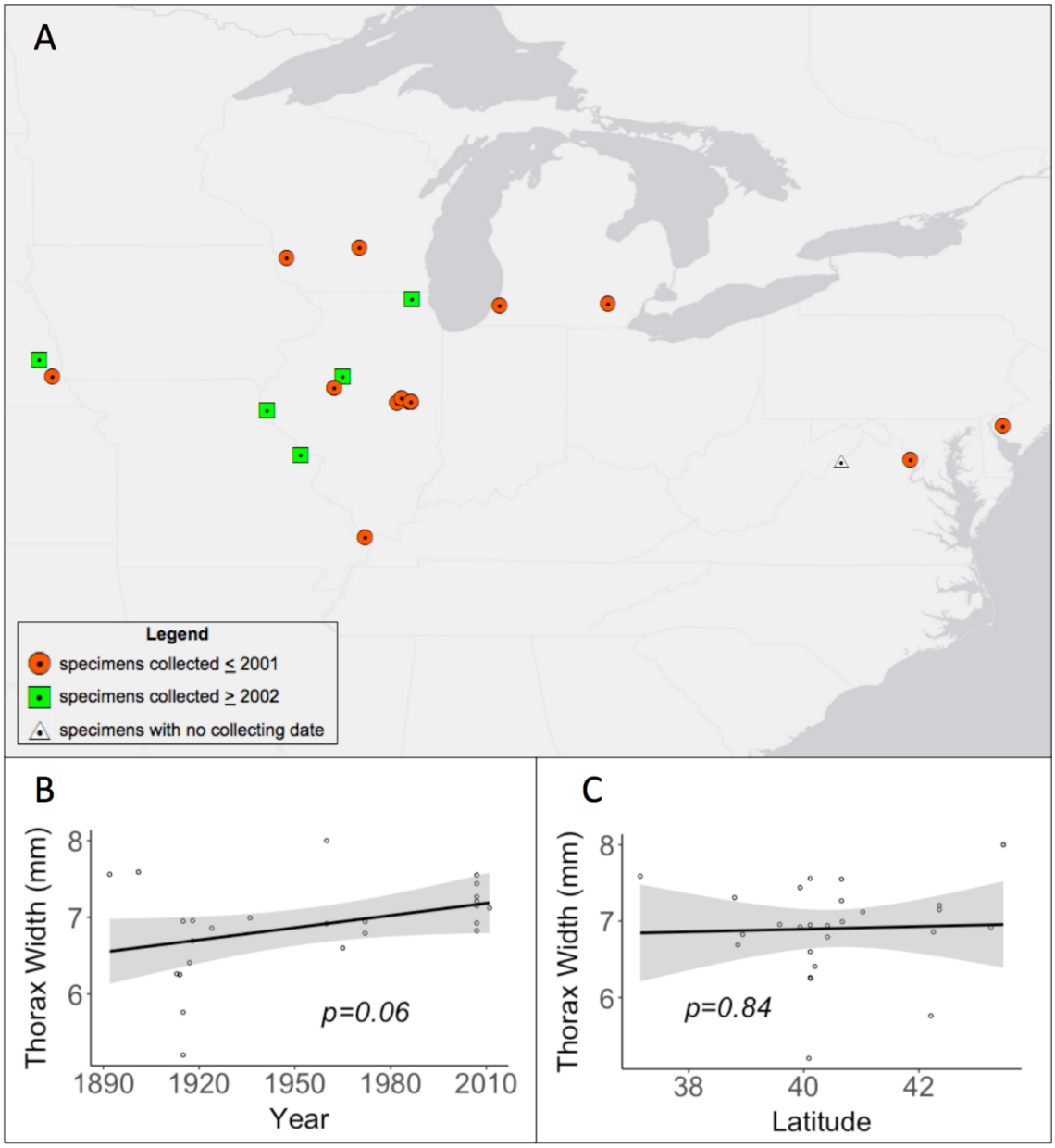

Fig S3. Spatial and temporal distribution of *Bombus auricomus* specimens used in analyses. A) Map of *Bombus auricomus* specimens' collecting locations. B) Temporal trend of *Bombus auricomus* body size. C) Latitudinal trend of *Bombus auricomus* body size.

Species: *Bombus bifarius*

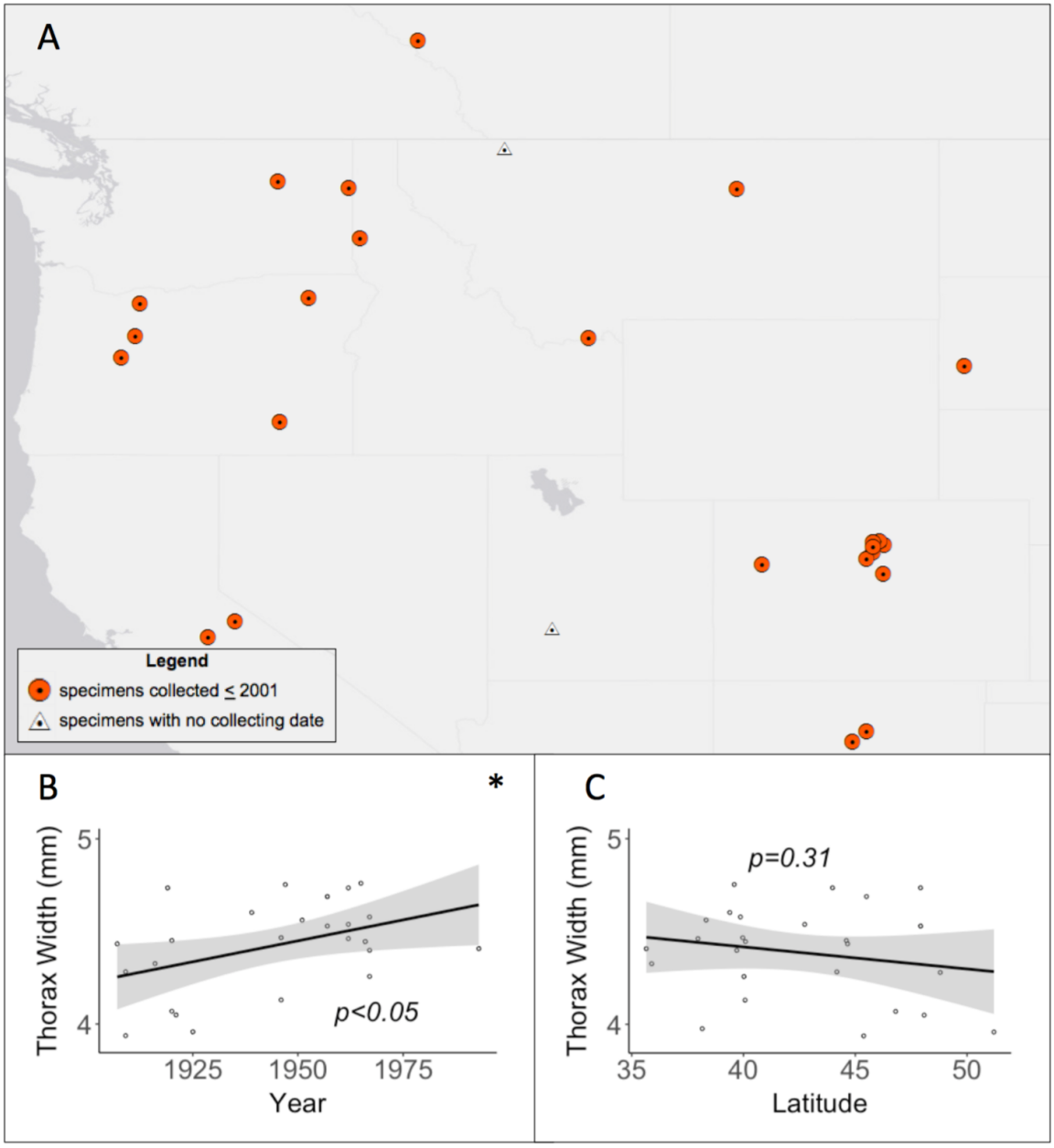

Fig S4. Spatial and temporal distribution of *Bombus bifarius* specimens used in analyses. A) Map of *Bombus bifarius* specimens' collecting locations. B) Temporal trend of *Bombus bifarius* body size. C) Latitudinal trend of *Bombus bifarius* body size. Asterisk (*) denotes statistical significance.

Species: *Bombus bimaculatus*

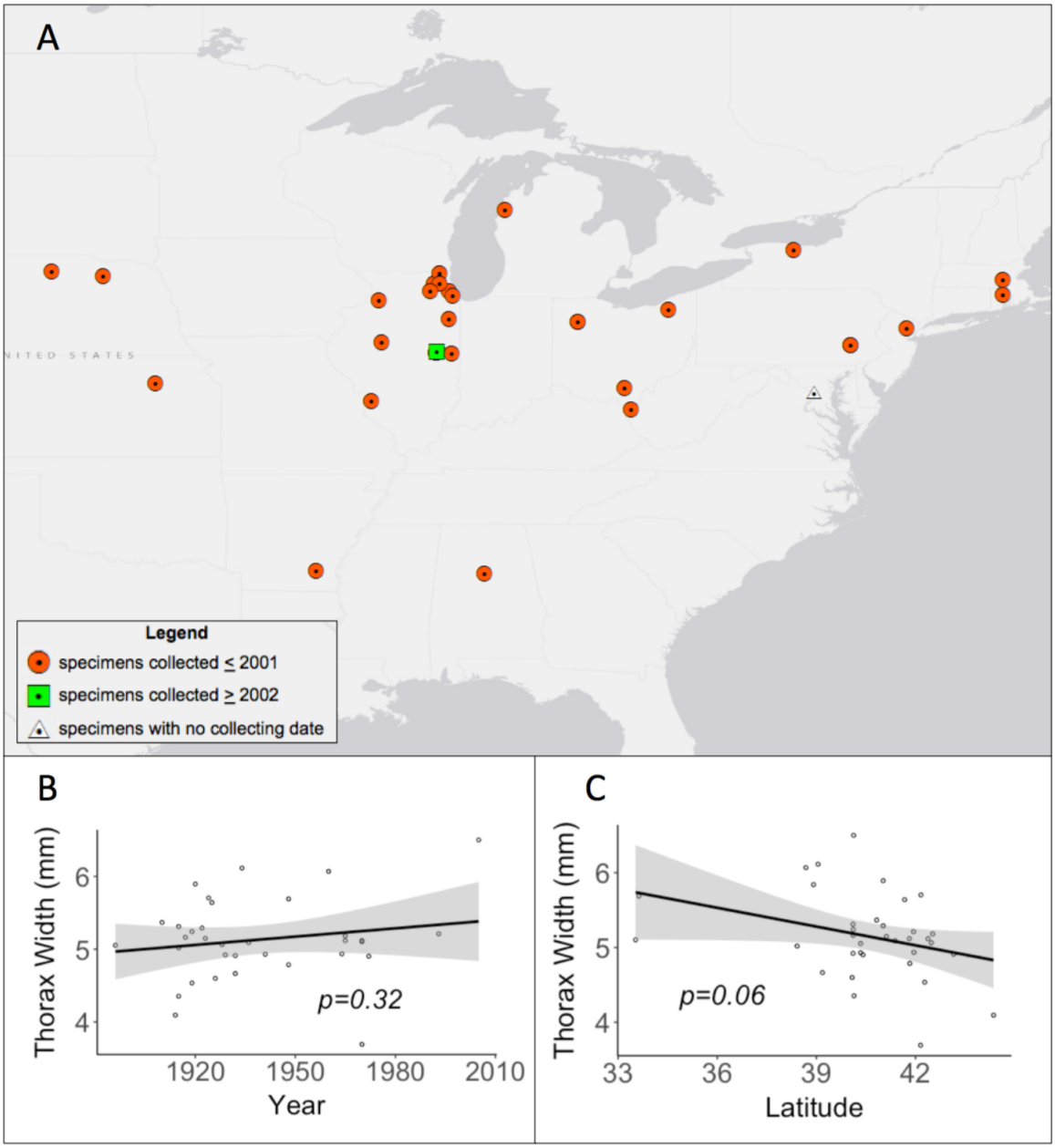

Fig S5. Spatial and temporal distribution of *Bombus bimaculatus* specimens used in analyses. A) Map of *Bombus bimaculatus* specimens' collecting locations. B) Temporal trend of *Bombus bimaculatus* body size. C) Latitudinal trend of *Bombus bimaculatus* body size.

Species: *Bombus borealis*

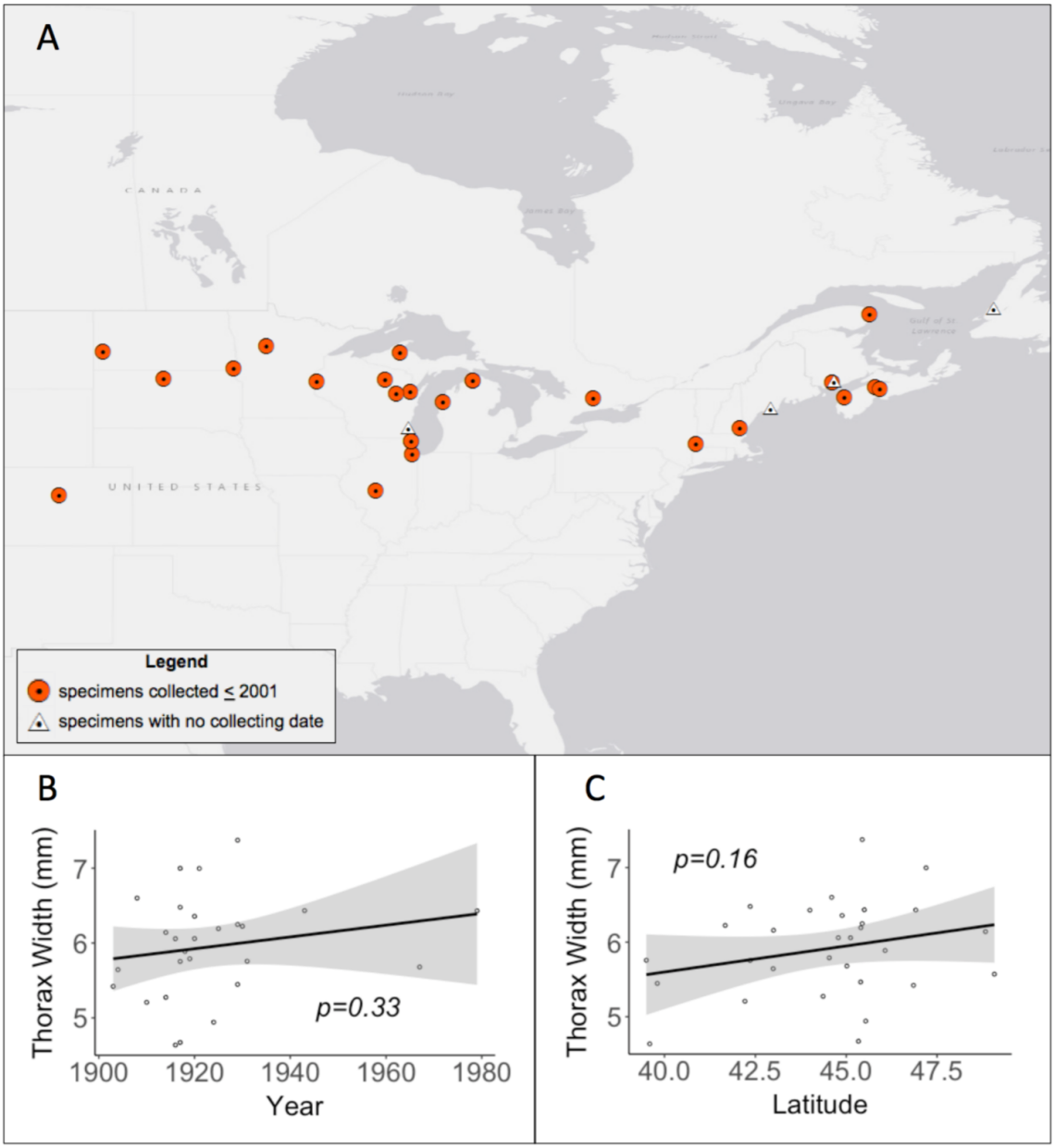

Fig S6. Spatial and temporal distribution of *Bombus borealis* specimens used in analyses. A) Map of *Bombus borealis* specimens' collecting locations. B) Temporal trend of *Bombus borealis* body size. C) Latitudinal trend of *Bombus borealis* body size.

Species: *Bombus caliginosus*

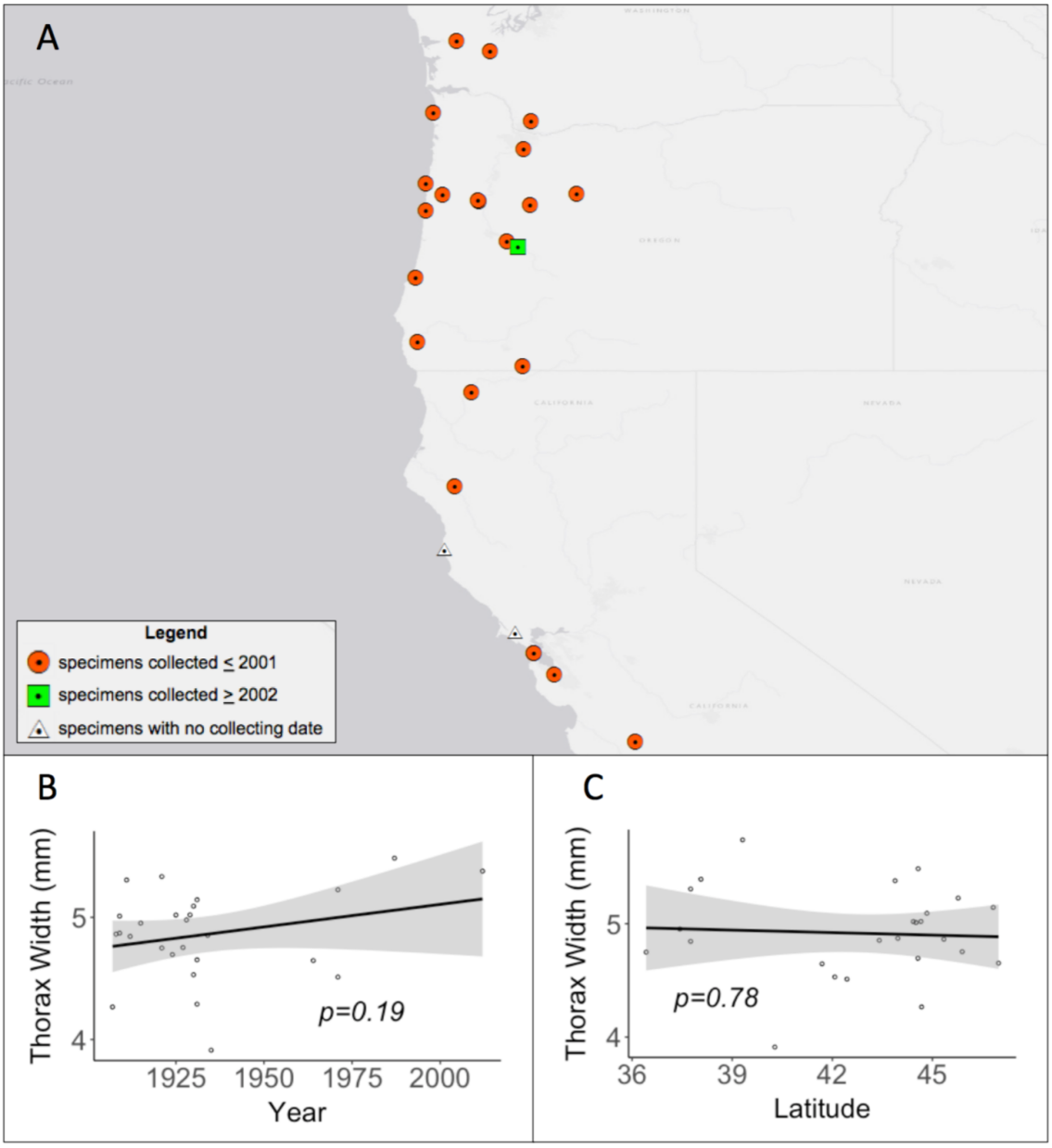

Fig S7. Spatial and temporal distribution of *Bombus caliginosus* specimens used in analyses. A) Map of *Bombus caliginosus* specimens' collecting locations. B) Temporal trend of *Bombus caliginosus* body size. C) Latitudinal trend of *Bombus caliginosus* body size.

Species: *Bombus centralis*

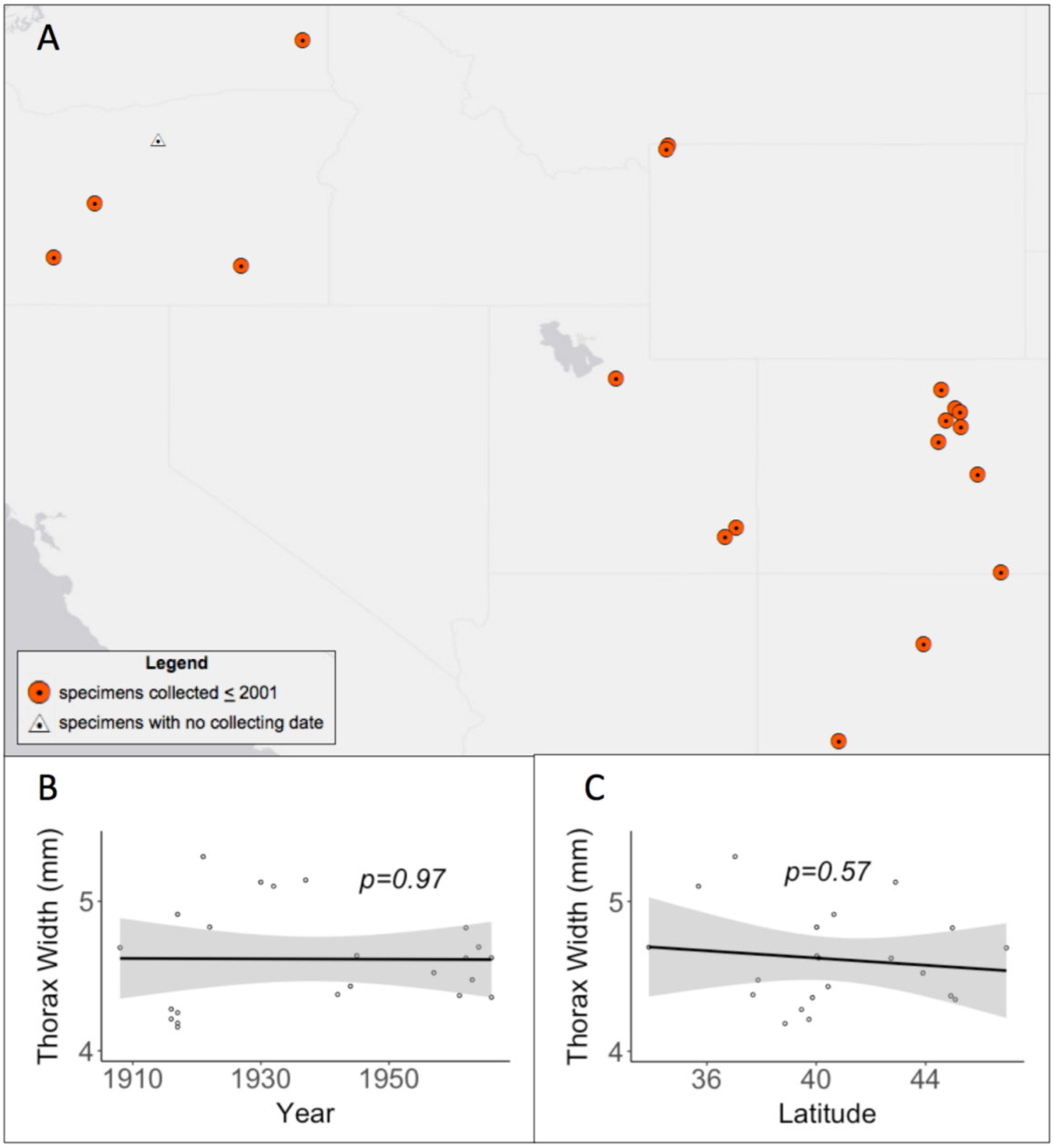

Fig S8. Spatial and temporal distribution of *Bombus centralis* specimens used in analyses. A) Map of *Bombus centralis* specimens' collecting locations. B) Temporal trend of *Bombus centralis* body size. C) Latitudinal trend of *Bombus centralis* body size.

Species: *Bombus crotchii*

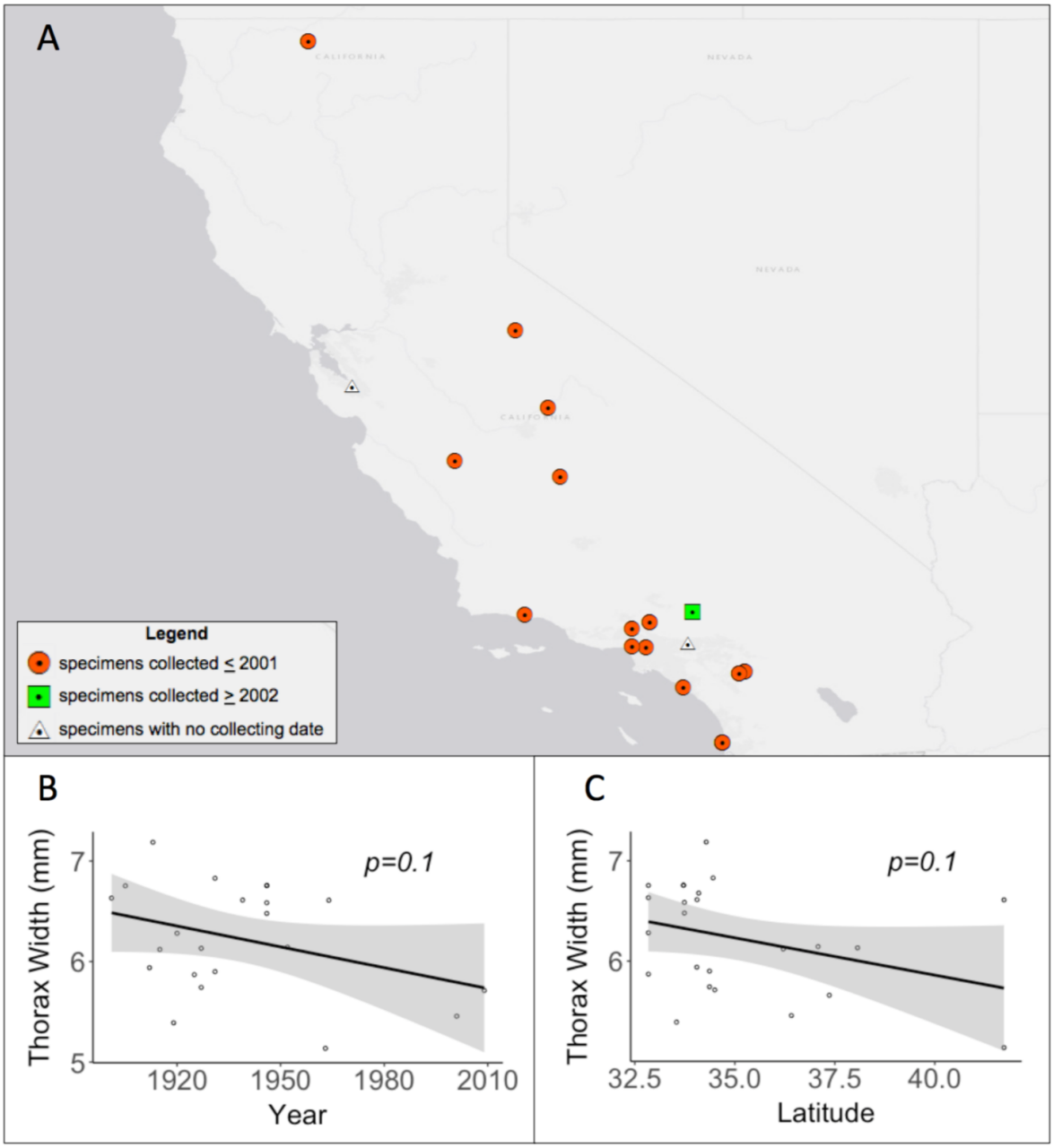

Fig S9. Spatial and temporal distribution of *Bombus crotchii* specimens used in analyses. A) Map of *Bombus crotchii* specimens' collecting locations. B) Temporal trend of *Bombus crotchii* body size. C) Latitudinal trend of *Bombus crotchii* body size.

Species: *Bombus fervidus*

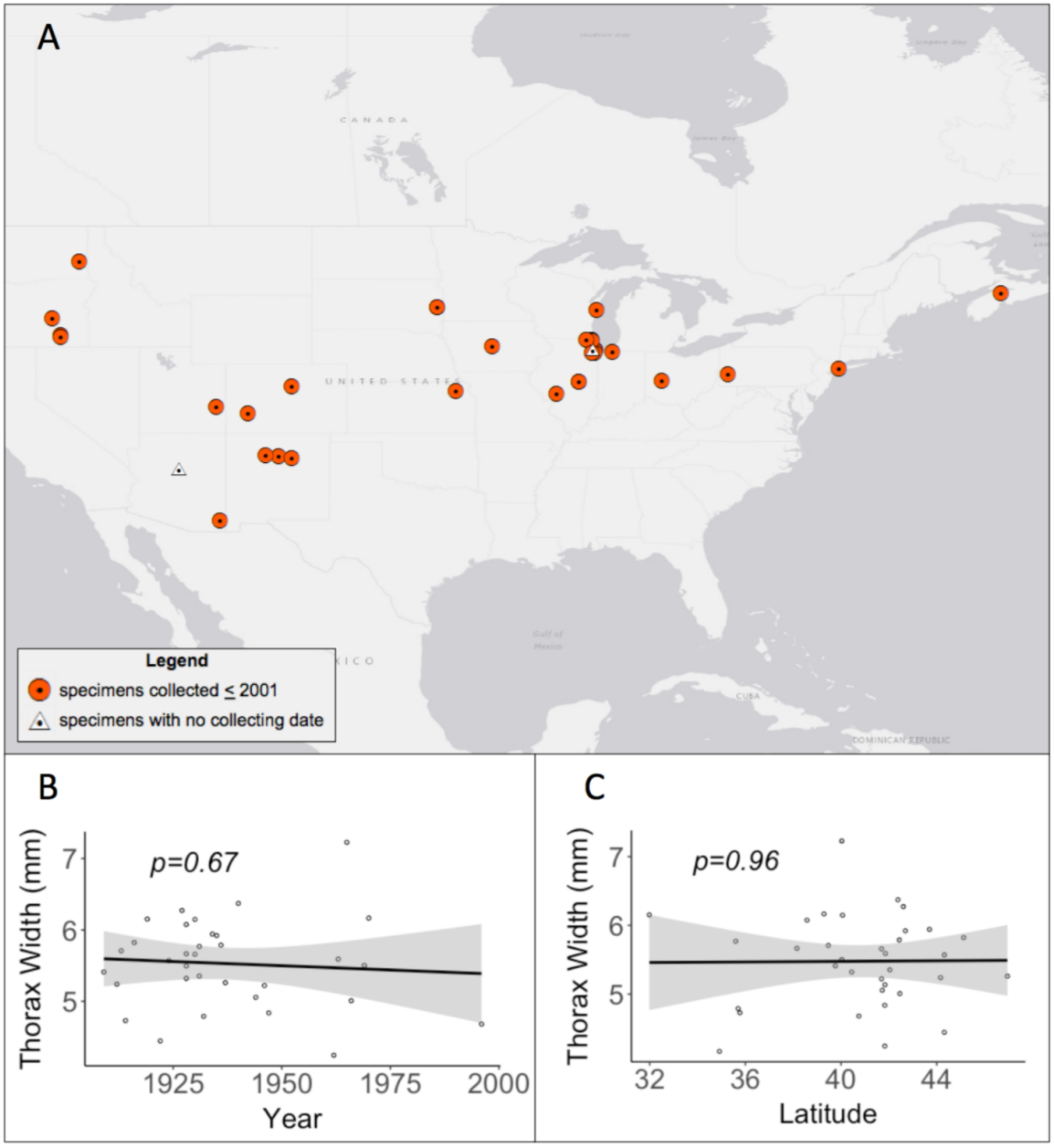

Fig S10. Spatial and temporal distribution of *Bombus fervidus* specimens used in analyses. A) Map of *Bombus fervidus* specimens' collecting locations. B) Temporal trend of *Bombus fervidus* body size. C) Latitudinal trend of *Bombus fervidus* body size.

Species: *Bombus flavifrons*

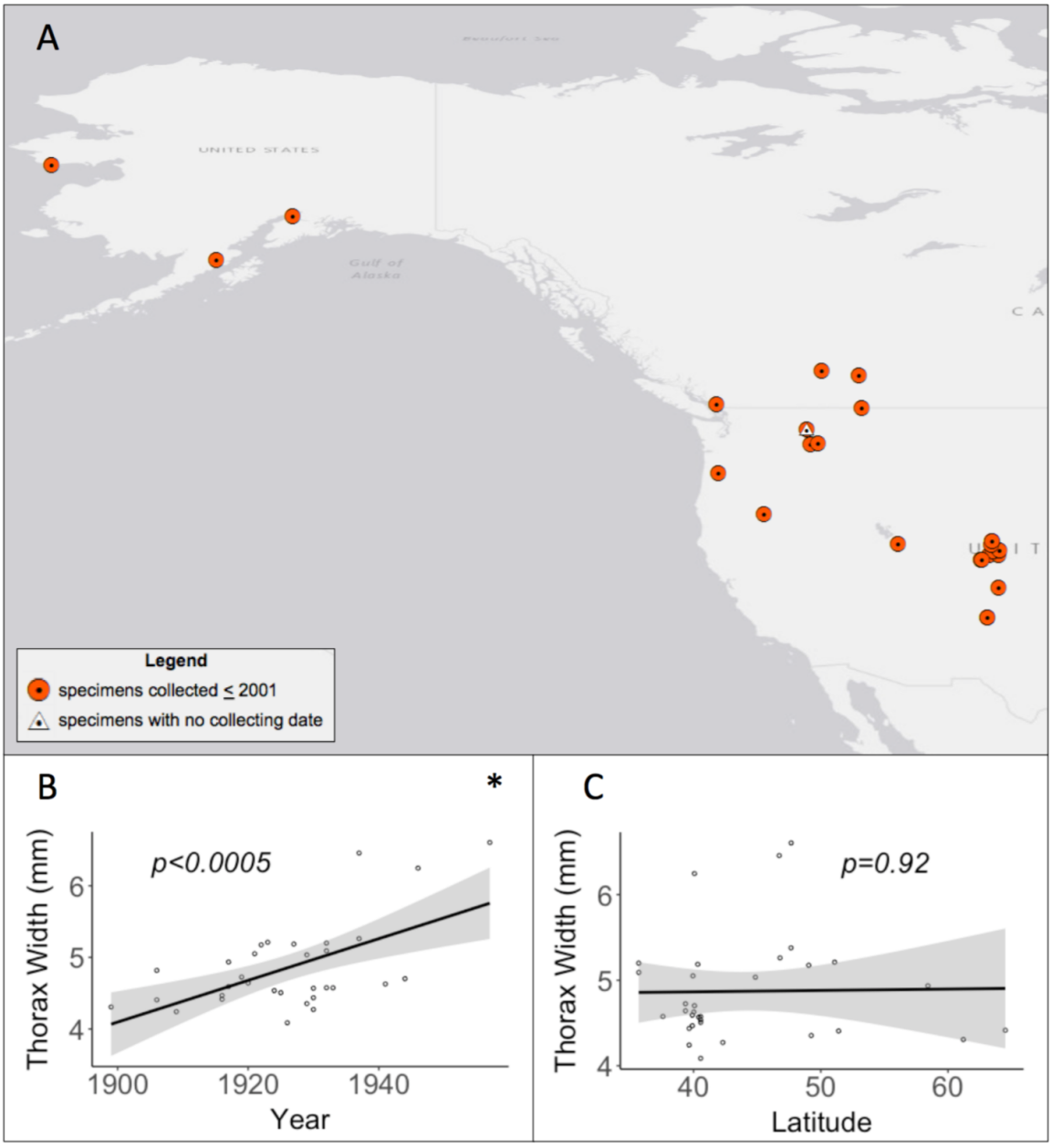

Fig S11. Spatial and temporal distribution of *Bombus flavifrons* specimens used in analyses. A) Map of *Bombus flavifrons* specimens' collecting locations. B) Temporal trend of *Bombus flavifrons* body size. C) Latitudinal trend of *Bombus flavifrons* body size. Asterisk (*) denotes statistical significance.

Species: *Bombus fraternus*

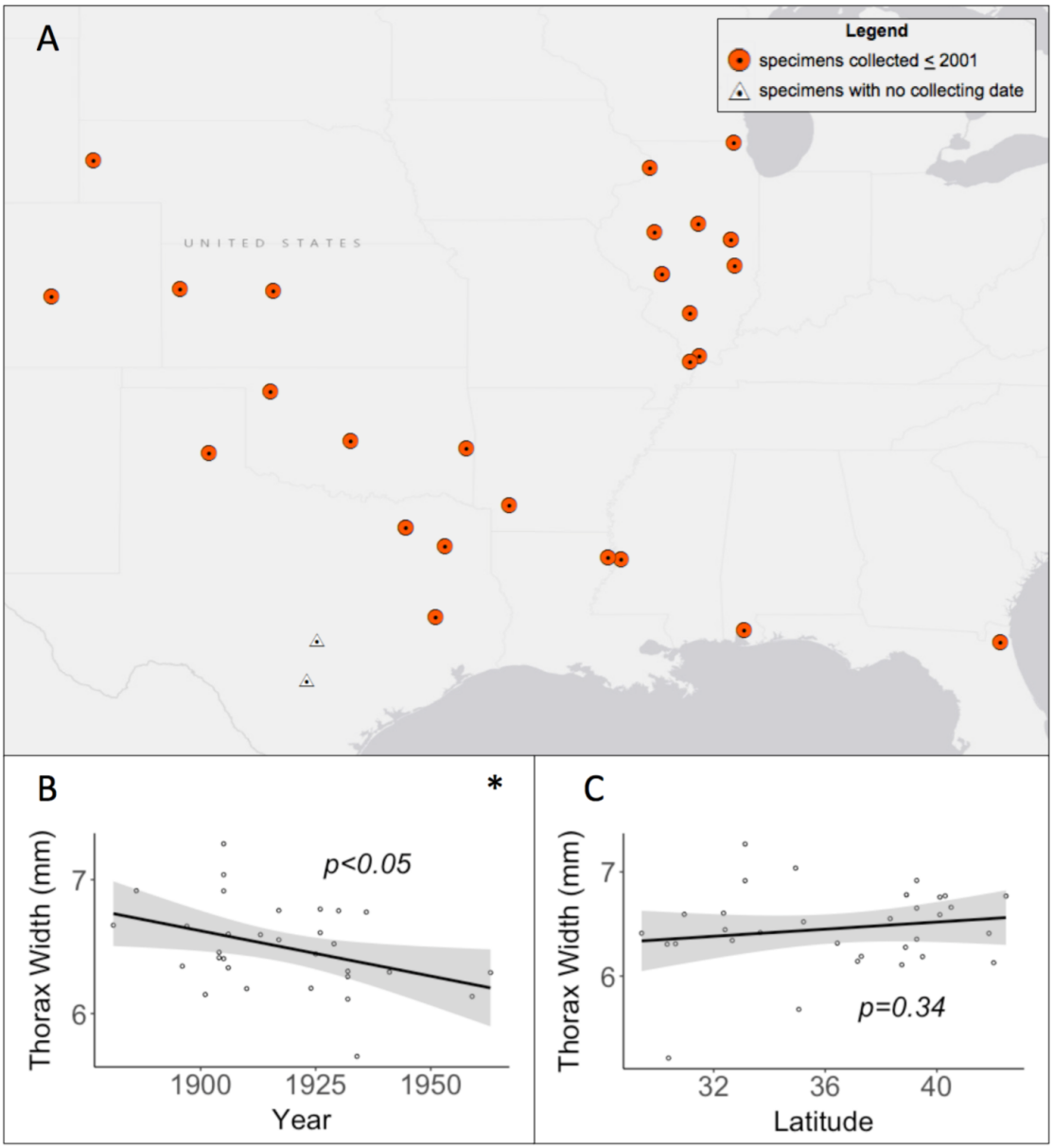

Fig S12. Spatial and temporal distribution of *Bombus fraternus* specimens used in analyses. A) Map of *Bombus fraternus* specimens' collecting locations. B) Temporal trend of *Bombus fraternus* body size. C) Latitudinal trend of *Bombus fraternus* body size. Asterisk (*) denotes statistical significance.

Species: *Bombus frigidus*

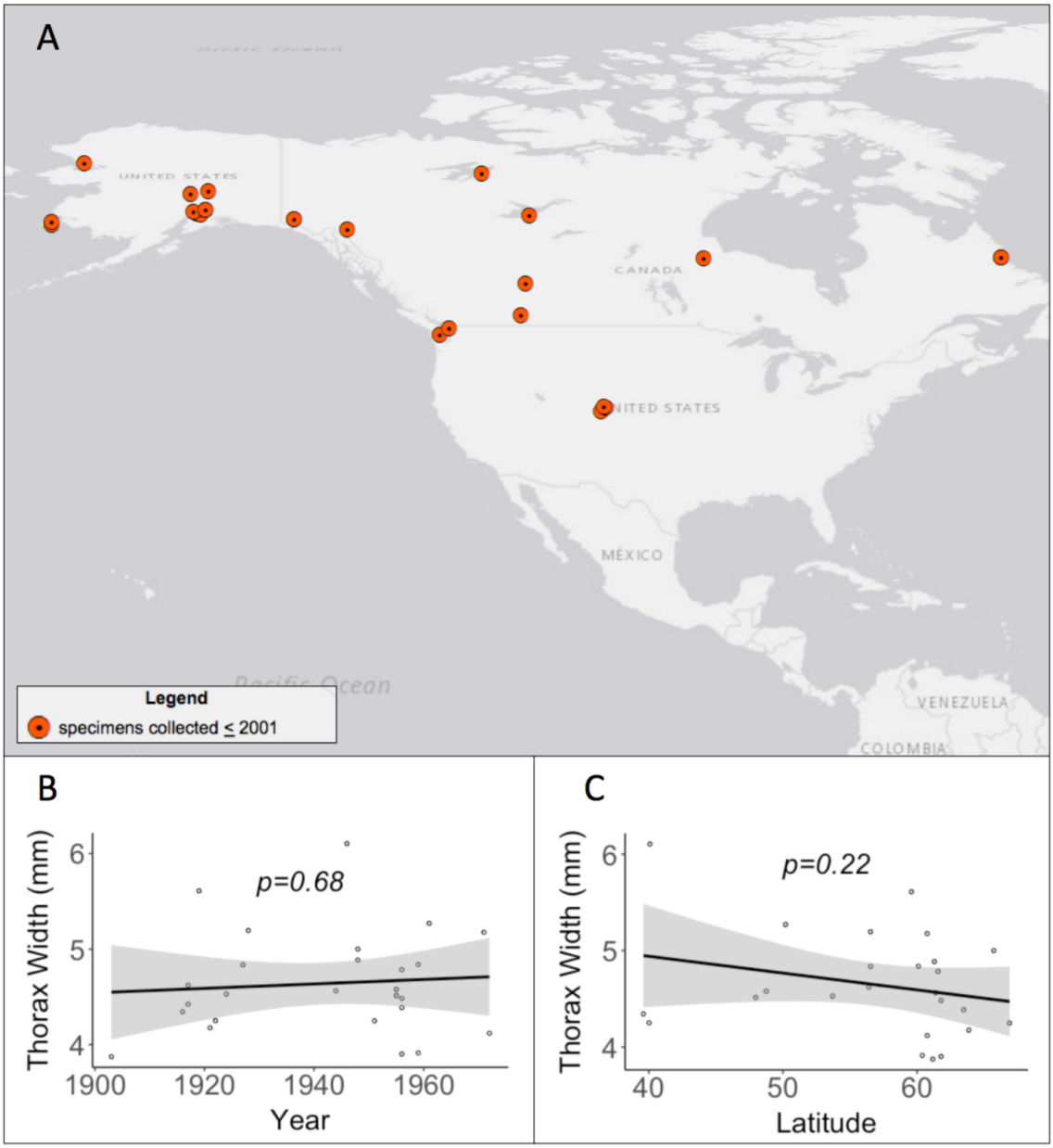

Fig S13. Spatial and temporal distribution of *Bombus frigidus* specimens used in analyses. A) Map of *Bombus frigidus* specimens' collecting locations. B) Temporal trend of *Bombus frigidus* body size. C) Latitudinal trend of *Bombus frigidus* body size.

Species: *Bombus griseocollis*

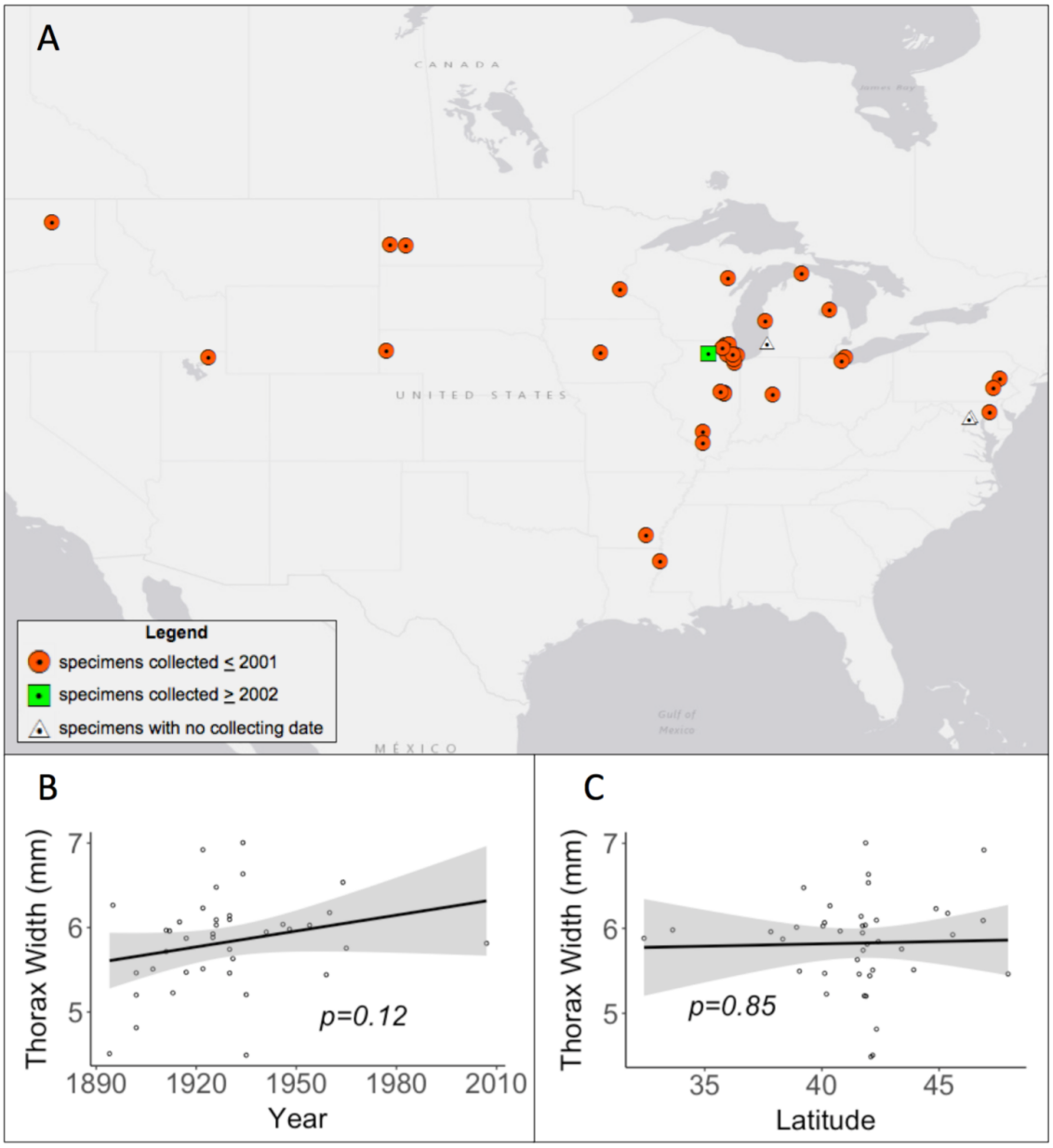

Fig S14. Spatial and temporal distribution of *Bombus griseocollis* specimens used in analyses. A) Map of *Bombus griseocollis* specimens' collecting locations. B) Temporal trend of *Bombus griseocollis* body size. C) Latitudinal trend of *Bombus griseocollis* body size.

Species: *Bombus huntii*

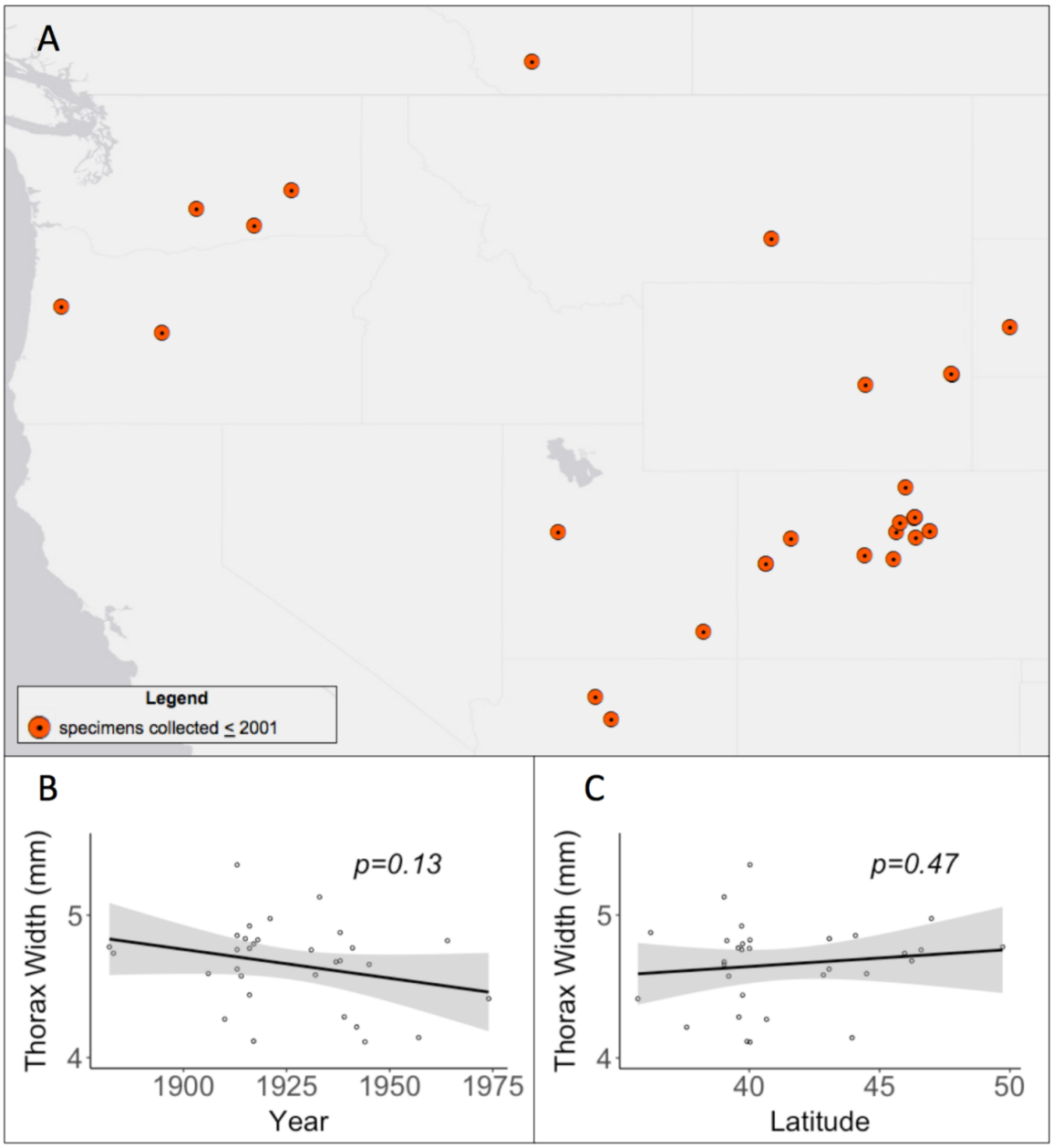

Fig S15. Spatial and temporal distribution of *Bombus huntii* specimens used in analyses. A) Map of *Bombus huntii* specimens' collecting locations. B) Temporal trend of *Bombus huntii* body size. C) Latitudinal trend of *Bombus huntii* body size.

Species: *Bombus impatiens*

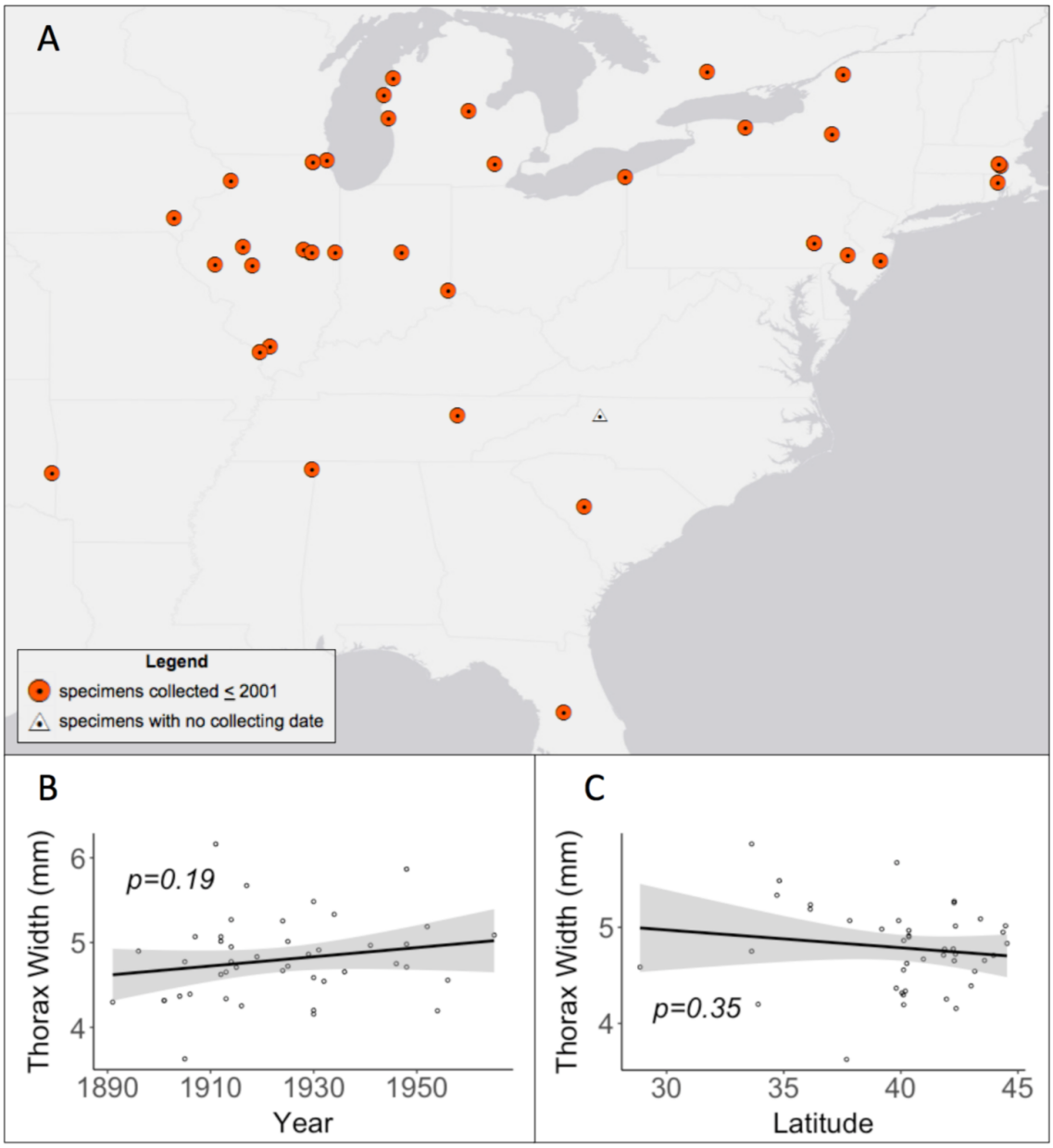

Fig S16. Spatial and temporal distribution of *Bombus impatiens* specimens used in analyses. A) Map of *Bombus impatiens* specimens' collecting locations. B) Temporal trend of *Bombus impatiens* body size. C) Latitudinal trend of *Bombus impatiens* body size.

Species: *Bombus melanopygus*

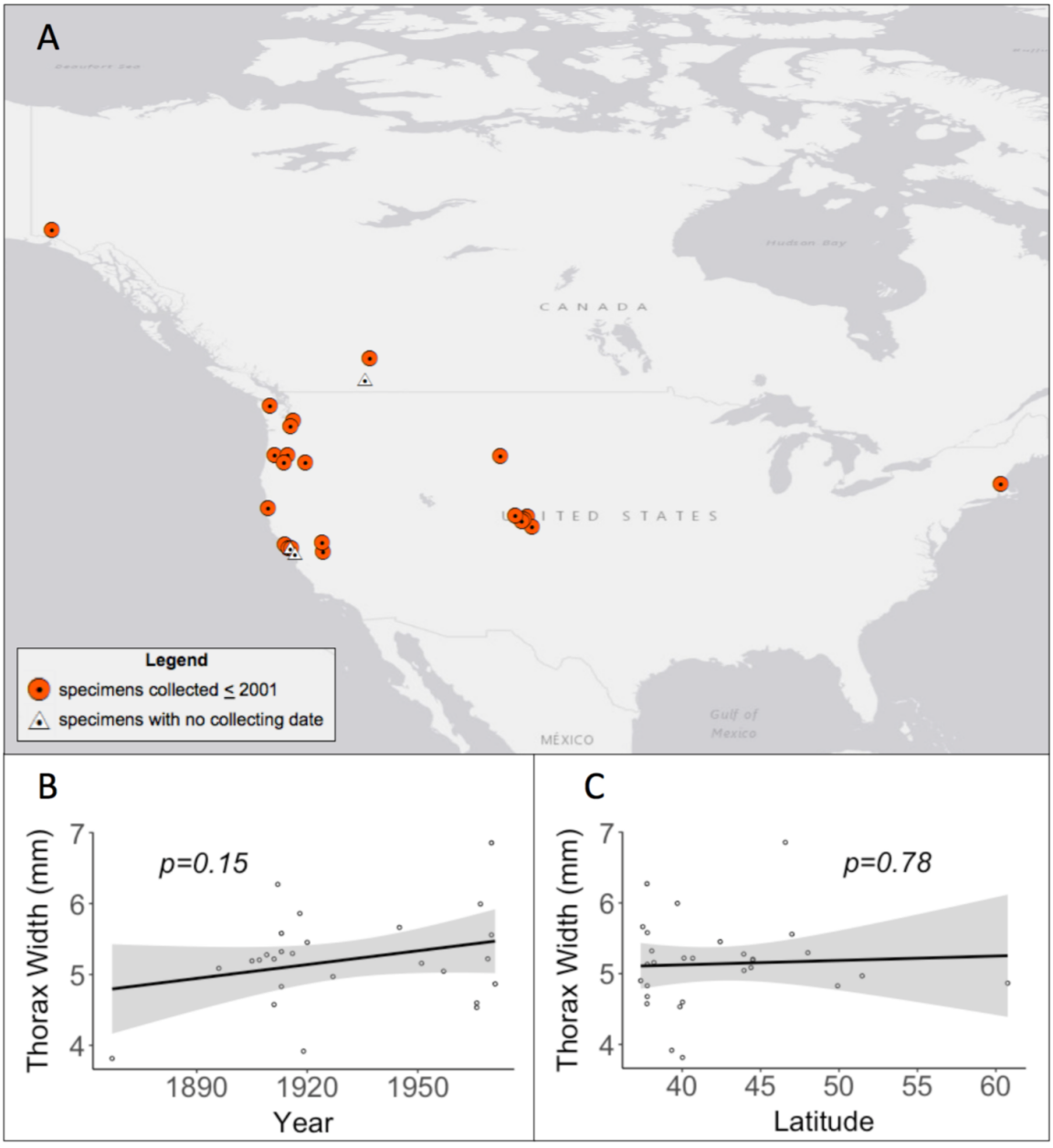

Fig S17. Spatial and temporal distribution of *Bombus melanopygus* specimens used in analyses. A) Map of *Bombus melanopygus* specimens' collecting locations. B) Temporal trend of *Bombus melanopygus* body size. C) Latitudinal trend of *Bombus melanopygus* body size.

Species: *Bombus mixtus*

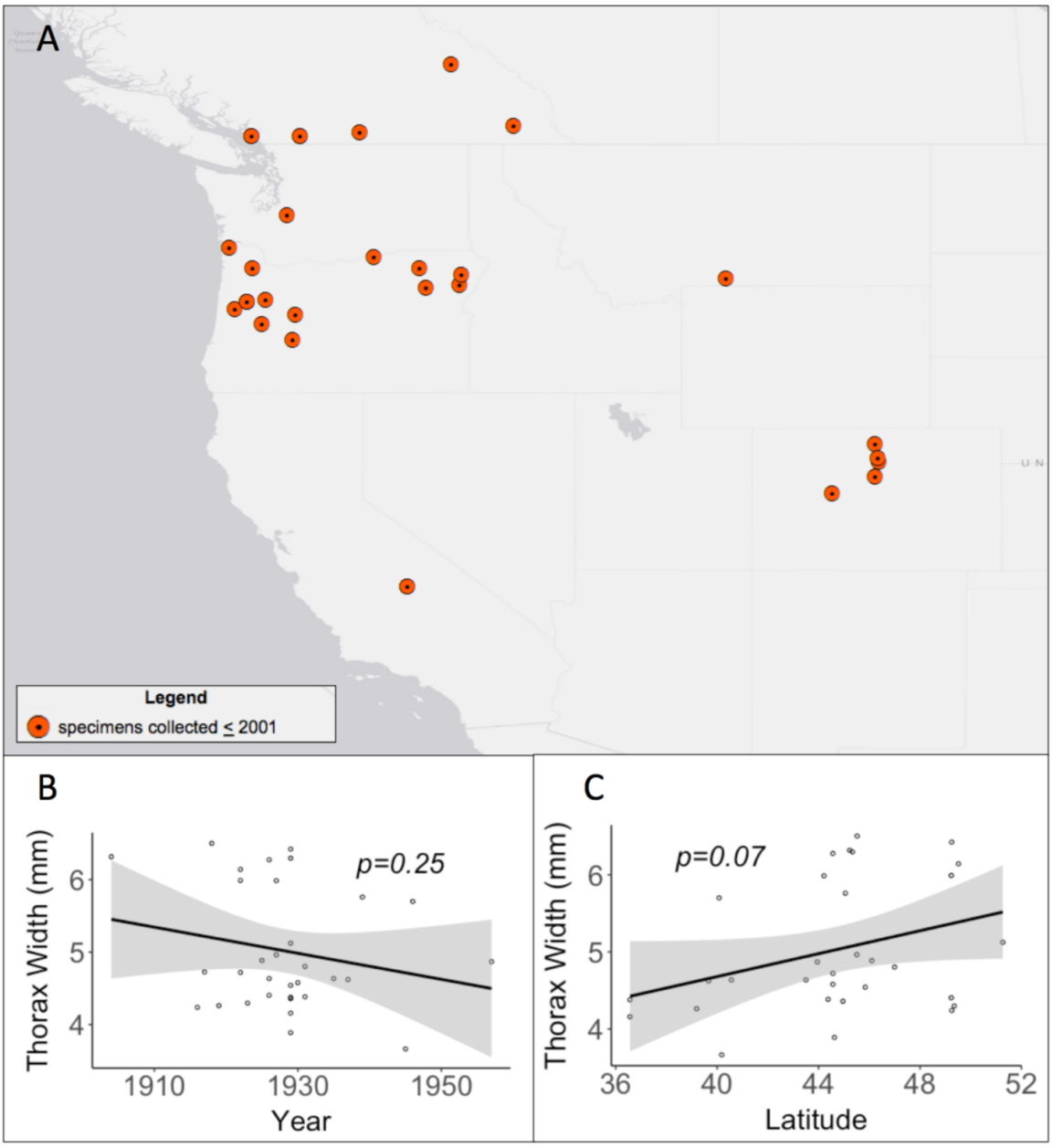

Fig S18. Spatial and temporal distribution of *Bombus mixtus* specimens used in analyses. A) Map of *Bombus mixtus* specimens' collecting locations. B) Temporal trend of *Bombus mixtus* body size. C) Latitudinal trend of *Bombus mixtus* body size.

Species: *Bombus morrisoni*

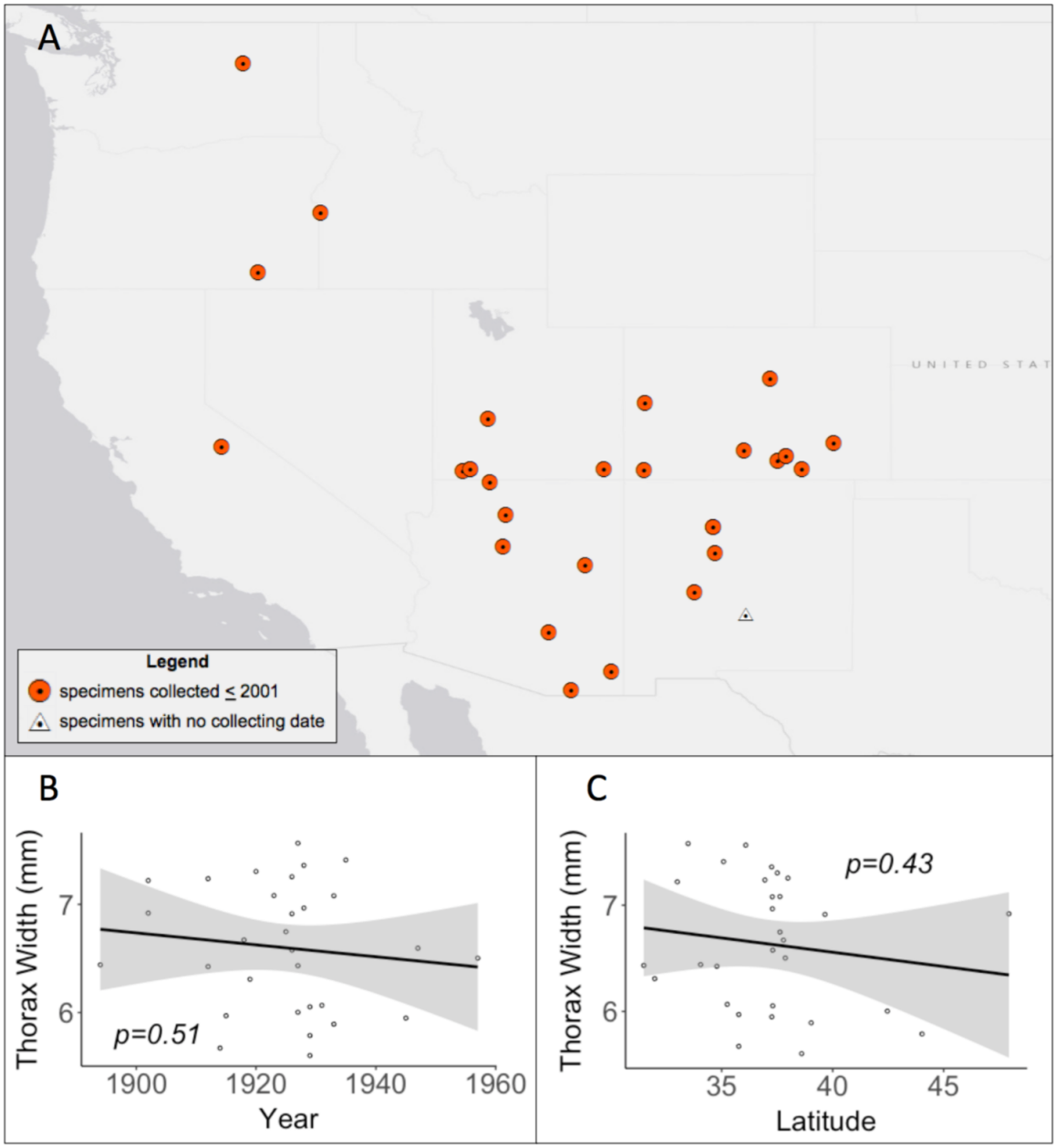

Fig S19. Spatial and temporal distribution of *Bombus morrisoni* specimens used in analyses. A) Map of *Bombus morrisoni* specimens' collecting locations. B) Temporal trend of *Bombus morrisoni* body size. C) Latitudinal trend of *Bombus morrisoni* body size.

Species: *Bombus nevadensis*

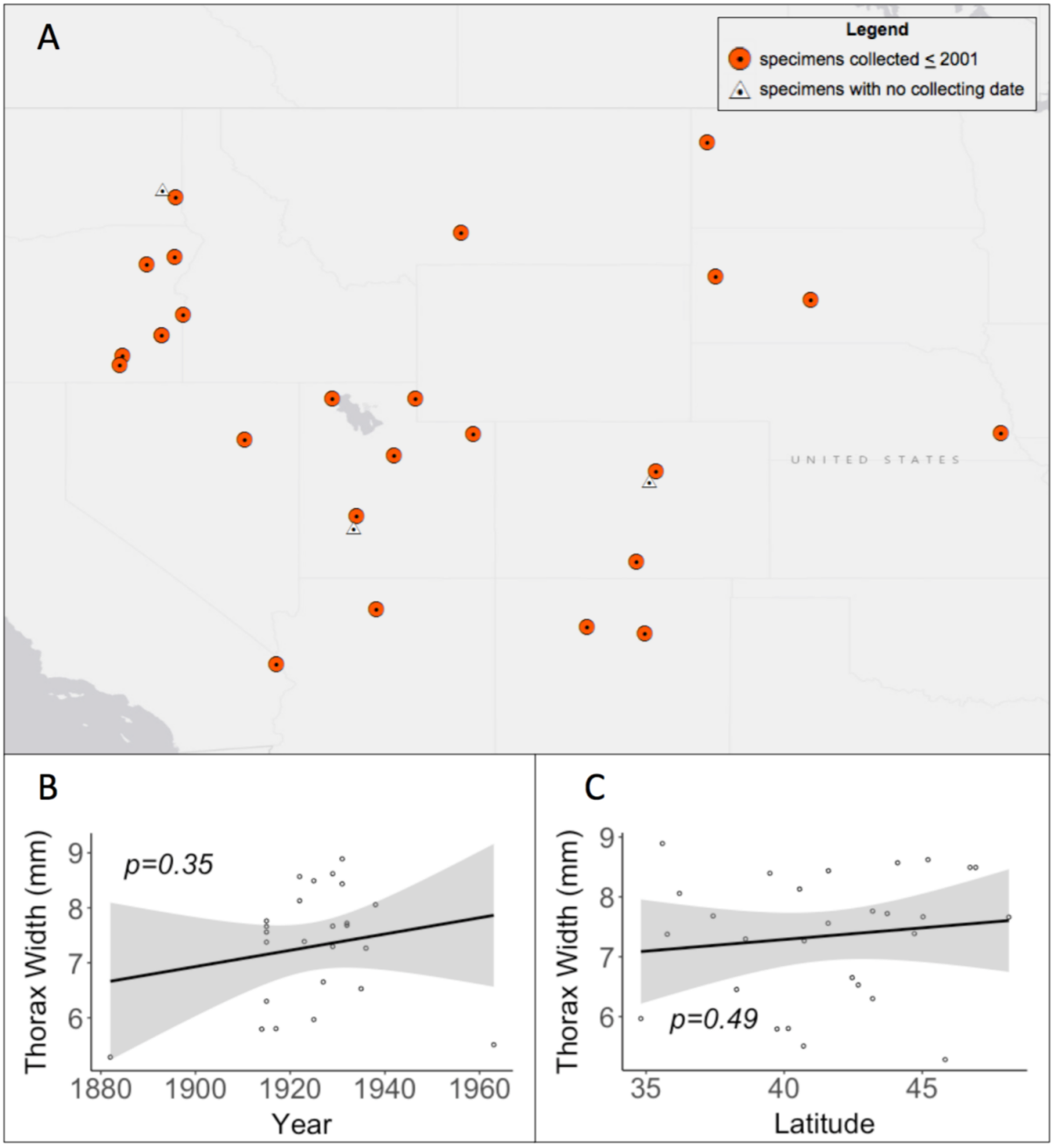

Fig S20. Spatial and temporal distribution of *Bombus nevadensis* specimens used in analyses. A) Map of *Bombus nevadensis* specimens' collecting locations. B) Temporal trend of *Bombus nevadensis* body size. C) Latitudinal trend of *Bombus nevadensis* body size.

Species: *Bombus occidentalis*

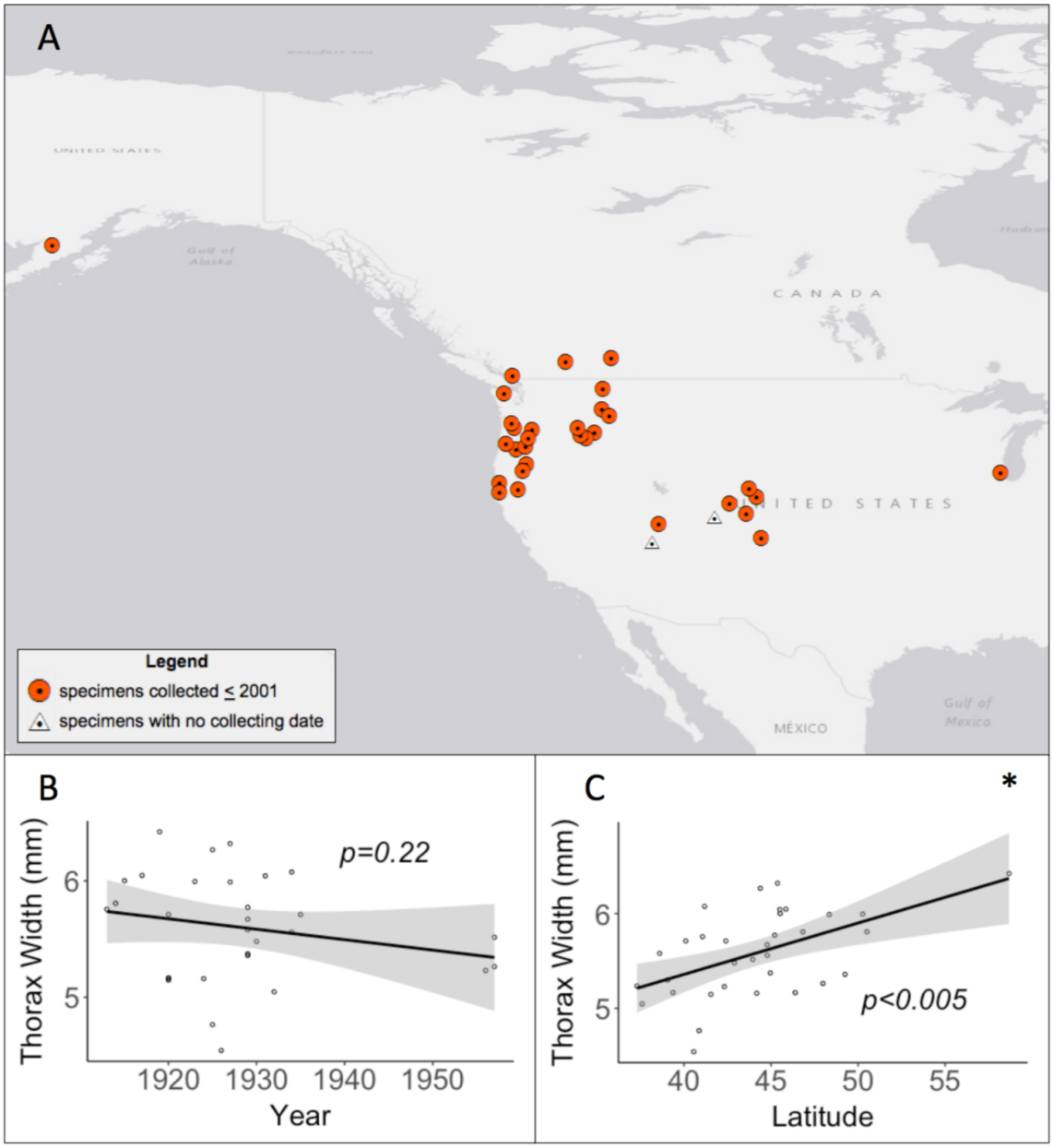

Fig S21. Spatial and temporal distribution of *Bombus occidentalis* specimens used in analyses. A) Map of *Bombus occidentalis* specimens' collecting locations. B) Temporal trend of *Bombus occidentalis* body size. C) Latitudinal trend of *Bombus occidentalis* body size. Asterisk (*) denotes statistical significance.

Species: *Bombus pensylvanicus*

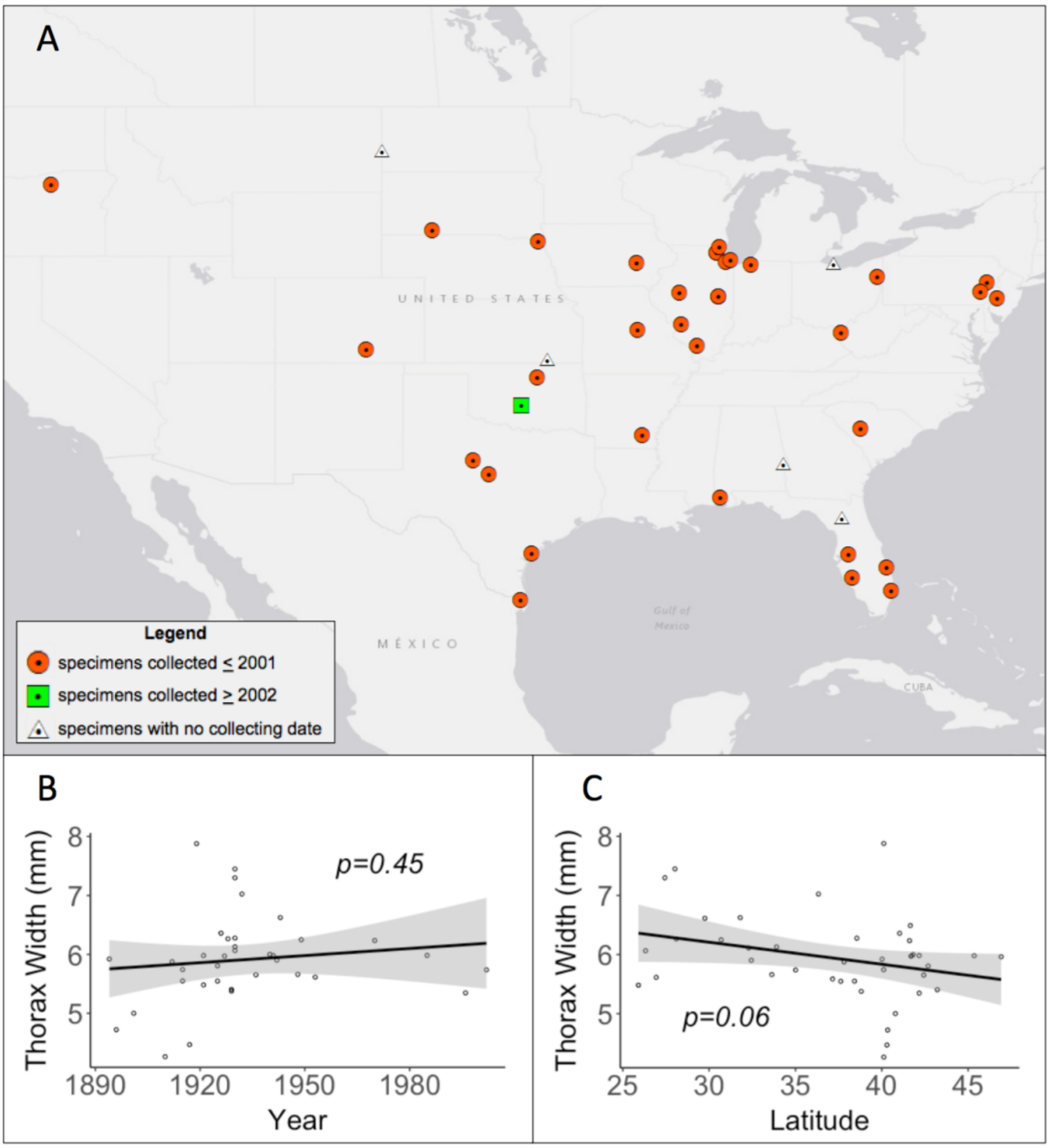

Fig S22. Spatial and temporal distribution of *Bombus pensylvanicus* specimens used in analyses. A) Map of *Bombus pensylvanicus* specimens' collecting locations. B) Temporal trend of *Bombus pensylvanicus* body size. C) Latitudinal trend of *Bombus pensylvanicus* body size.

Species: *Bombus perplexus*

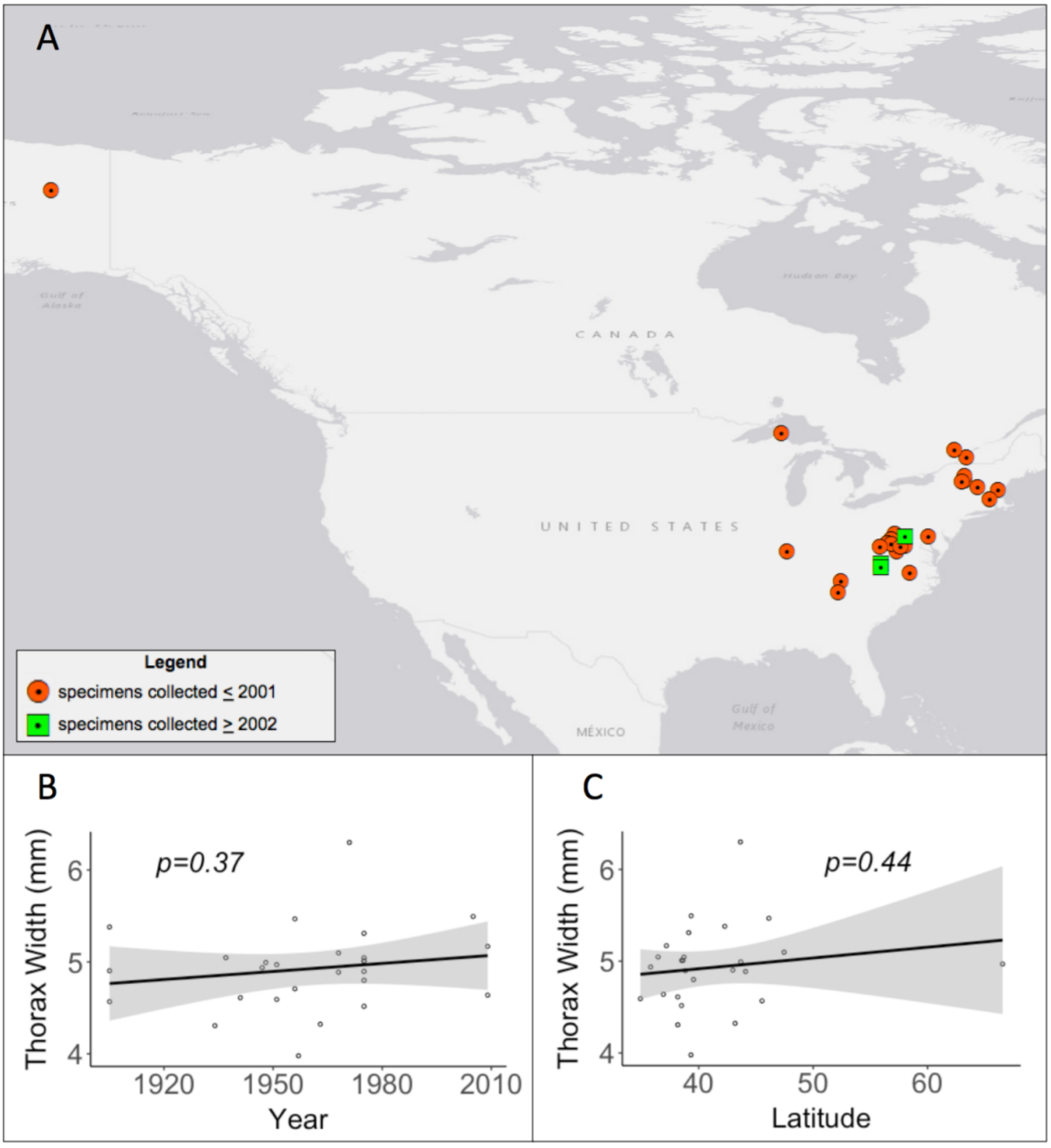

Fig S23. Spatial and temporal distribution of *Bombus perplexus* specimens used in analyses. A) Map of *Bombus perplexus* specimens' collecting locations. B) Temporal trend of *Bombus perplexus* body size. C) Latitudinal trend of *Bombus perplexus* body size.

Species: *Bombus rufocinctus*

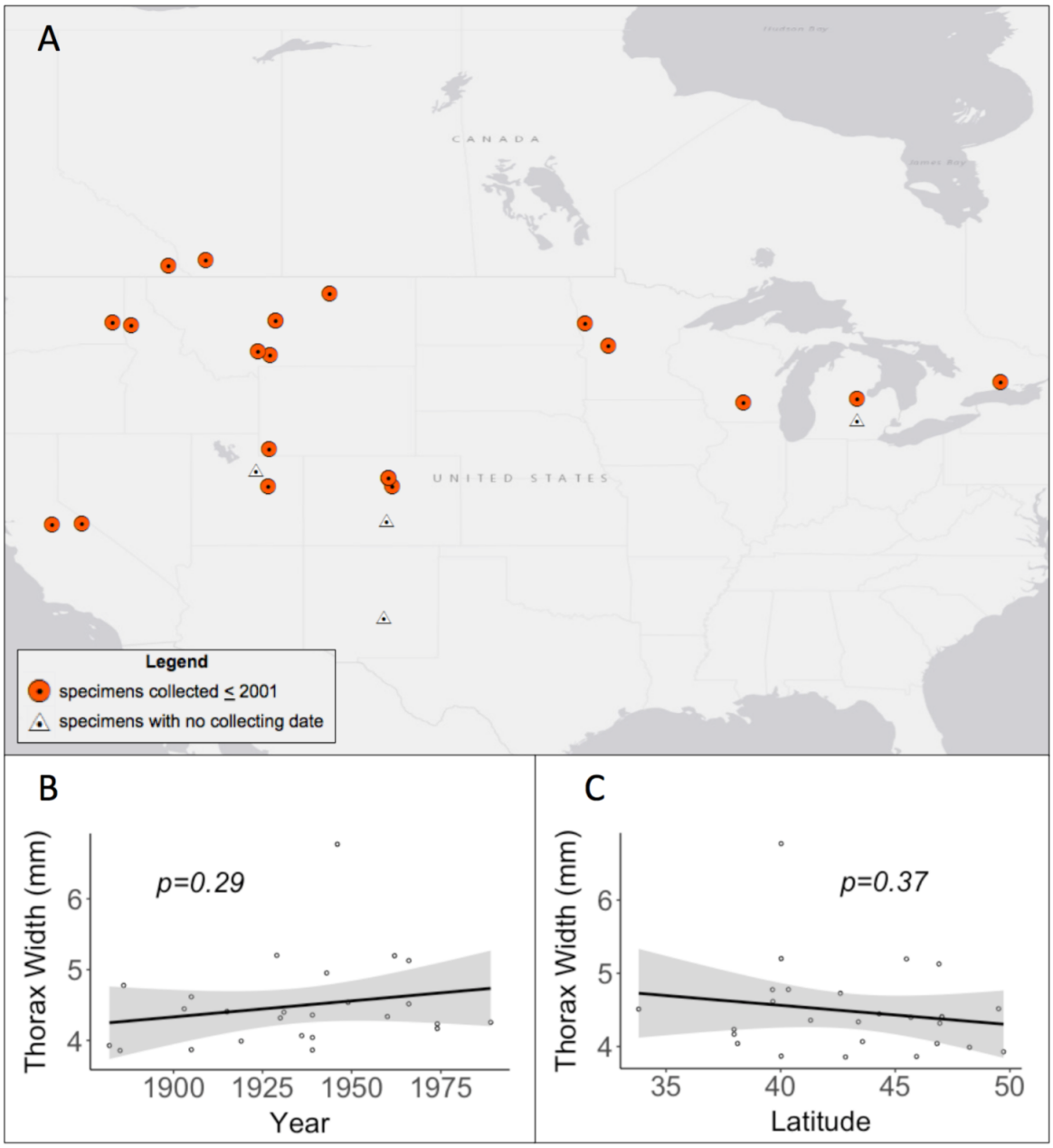

Fig S24. Spatial and temporal distribution of *Bombus rufocinctus* specimens used in analyses. A) Map of *Bombus rufocinctus* specimens' collecting locations. B) Temporal trend of *Bombus rufocinctus* body size. C) Latitudinal trend of *Bombus rufocinctus* body size.

Species: *Bombus sandersoni*

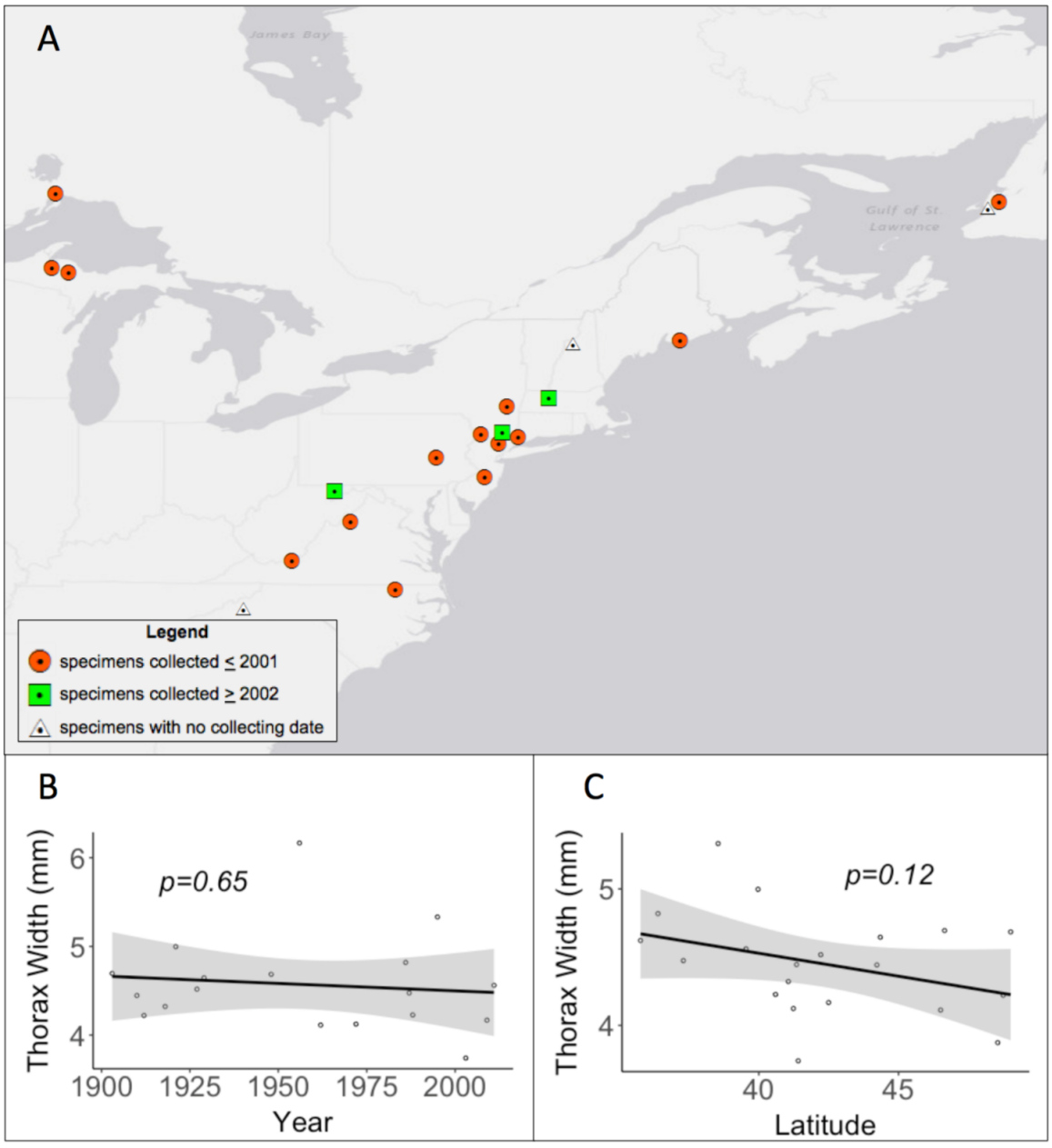

Fig S25. Spatial and temporal distribution of *Bombus sandersoni* specimens used in analyses. A) Map of *Bombus sandersoni* specimens' collecting locations. B) Temporal trend of *Bombus sandersoni* body size. C) Latitudinal trend of *Bombus sandersoni* body size.

Species: *Bombus sylvicola*

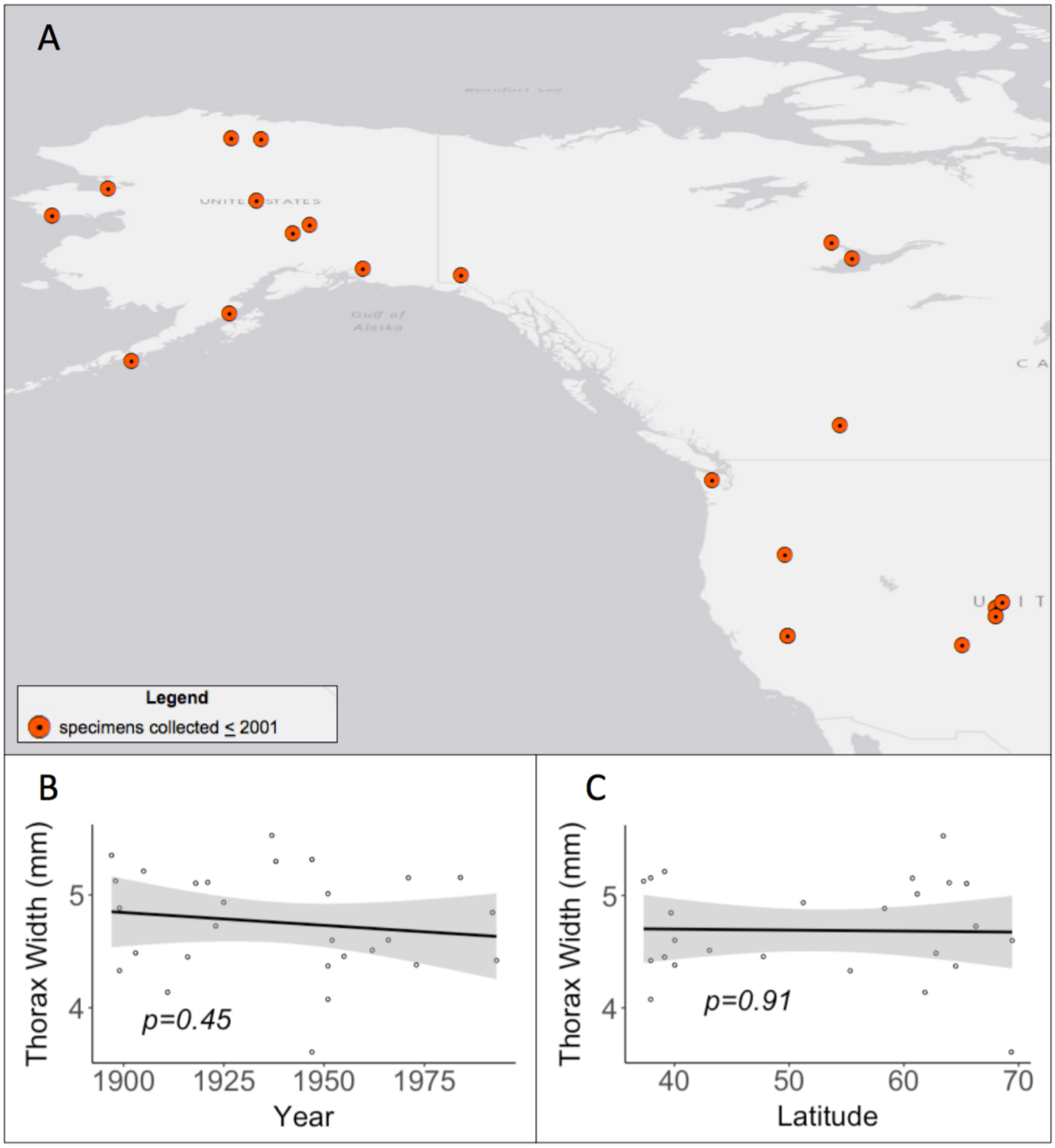

Fig S26. Spatial and temporal distribution of *Bombus sylvicola* specimens used in analyses. A) Map of *Bombus sylvicola* specimens' collecting locations. B) Temporal trend of *Bombus sylvicola* body size. C) Latitudinal trend of *Bombus sylvicola* body size.

Species: *Bombus ternarius*

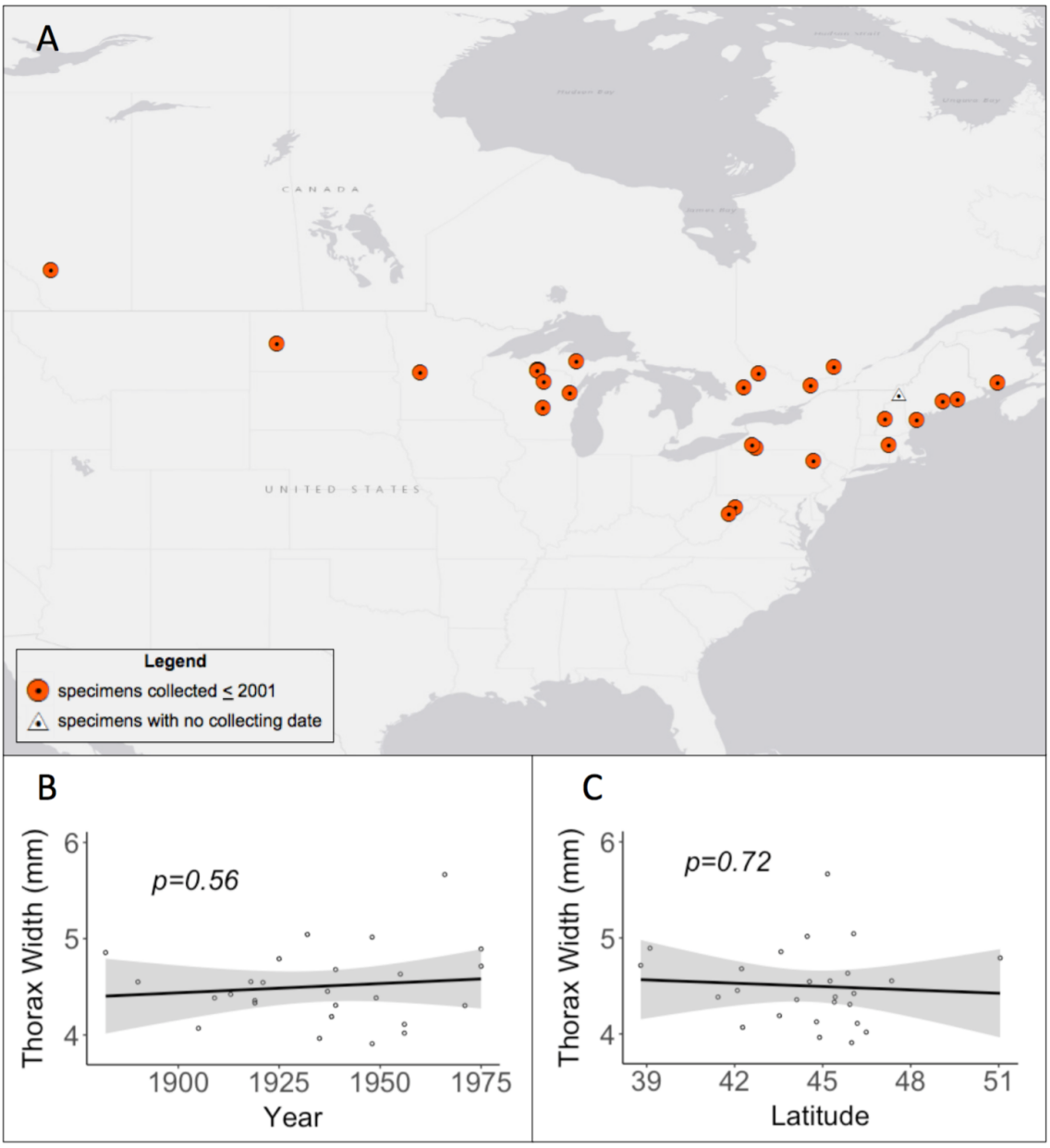

Fig S27. Spatial and temporal distribution of *Bombus ternarius* specimens used in analyses. A) Map of *Bombus ternarius* specimens' collecting locations. B) Temporal trend of *Bombus ternarius* body size. C) Latitudinal trend of *Bombus ternarius* body size.

Species: *Bombus terricola*

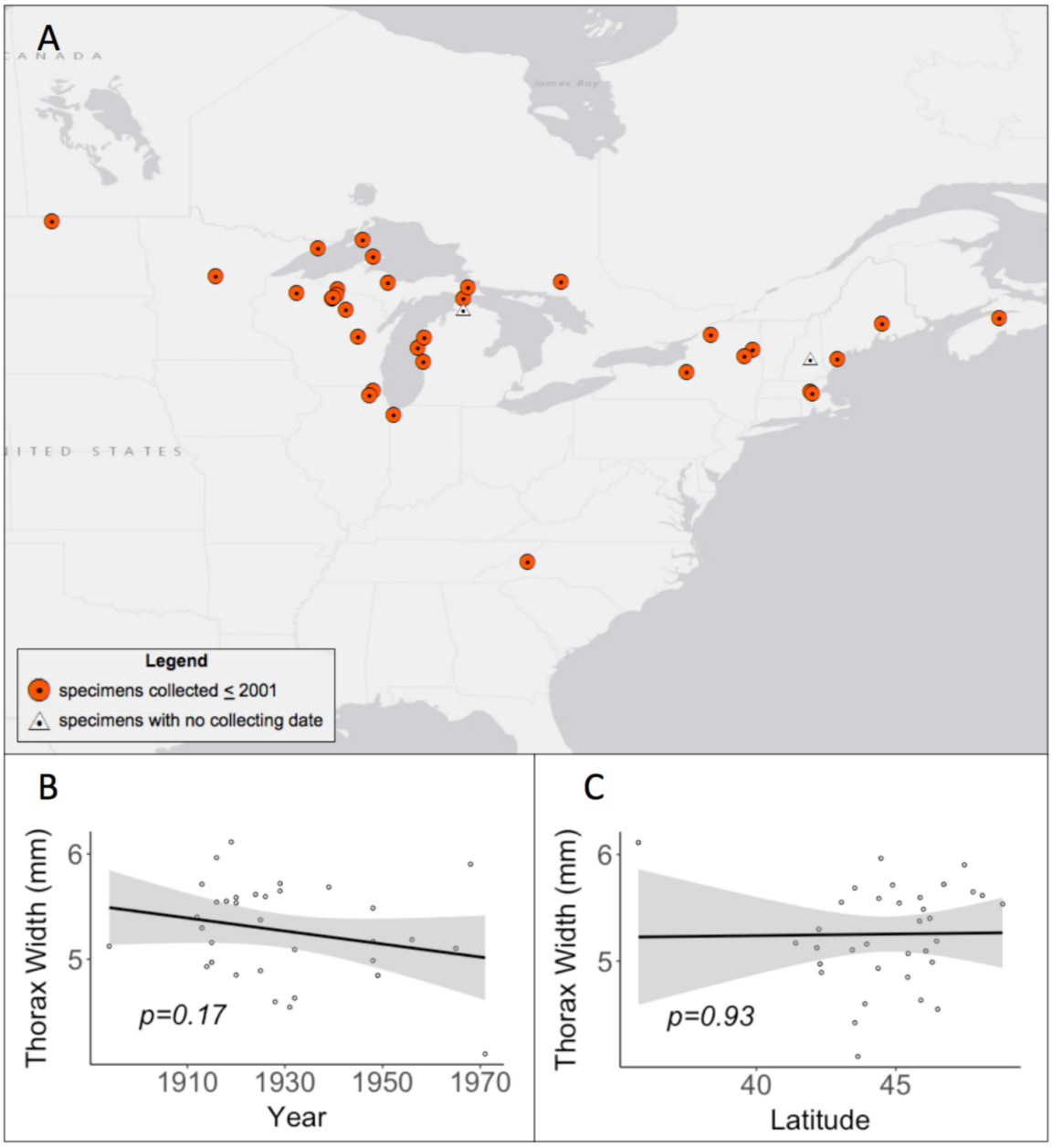

Fig S28. Spatial and temporal distribution of *Bombus terricola* specimens used in analyses. A) Map of *Bombus terricola* specimens' collecting locations. B) Temporal trend of *Bombus terricola* body size. C) Latitudinal trend of *Bombus terricola* body size.

Species: *Bombus vagans*

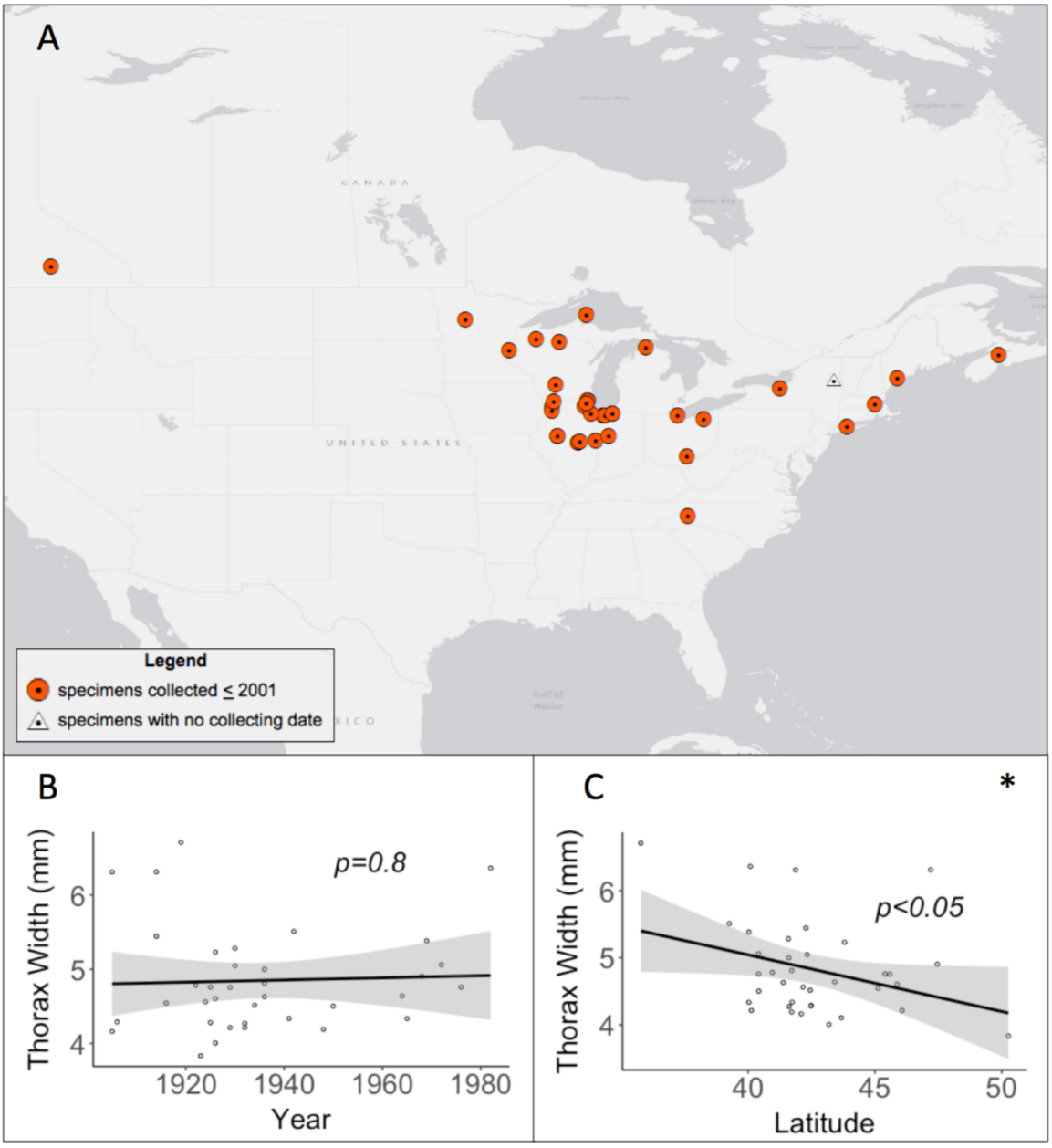

Fig S29. Spatial and temporal distribution of *Bombus vagans* specimens used in analyses. A) Map of *Bombus vagans* specimens' collecting locations. B) Temporal trend of *Bombus vagans* body size. C) Latitudinal trend of *Bombus vagans* body size. Asterisk (*) denotes statistical significance.

Species: *Bombus vandykei*

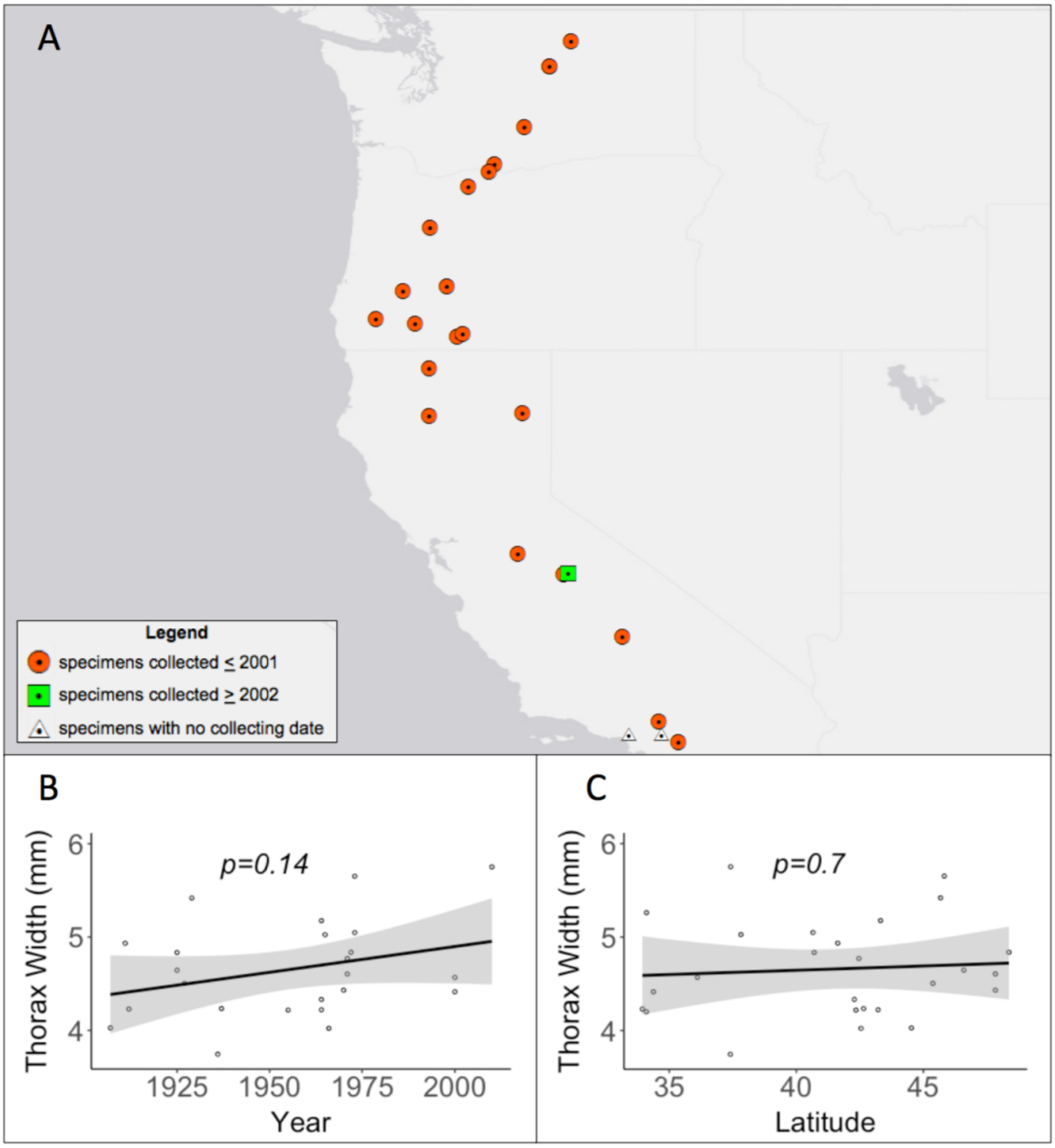

Fig S30. Spatial and temporal distribution of *Bombus vandykei* specimens used in analyses. A) Map of *Bombus vandykei* specimens' collecting locations. B) Temporal trend of *Bombus vandykei* body size. C) Latitudinal trend of *Bombus vandykei* body size.

Species: *Bombus vosnesenskii*

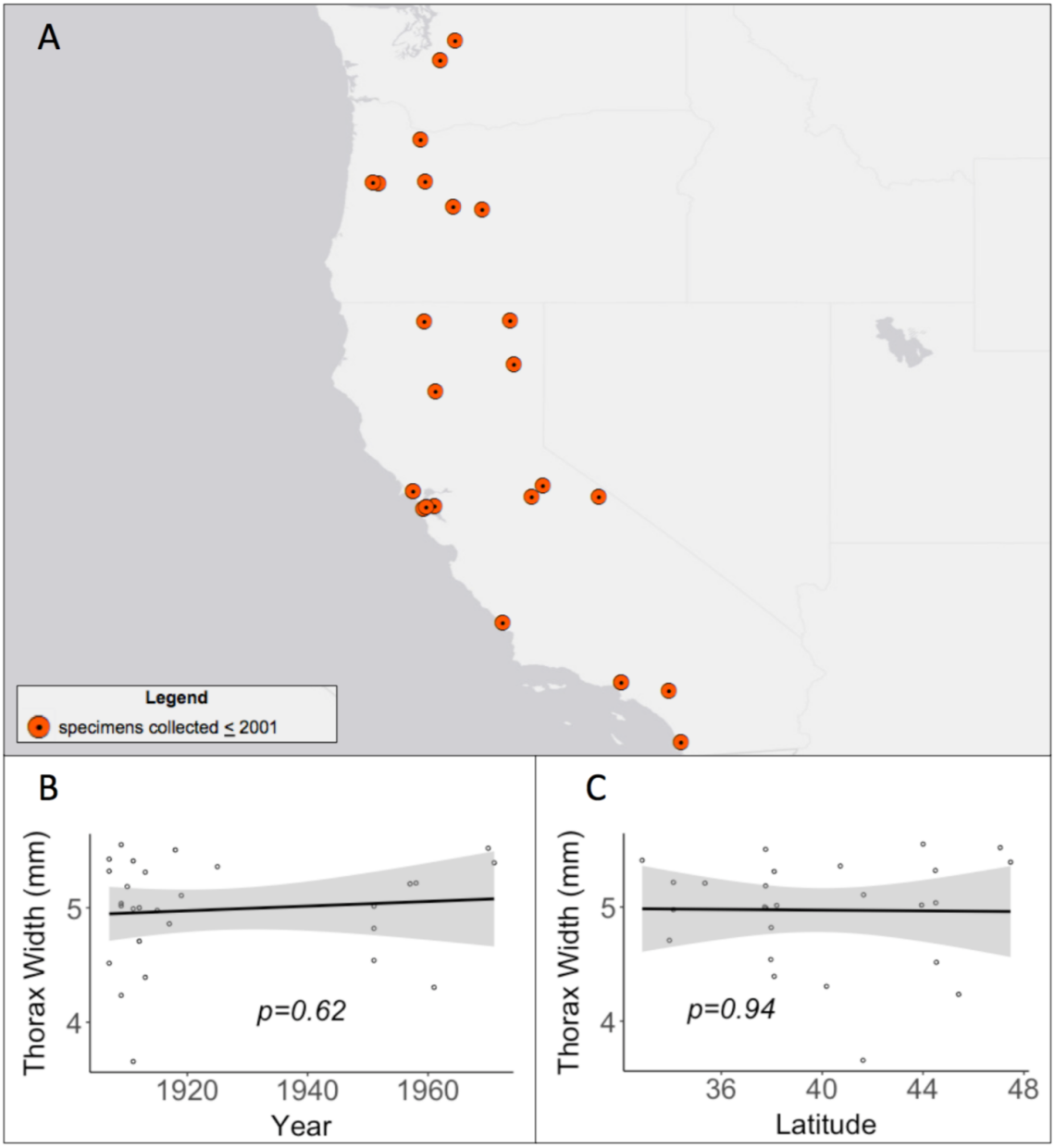

Fig S31. Spatial and temporal distribution of *Bombus vosnesenskii* specimens used in analyses. A) Map of *Bombus vosnesenskii* specimens' collecting locations. B) Temporal trend of *Bombus vosnesenskii* body size. C) Latitudinal trend of *Bombus vosnesenskii* body size.

Chapter II: Bumble bees exhibit intraspecific body size spatial structuring despite low genetic differentiation

Matthew W. Austin[a,b,*], Amber D. Tripodi[c], James P. Strange[c,d], and Aimee S. Dunlap[a,b]

[a]Department of Biology, University of Missouri – St. Louis, St. Louis, MO, USA

[b]Whitney R. Harris World Ecology Center, St. Louis, MO, USA

[c]USDA-ARS-Pollinating Insect Research Unit, Logan, UT, USA

[d]Department of Entomology, The Ohio State University, Columbus, OH, USA

*Corresponding author:

AustinMattW@gmail.com

Abstract

Biodiversity loss among pollinating insects has precipitously increased due to anthropogenic environmental changes. Among these taxa, the most comprehensive estimates of decline are for bees, for which human land use is the predominant driver of decline. Prior studies have demonstrated that human-modified environments can structure bee communities interspecifically, based on the matching of functional traits to local environments. However, little is known about whether bee functional traits can be structured intraspecifically across human-modified landscapes. Here, we study five bumble bee (Apidae: *Bombus*) species across an urban gradient in the greater Saint Louis, Missouri region in the North American Midwest and ask the following questions: (1) Can bumble bees exhibit intraspecific spatial structuring of body size, a developmentally plastic and ecologically consequential functional trait of bees? And, if so, (2) does this body size structure coincide with population genetic structure? We additionally estimate genetic diversity, inbreeding, and colony density of these species - three factors that can affect extinction risk. Using microsatellite genotyping and direct measurements of body size, we find that two of these species (*Bombus impatiens* and *Bombus pensylvanicus*) exhibit intraspecific spatial structuring of body size, despite a lack of population genetic structure. We also reaffirm reports of low genetic diversity in *B. pensylvanicus* and find evidence of inbreeding in *Bombus griseocollis*. Collectively, our results have implications for the conservation of threatened species and suggest that human-modified environments can induce landscape-level structuring within-species of developmentally plastic functional traits.

Keywords*: body size,* Bombus*, microsatellite, North American Midwest, pollinator decline, urban gradient*

Introduction

In the Anthropocene, we have witnessed precipitous declines of biodiversity (Corlett 2015), with approximately 1 million extant species currently in threat of extinction (IPBES 2019). Anthropogenic effects on the globe are widely recognized as the primary drivers of this biodiversity loss (IPBES 2019). Humans have transformed up to one-half of global land surfaces (Vitousek et al. 1997), thereby fragmenting previously continuous habitat and presenting many species with environments unencountered in their evolutionary past (Wong & Candolin 2015). Anthropogenic change may increase extinction risk by inducing mismatch between functional traits and the environment, if such traits are not sufficiently plastic (e.g., Hale & Swearer 2016). Additionally, through isolating subpopulations by creating barriers to dispersal, such habitat fragmentation may induce distinct genetic phenomena (e.g., increased population differentiation), which can further exacerbate declines (e.g., Charman et al. 2010). As functional traits mediate population performance via effects on fitness (Violle et al. 2007), while population genetics indicate long-term population stability (e.g., Husemann et al. 2016), effective conservation efforts are strengthened by integrative assessments of population genetics and how functional traits are distributed in human-modified environments.

Biodiversity loss among pollinating insects is particularly important for empirical inquiry, as insects are primarily responsible for the pollination of wild plants and agricultural crops (Wagner 2020). Of the pollinating insects, the most comprehensive estimates of decline are for bees (Goulson et al. 2015) and butterflies (e.g., Thogmartin et al. 2017). Various bee taxa have experienced range contractions (e.g., Cameron et al. 2011), abundance declines (e.g., Cameron et al. 2011), and local extinctions (e.g., Burkle

et al. 2013; Grixti et al. 2009), thereby resulting in species richness losses. Among these taxa are the bumble bees (Hymenoptera: Apidae: *Bombus*), a monophyletic group of eusocial bees primarily native to temperate and subpolar regions of the Northern Hemisphere (Goulson 2010). Bumble bees have undergone precipitous declines throughout their native range (e.g., Colla et al. 2012; Hatfield et al. 2015), with estimates suggesting that approximately one-third of bumble bee species are in decline (Arbetman et al. 2017). Anthropogenic habitat modification is widely recognized as a predominant driver of these declines (Goulson et al. 2015), with habitat loss reducing the availability of forage and nesting sites (Goulson et al. 2015), fragmentation inducing heterogeneity in species occurrences (e.g., Bommarco et al. 2010), and population success differing between rural and urban areas (Hall et al. 2016).

Previous studies have demonstrated that human-modified environments can structure bee communities interspecifically, based on the matching of functional traits to local environments (e.g., Banaszak-Cibicka & Zmihorski 2012; Wilson & Jamieson 2019). However, whether human-modified environments can structure bee functional traits intraspecifically is largely unknown. In bees, body size is one functional trait that has considerable ecological consequences. At the community-level, body size influences pollination system connectivity by dictating which floral species a bee can forage from (Peat et al. 2005). At the individual-level, body size influences a suite of characteristics, including dispersal distance (Greenleaf et al. 2007), foraging efficiency (Spaethe and Weidenmüller 2002), and resistance to starvation (Couvillon and Dornhaus 2010). In bumble bees, body size is developmentally plastic, with higher rates of larval feeding yielding larger adult workers (Pendrel and Plowright 1981; Sutcliffe and Plowright

1988). This plasticity can result in up to 10-fold differences in worker body size within colonies, despite workers from monogamous queens being highly related (r=0.75) (Couvillon and Dornhaus 2009; Goulson 2010). Furthermore, body size may influence bumble bees' susceptibility to decline; species with larger average body size (Bartomeus et al. 2013) or lower variation in body size (Austin & Dunlap 2019) appear more susceptible to negative effects of human activity. Despite the known ecological implications of bumble bee body size, we lack an understanding of how body size can be structured within-species across human-modified environments.

Conservation efforts are strengthened by considering how functional traits are spatially structured. Understanding the link between environment and phenotype is critical for habitat restoration (e.g., Watters et al. 2003) and species relocations (e.g., Haddaway et al. 2012). Additionally, phenotypic divergence between subpopulations may indicate variance in environmental quality and differential extinction risk among subpopulations (e.g., Lema & Nevitt 2006). Coupling functional trait investigations with population genetics can elucidate whether phenotypic divergence mirrors patterns of population genetic structuring. If these mirror one another, phenotypic divergence may indicate divergent selection between subpopulations, while phenotypic divergence without genetic structure may indicate plasticity in local environments despite high rates of gene flow (Crispo et al. 2008). This is important as divergent selection can alter the delineation of evolutionary significant units (Fraser & Bernatchez 2001) and the degree to which functional traits are plastic can affect range shifts, extinction, and persistence of threatened species (Nicotra et al. 2010; Hale & Swearer 2016). Conservation efforts can be further strengthened by population genetics by estimating factors that may contribute

to extinction risk, including inbreeding, reduced genetic diversity, and low effective population size (Spielman et al. 2004). Various conservation-genetic techniques have been developed to study bee ecology and evolution (e.g., Woodard et al. 2015). Genotyping of microsatellites has proven particularly versatile (e.g., Charman et al. 2010; Lozier et al. 2011) and is a robust method for detecting genetic effects of recent habitat fragmentation, even in species with high gene flow (Williams et al. 2003).

Here, we investigate body size spatial structuring and population genetics in five bumble bee species across the greater Saint Louis, Missouri region: *Bombus auricomus*, *Bombus bimaculatus*, *Bombus griseocollis*, *Bombus impatiens*, and *Bombus pensylvanicus*. These species have experienced divergent population trends over the past two centuries in North America; *B. auricomus* and *B. pensylvanicus* have decreased relative abundance, while *B. impatiens*, *B. bimaculatus*, and *B. griseocollis* have experienced abundance increases (Hatfield et al. 2015). The International Union for Conservation of Nature (IUCN) Red List categorizes all of these species as "Least Concern" with stable population trends, except for *B. pensylvanicus*, which is listed as "Vulnerable" with a declining population trend (IUCN 2019). Recent data suggest a listing of "Critically Endangered" for *B. pensylvanicus* in Canada, following IUCN Red List criteria (MacPhail et al. 2019). By estimating population genetics using microsatellites and analyzing intraspecific spatial structure of body size, we provide an integrative, comparative assessment of conservation genetics and trait variation in a group of at-risk pollinating insects. We ask the following questions: (1) do these species exhibit intraspecific spatial structure in body size and, if so, (2) does this body size structure coincide with population genetic structure? We additionally estimate genetic

diversity, inbreeding, and colony density for these species throughout the greater Saint Louis region, as these factors can help inform conservation efforts. As anthropogenic changes to the biosphere continue to drive biodiversity loss, it is of paramount importance to understand functional trait variability and conservation genetics of groups at risk of extinction.

Material and Methods

Study Sites and Sampling

We sampled bumble bees in the greater Saint Louis, Missouri region in 2018, throughout the entire period of colony activity for each species. The five focal bumble bee species in this study (*B. auricomus*, *B. bimaculatus*, *B. griseocollis*, *B. impatiens*, *B. pensylvanicus*) can all be reliably found throughout this area (Camilo et al. 2018). We sampled bumble bees weekly from each of four sites: Calvary Cemetery (CC), EarthDance Farms (ED), Castlewood State Park (CW) (permission by Missouri Department of Natural Resources, Application for Research in Missouri State Parks 2018; Christopher Crabtree *personal communication*), and Shaw Nature Reserve (SNR) (Fig 1A). These sites occur along a gradient from Saint Louis city to an area west of Saint Louis, which follows a trend of decreasing human population density (number of people/km^2) with increased distance from Saint Louis (Fig 1B). As human population density is a commonly used metric for anthropogenic influence on the environment (e.g., Thompson & Jones 1999; Fontana et al. 2011), we consider our sites as occurring along an urban gradient, where sites occurring in localities with greater human population density are considered more urban (Fig 1B; see Supplemental Materials for density calculations and site descriptions). As

the minimum distance separating any two of these sites is greater than the typical dispersal distance of queen bumble bees (Lepais et al. 2010), we treat all conspecific bees per individual site as a putative subpopulation.

We opportunistically collected bees by hand-netting and immediately transferred them to individual ventilated vials. For all bees collected while actively foraging on a flower, we recorded the floral genus the bee was foraging on. We employed non-lethal sampling (Holehouse et al. 2003) and released bees following data collection. Before release, we identified bees to species and sex, removed a mid-leg tarsus from each bee and immediately stored it in 100% ethanol for microsatellite genotyping. For a subset of bees, we also measured thorax width using digital calipers [standard practice for measurements of bee body size (Cane 1987; Goulson 2010)] prior to release.

Microsatellite Genotyping

We performed DNA extraction and PCR amplification at the University of Missouri - St. Louis. Immediately prior to DNA extraction, we dried mid-leg tarsus samples and transferred each sample to a 96 well plate. In between samples, we immersed the forceps used for this work in 95% ethanol to prevent cross contamination. We followed a Chelex-based DNA extraction protocol (Walsh et al. 1991), whereby we added 150 μL Chelex 100 and 5 μL Proteinase K to each sample, and subsequently incubated samples in a Bio-Rad T100 Thermal Cycler with the following conditions: (1) 55°C for 1 h, (2) 99°C for 15 min, (3) 37°C for 1 min, and (4) 99°C for 15 min. Prior to PCR amplification, we stored extracted DNA samples at -20°C.

We genotyped each sample at 18 dye-labeled microsatellite loci (Estoup et al. 1995; Estoup et al. 1996; Funk et al. 2006; Stolle et al. 2009). Not all loci were successfully amplified or reliably scored within each species, so each species had its own complement of loci used for analyses (Table S1). We ran two multiplex PCRs per sample (i.e., plexes A and B), with six to nine microsatellite primers in each multiplex. Each multiplex reaction mixture contained 1 μL Chelex DNA extraction supernatant, 2 μL Promega 5x buffer, 0.56 μL $MgCl_2$ 25 mM, 0.6 μL dNTP, 0.2 μL bovine serum albumin, 0.08 μL Taq polymerase, 2.28-3.08 μL H_2O, and 0.045-0.400 μL of each primer. Each sample had a total reaction mixture volume of 10 μL, contained in a new well of a 96 well plate. We performed each PCR using a Bio-Rad T100 Thermal Cycler with the following conditions: (1) 95°C hot start, (2) initial denaturation at 95°C for 3.5 min, (3) 31 cycles of 95°C for 30 sec, 55°C (plex A) or 58°C (plex B) for 1.25 min, 72°C for 45 sec, and (4) final extension of 72°C for 15 min. Subsequently, we sent 2 μL of each PCR product to the University of Missouri DNA Core for fragment analysis, where DNA Core staff added formamide and an internal size standard (600 LIZ). We scored alleles using Geneious 11.0.4 with the Microsatellite Plugin (Kearse et al. 2012). Following microsatellite genotyping, we verified species identifications based on genetic signatures. Furthermore, we discarded from downstream genetic analyses all individuals and loci with 20% or greater genotyping failure per species.

Colony Density

Colony density (i.e., the number of colonies per subpopulation) is considered a measure of how well a given site supports a species (Geib et al. 2015). We estimated colony

density (N_c) for each subpopulation following methods described by Geib et al. (2015). Following these methods, N_c serves as a surrogate for effective population size (N_e), wherein the number of colonies per subpopulation is estimated based on genetic reconstructions of female sibships (Wang 2004). Prior to estimating N_c, we removed loci per species that had ≥25% null allele frequency following Chakraborty et al. (1992), using the R package PopGenReport version 2.0 (Gruber & Adamack 2014), and did not calculate N_c for any subpopulation with ≤15 successfully genotyped females. See Supplemental Materials for full methods of N_c calculations.

Population Genetic Analyses

We included only one randomly chosen sister per colony for population genetic analyses. After retaining one sister per colony, we checked loci for linkage disequilibrium (LD) using the R package Genepop '007 (Rousset 2008). If we found two or more loci to be in significant LD (p-value < 0.05), we retained only one of these loci for further genetic analyses. We tested individual loci for Hardy-Weinberg equilibrium (HWE) using the R package PopGenReport version 2.0 (Gruber & Adamack 2014).

Following these quality control measures, we calculated allelic richness (i.e., mean allele number per locus; *AR*) per subpopulation and global *AR* per species (i.e., species-level *AR* grouping samples across sites). As *AR* can be sensitive to variances in sample size, sample size rarefaction is the preferred method of standardizing *AR* for comparative studies (Leberg 2002). Prior to calculating *AR* values, we rarefied subpopulation sample sizes to the lowest subpopulation sample size across all five species, using the R package hierfstat (Goudet 2005). For global measures of *AR*, we

rarefied each species' sample size to the sample size of the species with the lowest overall sample size.

To assess genetic differentiation among intraspecific subpopulations, we calculated F_{ST} across all loci per species (Weir & Cockerham 1984) in FSTAT (version 2.9.4). To ensure that our data had sufficient statistical power to detect true genetic differentiation, we performed a power simulation per species with the program POWSIM (version 4.1), which tests the null hypothesis of no genetic differentiation between subpopulations, given different combinations of samples size, loci, and alleles (Ryman & Palm 2006). See Supplemental Materials for full power analysis methods.

We assessed each species for possible inbreeding by (1) calculating the inbreeding coefficient, F_{IS}, across all loci per species (Weir & Cockerham 1984) in FSTAT (version 2.9.4), and (2) inspecting males for diploidy. In bee populations, diploid male frequency increases with inbreeding due to increased rates of homozygosity at the complementary sex determination locus (Zayed & Packer 2001). To assess male diploidy, for each male bee we recorded whether each successfully genotyped locus was scored as homozygous or heterozygous. Following Darvill et al. (2006), we then recorded a male as diploid if three or more of his loci were scored as heterozygous. For calculations of F_{ST}, F_{IS}, and subpopulation *AR*, we removed all individuals from populations with <25 samples following our quality control measures (Hale et al. 2012). However, we did not remove individuals from populations with a low sample size for our calculations of global *AR*.

Body Size Variation Analyses

For all body size variation analyses, we included only one randomly chosen sister per colony and excluded all subpopulations that included ≤15 workers with thorax width measurements. Given our weekly sampling protocol across sites, these measurements collectively represent body size variation across each species' entire period of colony activity. To determine whether our focal bumble bee species exhibit intraspecific spatial structure in body size, we compared intraspecific subpopulations for significantly different average body sizes. We first ran an analysis of variance (ANOVA) with thorax width as the response variable, and site and species as categorical predictors. Subsequently, we ran contrasts between least squares means for each unique pairing of intraspecific subpopulations. We used a Bonferroni corrected α-value to determine statistical significance of these contrasts. To compute these contrasts, we used the R package lsmeans version 2.30 (Lenth 2016).

Results

Sampling and Genotyping

Across all species and sites, we collected 839 bees; 774 females and 65 males. Sample sizes are variable across species and sites (Tables S2 and S3), ranging from conspecific bees being absent or found in low abundance to upwards of 70 conspecific bees being collected at a site. Following all genotyping quality control measures, each species had a minimum of 10 loci used in population genetic analyses (Fig S1; Table S1). A description of these quality control results and loci retained per species can be found in the Supplemental Materials.

Colony Density

Each species has variable colony densities across sites. N_c ranges from a minimum of 19.6 (*B. pensylvanicus* at ED) to a maximum of 98.7 (*B. bimaculatus* at CW). We could not calculate N_c for *B. auricomus*, *B. griseocollis*, or *B. pensylvanicus* at CW, and for *B. bimaculatus* at CC, due to the number of successfully genotyped females ≤15 for these subpopulations. See Table 1 for N_c estimates per subpopulation and Table S2 for additional breakdown of how N_c estimates were calculated.

Population Genetic Analyses

Throughout the greater Saint Louis region, genetic differentiation between intraspecific subpopulations is low to absent in each species, with $F_{ST} \leq 0.002$ in each species and all 95% CIs including zero. Each power simulation revealed statistical power >0.99 for detecting an F_{ST}=0.05 using both chi-square and Fisher's exact tests. Accordingly, our sampling protocol had a >99% probability of detecting true F_{ST} values of 0.05. F_{IS} values are more variable, ranging from a minimum of 0.023 (*B. bimaculatus*) to a maximum of 0.151 (*B. griseocollis*). Zero is only included in the F_{IS} 95% CI of *B. bimaculatus*. All males collected are haploid, except in *B. griseocollis* for which 21 of 25 collected males (84%) are diploid (i.e., ≥3 loci scored as heterozygous) (Table S3). Global *AR* calculations were rarefied to a sample size of 88 per species, following *B. pensylvanicus* having the lowest overall sample size (i.e., 88 female genotypes retained * 2 alleles/female = 176 alleles). Subpopulation *AR* calculations were rarefied to a subpopulation size of 28, as the subpopulation included in genetic analyses with the lowest sample size was *B. pensylvanicus* at CC (i.e., 28 female genotypes retained * 2

alleles/female = 56 alleles). *AR* varies interspecifically (i.e., between species' global *AR* values) and between intraspecific loci (Fig S1). *Bombus pensylvanicus* has the lowest *AR* across all species (global *AR* = 6.29 $\pm$ 1.42 SE) and *B. impatiens* has the highest *AR* (global *AR* = 10.24 $\pm$ 2.21 SE). We could not calculate F_{ST}, F_{IS}, and site-specific *AR* for *B. auricomus*, *B. griseocollis*, or *B. pensylvanicus* at CW, *B. bimaculatus* at CC, and *B. pensylvanicus* at ED due to <25 genotypes remaining in each of these subpopulations following our quality control measures. See Table 1 for these population genetic statistics across sites and species.

Body Size Variation Analyses

We find evidence for spatial structuring of intraspecific body size for bumble bees in the greater Saint Louis region. Our full ANOVA shows significant effects of species, site, and their interaction on worker thorax width (species, site, and species*site all $p < 0.0001$). Average body size significantly differs between intraspecific subpopulations of *B. impatiens* and *B. pensylvanicus*. Specifically, for *B. impatiens*, worker body size is larger on average at CC than at CW (contrast of least square means $p < 0.0001$) (Fig 2). For *B. pensylvanicus* worker body size is larger on average at SNR than at CC or ED (both contrasts of least square means $p < 0.0001$) (Fig 2). No other species shows significant spatial structuring of average body size (all contrasts of least square means $p > 0.006$) (Table S4). The Bonferroni adjusted α-value used for determining statistical significance between average body size contrasts is α=0.00278 (i.e., 0.05/18 contrasts) (Table S4). We did not include *B. auricomus*, *B. griseocollis*, or *B. pensylvanicus* at CW, and *B. bimaculatus* at CC in these analyses due to $\leq$15 workers having thorax width

measurements at these subpopulations. See Table S5 for all worker thorax width sample sizes and body size means per subpopulation.

Discussion

Studying five bumble bee species across four sites in the greater Saint Louis region, we find evidence for intraspecific spatial structuring of body size, despite genetic homogeneity among subpopulations. Specifically, two species, *B. impatiens* and *B. pensylvanicus*, exhibit spatial body size structuring; however, the direction of this spatial structuring is not consistent between species (i.e., sites with increased urbanization are associated with larger *B. impatiens* and smaller *B. pensylvanicus*). As our study sites occur along an urban gradient from the city of Saint Louis to a rural area west of the city (Fig 1), these results suggest that human-modified environments can drive body size differences between intraspecific subpopulations of pollinating insects. This work builds upon a body of literature documenting the functional trait variability (e.g., Albert et al. 2010; Brousseau et al. 2018) and conservation genetics (e.g., Charman et al. 2010; Geib et al. 2015) of groups at risk of extinction, while demonstrating that urbanization can structure bee communities intraspecifically.

Two non-mutually exclusive explanations may account for the observed intraspecific spatial structuring of body size: phenotypic plasticity or local adaptation. We argue that this result is likely a consequence of plasticity as opposed to adaptation for two primary reasons. First, we do not find evidence for genetic structure in any of our studied species; i.e., all F_{ST} values are low (all $F_{ST} \leq 0.002$; Table 1) and our power analyses indicate that our data had sufficient statistical power to detect true genetic differentiation,

if it were present. This suggests high rates of intraspecific gene flow throughout the greater Saint Louis region. High rates of gene flow often limit subpopulations from adapting to their local environments, by homogenizing traits throughout a metapopulation (Fitzpatrick et al. 2017). Second, body size is an exceptionally plastic trait in bumble bees, with 10-fold differences in body size occurring among highly related intra-colony workers (r=0.75) (Couvillon and Dornhaus 2009; Goulson 2010). Plasticity can shield a population from local adaptation by moving the population toward an adaptive peak, thus enabling persistence in a changed environment without adaptive genetic change (Price et al. 2003). Accordingly, the lack of genetic structure, coupled with the known plasticity of bumble bee body size, support the observed body size spatial structuring being a result of plastic responses to local environments, as opposed to adaptive genetic divergence. However, we cannot definitively rule out the possibility of local adaptation; in rare cases, subpopulations can become locally adapted even while gene flow is maintained (e.g., Liu et al. 2016). It is possible that recent habitat fragmentation has induced strong differential selection between subpopulations, though sufficient time has not passed for population genetics to reflect this. Although, this may be an unlikely explanation of our results, as microsatellites can document genetic effects of recent fragmentation in species of pollinating insects with high gene flow (Williams et al. 2003).

Several environmental factors may drive this observed spatial structuring of body size. In bumble bees, worker larvae fed a higher quality diet or at higher rates develop into larger adults (Pendrel and Plowright 1981; Sutcliffe and Plowright 1988). It is possible that body size spatial structuring results from differences in nutritional quality and/or quantity among sites, whereby large size is promoted by high nutritional

quality/quantity (or small size results from a constraint of low nutritional quality/quantity). While we did not directly quantify nutrition in this study, our data suggest this may be a likely explanation of our results. First, in all cases where average body size significantly differed between intraspecific sites (i.e., between CC and CW for *B. impatiens* and between SNR and both CC and ED for *B. pensylvanicus*; Fig 2), conspecific females were observed foraging from a higher richness of floral genera at the sites where body size was larger (Table S6). As bees often optimize nutritional intake by foraging from a variety of floral species (Vaudo et al. 2015), this may correspond to bees having more balanced diets at sites with a higher richness of exploitable floral genera. Second, at all sites where average body size was larger intraspecifically, not only were more floral genera exploited, but colony density was higher as well (Table 1). Numerous studies indicate that colony success is dependent on nutritional availability at a site (e.g., Woodard & Jha 2017; Vaudo et al. 2015). Thus, the higher colony density observed at sites with a greater richness of exploited floral genera supports the idea that these sites conferred greater nutritional quality and/or quantity. It is notable that the greatest magnitude of body size spatial structuring was observed in *B. pensylvanicus*. Numerous reports have suggested *B. pensylvanicus* is the species most at risk of extinction among those studied (IUCN 2019; MacPhail et al. 2019) and our finding of *B. pensylvanicus* having the lowest genetic diversity among bumble bee species throughout the greater Saint Louis region (lowest *AR* in both 2018 and 2017; see Supplemental Materials for description of 2017 population genetics; Table 1; Table S7) reaffirms these reports. Interestingly, however, at SNR - the site where *B. pensylvanicus* was largest intraspecifically and was found feeding from a comparatively high number of floral

genera - *B. pensylvanicus* colony density was highest both intraspecifically (i.e., across sites) and interspecifically (i.e., highest interspecific colony density at SNR in both 2018 and 2017) (Tables 1 and S7). This may suggest that sites with high floral species richness provide robust support to *B. pensylvanicus* populations, thus indicating the potential role that floral enrichment can play in supporting populations of threatened bumble bee species.

The importance of investigating functional trait diversity of threatened species has been increasingly recognized as conservation program efficacy depends on environmental effects on the development and expression of phenotype (e.g., Watters et al. 2003; Keller & Waller 2002) and plasticity is a primary response of species to global change (e.g., Wong & Candolin 2015). The spatial structuring of body size we observed suggests that human-modified environments can induce landscape-level structuring of developmentally plastic functional traits. Conservation programs should be cognizant of when traits are developmentally, but irreversibly, plastic. For example, Lema & Nevitt (2006) document that pupfish (*Cyprinodon* spp.) exhibit a developmentally plastic small body size as a result of high water temperature and low food availability. They suggest that management programs consider this by captively breeding pupfish in similar conditions to the population they will be reintroduced to, so that large individuals with high dietary requirements are not reintroduced into a food-limited environment (Lema & Nevitt 2006). Similarly, if the spatial structuring of body size we observed resulted from nutritional differences among sites, this may suggest that spatial structuring of bumble bee body size can be used to indicate variance in environmental quality, with subpopulations with relatively smaller average body sizes being targeted for floral

enrichment. However, alternative explanations may underlie the observed spatial structuring of body size. For example, spatial heterogeneity in environmental contaminants could differentially expose subpopulations to pollutants, which may have downstream effects on foraging behavior (Sivakoff & Gardiner 2017) and the development of adult body size (Whitehorn et al. 2018). Alternatively, in urban areas, increased metabolic demands imposed by the urban-heat-island (UHI) effect are expected to drive shifts toward smaller body size in certain taxa (Merckx et al. 2018). The direction of *B. pensylvanicus* body size spatial structuring across the urban gradient follows the predicted direction under the UHI effect, analogous to the Brazilian stingless bee, *Melipona fasciculata* (Oliveira et al. 2019); however, the spatial structuring of *B. impatiens* body size follows the opposite pattern. Furthermore, in taxa where body size positively correlates with dispersal distance, habitat fragmentation may drive increased body size to promote movement of individuals between habitat patches (Warzecha et al. 2016; Merckx et al. 2018). However, similar to the UHI effect, the contrasting directions of spatial body size structuring found for *B. impatiens* and *B. pensylvanicus* complicate this as a likely explanation for our results. As myriad environmental factors may interact to affect spatial structuring of bumble bee body size, to gain a comprehensive understanding of this system, future studies should directly quantify nutrition, environmental factors, and fragmentation across subpopulations.

Our results exemplify the importance of simultaneously investigating functional trait variability and conservation genetics of groups at risk of extinction. While *B. griseocollis* does not exhibit spatial structuring of body size throughout the greater Saint Louis area, we find evidence that *B. griseocollis* is potentially inbred in this region.

Bombus griseocollis had the highest inbreeding coefficient (F_{IS}) and the second lowest global *AR* among the studied species (Table 1) and 84% of sampled *B. griseocollis* males were diploid in 2018 (Table S3). In haplodiploid bees, males develop via either (1) parthenogenesis, in which hemizygosity at the sex-determining locus produces a viable, haploid male, or (2) a fertilized egg, in which homozygosity at the sex-determining locus produces a sterile, diploid male (Zayed & Packer 2001). As inbreeding promotes an increased proportion of homozygosity (Keller & Waller 2002), diploid males may occur at higher frequencies in inbred haplodiploid populations. While additional sampling in the Midwest is needed, in replicate years and populations, these results suggest relatively high rates of inbreeding in Saint Louis *B. griseocollis* populations, despite *B. griseocollis* being broadly distributed and abundant throughout much of the United States (Strange & Tripodi 2019) and listed as "Least Concern" by the IUCN Red List (IUCN 2019). Indeed, future research on *B. griseocollis* populations is needed, as understanding why the observed rates of *B. griseocollis* male diploidy are so high will be critical to implementing effective conservation programs. Collectively, our results indicate the utility of simultaneously investigating phenotypic and genetic variation of threatened species, as phenotypic and genetic signatures of population stability can occur independently of one another and together provide a more complete understanding of population stability across heterogeneous landscapes.

Conclusions

The conservation of threatened species is strengthened by integrative assessments of functional trait variability and population genetics. We document that bumble bees can

exhibit intraspecific body size spatial structuring, despite subpopulations being genetically homogenous. These results suggest that urbanization can induce landscape-level structuring of functional traits that are developmentally plastic, potentially due to nutritional differences across sites. We additionally find evidence that (1) *B. pensylvanicus* has comparatively low genetic diversity, reaffirming findings from previous studies (e.g., Lozier et al. 2011; Cameron et al. 2011) and (2) *B. griseocollis* is inbred in the greater Saint Louis region. Collectively, these results are informative for the development of bumble bee conservation programs and add to a growing body of literature on how threatened species are affected by human-modified environments. Anthropogenic effects on the environment are threatening approximately 1 million extant species with extinction (IPBES 2019). To aid the conservation of these at-risk groups, it is imperative to concurrently assess genetic and phenotypic variability within species at a variety of spatial scales.

Acknowledgements

Our gratitude is extended to Joyce Knoblett of the USDA-ARS Pollinating Insect - Biology, Management, Systematics Research Unit for assisting with the molecular work of our 2017 samples. We thank Shaw Nature Reserve, EarthDance Farms, Calvary Cemetery, Castlewood State Park, and their employees for access to these sites, which made this work possible. In particular, we acknowledge James Trager (Shaw Nature Reserve) and Eric Stevens (EarthDance) who provided helpful insight to learning about the biota of their respective sites. We thank the Missouri Department of Natural Resources for granting us permission to sample at Castlewood State Park. We thank

Patricia Parker and various members of ASD's laboratory for their valuable feedback and discussion on this project. This study was funded by grants from the Whitney R. Harris World Ecology Center, Trans World Airlines, and the Webster Groves Nature Study Society to MWA. MWA was partially supported by the Raven Fellowship and the Graduate Dissertation Fellowship through the University of Missouri - St. Louis.

Funding

This work was funded by grants to MWA from the Whitney R. Harris World Ecology Center, Trans World Airlines, and the Webster Groves Nature Study Society. MWA was partially funded by the Raven Fellowship (spring 2019) and the Graduate Dissertation Fellowship (fall 2019, spring 2020) through the University of Missouri - St. Louis.

References

Albert, C.H., Thuiller, W., Yoccoz, N.G., Soudant, A., Boucher, F., Saccone, P., & Lavorel, S. 2010. Intraspecific functional variability: Extent, structure and sources of variation. Journal of Ecology 98:604-613.

Arbetman, M.P., Gleiser, G., Morales, C.L., Williams, P., & Aizen, M.A. 2017. Global decline of bumblebees is phylogenetically structured and inversely related to species range size and pathogen incidence. Proceedings of the Royal Society B: Biological Sciences 284:1859.

Austin, M.W. & Dunlap, A.S. 2019. Intraspecific variation in worker body size makes North American bumble bees (*Bombus* spp.) less susceptible to decline. The American Naturalist 194:381-394.

Banaszak-Cibicka, W. & Żmihorski, M. 2012. Wild bees along an urban gradient: winners and losers. Journal of Insect Conservation 16:331-343.

Bartomeus, I., Ascher, J.S., Gibbs, J., Danforth, B.N., Wagner, D.L., Hedtke, S.M., & Winfree, R. 2013. Historical changes in northeastern US bee pollinators related to shared ecological traits. Proceedings of the National Academy of Sciences of the United States of America 110:4656-4660.

Bommarco, R., Biesmeijer, J.C., Meyer, B., Potts, S.G., Pöyry, J., Roberts, S.P.M., Steffan-Dewenter, I., & Öckinger, E. 2010. Dispersal capacity and diet breadth modify the response of wild bees to habitat loss. Proceedings of the Royal Society B: Biological Sciences 277:2075-2082.

Brousseau, P.M., Gravel, D., & Handa, I.T. 2018. On the development of a predictive functional trait approach for studying terrestrial arthropods. Journal of Animal Ecology 87:1209-1220.

Burkle, L.A., Marlin, J.C., & Knight, T.M. 2013. Plant-pollinator interactions over 120 years: Loss of species, co-occurrence, and function. Science 340:1611-1615.

Cameron, S.A., Lozier, J.D., Strange, J.P., Koch, J.B., Cordes, N., Solter, L.F., & Griswold, T.L. 2011. Patterns of widespread decline in North American bumble bees. Proceedings of the National Academy of Sciences of the United States of America 108:662-667.

Camilo, G.R., Muñiz, P.A., Arduser, M.S., & Spevak, E.M. 2018. A checklist of the bees (Hymenoptera: Apoidea) of St. Louis, Missouri, USA. Journal of the Kansas Entomological Society 90:175-188.

Cane, J.H. 1987. Estimation of bee size using intertegular span (Apoidea). Journal of the Kansas Entomological Society:145-147.

Chakraborty, R., Andrade, M.D.E., Daiger, S.P., & Budowle, B. 1992. Apparent heterozygote deficiencies observed in DNA typing data and their implications in forensic applications. Annals of Human Genetics 56: 45–57.

Charman, T.G., Sears, J., Green, R.E., & Bourke, A.F.G. 2010. Conservation genetics, foraging distance and nest density of the scarce Great Yellow Bumblebee (*Bombus distinguendus*). Molecular Ecology 19:2661-2674.

Colla, S.R., Gadallah, F., Richardson, L., Wagner, D., & Gall, L. 2012. Assessing declines of North American bumble bees (*Bombus* spp.) using museum specimens. Biodiversity and Conservation 21:3585-3595.

Corlett, R.T. 2015. The Anthropocene concept in ecology and conservation. Trends in Ecology and Evolution 30:36-41.

Couvillon, M.J. & Dornhaus, A. 2009. Location, location, location: Larvae position inside the nest is correlated with adult body size in worker bumble-bees (*Bombus impatiens*). Proceedings of the Royal Society B: Biological Sciences 276:2411-2418.

Couvillon, M.J. & Dornhaus, A. 2010. Small worker bumble bees (*Bombus impatiens*) are hardier against starvation than their larger sisters. Insectes Sociaux 57:193-197.

Crispo, E. 2008. Modifying effects of phenotypic plasticity on interactions among natural selection, adaptation and gene flow. Journal of Evolutionary Biology 21:1460-1469.

Darvill, B., Ellis, J.S., Lye, G.C., & Goulson, D. 2006. Population structure and inbreeding in a rare and declining bumblebee, *Bombus muscorum* (Hymenoptera: Apidae). Molecular Ecology 15:601-611.

Estoup, A., Scholl, A., Pouvreau, A., & Solignac, M. 1995. Monoandry and polyandry in bumble bees (Hymenoptera; Bombinae) as evidenced by highly variable microsatellites. Molecular Ecology 4:89-94.

Estoup, A., Solignac, M., Cornuet, J.M., Goudet, J., & Scholl, A. 1996. Genetic differentiation of continental and island populations of *Bombus terrestris* (Hymenoptera: Apidae) in Europe. Molecular Ecology 5:19-31.

Fitzpatrick, S.W., Handelsman, C.A., Torres-Dowdall, J., Ruell, E.W., Broder, E.D., Kronenberger, J.A., Reznik, D.N., Ghalambor, C.K., Angeloni, L.M., & Funk, W.C. 2017. Gene flow constrains and facilitates genetically based divergence in quantitative traits. Copeia 105:462-474.

Fontana, C.S, Burger, M.I. & Magnusson, W.E. 2011. Bird diversity in a subtropical South-American City: effects of noise levels, arborisation and human population density. Urban Ecosystems 14:341-360.

Fraser, D.J. & Bernatchez, L. 2001. Adaptive evolutionary conservation: towards a unified concept for defining conservation units. Molecular Ecology 10:2741-2752.

Funk, C.R., Schmid-Hempel, R., & Schmid-Hempel, P. 2006. Microsatellite loci for *Bombus* spp. Molecular Ecology Notes 6:83-86.

Geib, J.C., Strange, J.P., & Galen, A. 2015. Bumble bee nest abundance, foraging distance, and host-plant reproduction: Implications for management and conservation. Ecological Applications 25:768-778.

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

Goulson, D. 2010. Bumblebees: Behaviour, Ecology, and Conservation. Oxford University Press.

Goulson, D., Nicholls, E., Botías, C., & Rotheray, E.L. 2015. Bee declines driven by combined stress from parasites, pesticides, and lack of flowers. Science 347:1255957.

Greenleaf, S.S., Williams, N.M., Winfree, R., & Kremen, C. 2007. Bee foraging ranges and their relationship to body size. Oecologia 153:589-596.

Grixti, J.C., Wong, L.T., Cameron, S.A., & Favret, C. 2009. Decline of bumble bees (*Bombus*) in the North American Midwest. Biological Conservation 142:75-84.

Gruber, B. & Adamack, A.T. 2014. PopGenReport: simplifying basic population genetic analyses in R. Methods in Ecology and Evolution 5:384-387.

Haddaway, N.R., Mortimer, R.J.G., Christmas, M., Grahame, J.W., & Dunn, A.M. 2012. Morphological diversity and phenotypic plasticity in the threatened British white-clawed crayfish (*Austropotamobius pallipes*). Aquatic Conservation: Marine and Freshwater Ecosystems 22:220-231.

Hale, M.L., Burg, T.M., & Steeves, T.E. 2012. Sampling for microsatellite-based population genetic studies: 25 to 30 Individuals per population is enough to accurately estimate allele frequencies. PLOS ONE 7:e45170.

Hale, R. & Swearer, S.E. 2016. Ecological traps: current evidence and future directions. Proceedings of the Royal Society B: Biological Sciences 283:1-8.

Hall, D.M., Camilo, G.R., Tonietto, R.K., Ollerton, J., Ahrné, K., Arduser, M., Ascher, J.S., Baldock, K.C.R., Fowler, R., Frankie, G., Goulson, D., Gunnarsson, B., Hanley, M.E., Jackson, J.I., Langellotto, G., Lowenstein, D., Minor, E.S., Philpitt, S.M., Potts, S.G., Sirohi, M.H., Spevak, E.M., Stone, G.N., & Threlfall, C.G. 2017. The city as a refuge for insect pollinators. Conservation Biology 31:24-29.

Hatfield, R., Colla, S., Jepsen, S., Richardson, L., Thorp, R., & Jordan, S. 2015. IUCN assessments of North American *Bombus* spp. http://www.xerces.org/.

Holehouse, K.A., Hammond, R.L., & Bourke, A.F.G. 2003. Non-lethal sampling of DNA from bumble bees for conservation genetics. Insectes Sociaux 50:277-285.

Husemann, M., Zachos, F.E., Paxton, R.J., & Habel, J.C. 2016. Effective population size in ecology and evolution. Heredity 117:191-192.

IPBES. 2019. Summary for policymakers of the global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. IPBES secretariat, Bonn, Germany.

IUCN. 2019. The IUCN Red List of Threatened Species. Accessed 18 Dec. 2019. https://www.iucnredlist.org

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P., & Drummond, A. 2012. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28:1647-1649.

Keller, L.F. & Waller, D.M. 2002. Inbreeding effects in wild populations. Trends in Ecology & Evolution 17:230-241.

Leberg, P.L. 2002. Estimating allelic richness: Effects of sample size and bottlenecks. Molecular Ecology 11:2445-2449.

Lema, S.C. & Nevitt, G.A. 2006. Testing an ecophysiological mechanism of morphological plasticity in pupfish and its relevance to conservation efforts for endangered Devils Hole pupfish. Journal of Experimental Biology 209:3499-3509.

Lenth, R.V. 2016. Least-Squares Means: The R Package lsmeans. Journal of Statistical Software 69:1-33.

Lepais, O., Darvill, B., O'Connor, S., Osborne, J.L., Sanderson, R.A., Cussans, J., Goffe, L., Goulson, D. 2010. Estimation of bumblebee queen dispersal distances using sibship reconstruction method. Molecular Ecology 19:819-831.

Liu, B.-J., Zhang, B.-D., Xue, D.-X., Gao, T.-X., & Liu, J.-X. 2016. Population structure and adaptive divergence in a high gene flow marine fish: The small yellow croaker (*Larimichthys polyactis*). PLOS ONE 11:e0154020.

Lozier, J.D., Strange, J.P., Stewart, I.J., & Cameron, S.A. 2011. Patterns of range-wide genetic variation in six North American bumble bee (Apidae: *Bombus*) species. Molecular Ecology 20:4870-4888.

MacPhail, V.J., Richardson, L.L., & Colla, S.R. 2019. Incorporating citizen science, museum specimens, and field work into the assessment of extinction risk of the American Bumble bee (*Bombus pensylvanicus* De Geer 1773) in Canada. Journal of Insect Conservation 23:597-611.

Merckx, T., Kaiser, A., & Van Dyck, H. 2018. Increased body size along urbanization gradients at both community and intraspecific level in macro-moths. Global Change Biology 24:3837-3848.

Nicotra, A.B., Atkin, O.K., Bonser, S.P., Davidson, A.M., Finnegan, E.J., Mathesius, U., Poot, P., Purugganan, M.D., Richards, C.L., Valladares, F., van Kleunen, M. 2010. Plant phenotypic plasticity in a changing climate. Trends in Plant Science 15:684-692.

Oliveira, M.O., Brito, T.F., Campbell, A.J., & Contrera, F.A. L. 2019. Body size and corbiculae area variation of the stingless bee *Melipona fasciculata* Smith, 1854 (Apidae, Meliponini) under different levels of habitat quality in the eastern Amazon. Entomologia Generalis 39:45-52.

Peat, J., Tucker, J., & Goulson, D. 2005. Does intraspecific size variation in bumblebees allow colonies to efficiently exploit different flowers? Ecological Entomology 30:176-181.

Pendrel, B.A. & Plowright, R.C. 1981. Larval feeding by adult bumble bee workers (Hymenoptera: Apidae). Behavioral Ecology and Sociobiology 8:71-76.

Price, T.D., Qvarnström, A., & Irwin, D.E. 2003. The role of phenotypic plasticity in driving genetic evolution. Proceedings of the Royal Society B: Biological Sciences 270:1433-1440.

Rousset, F. 2008. genepop'007: a complete re-implementation of the genepop software for Windows and Linux. Molecular Ecology Resources 8:103-106.

Ryman, N. & Palm, S. 2006. POWSIM: a computer program for assessing statistical power when testing for genetic differentiation. Molecular Ecology Notes 6:600-602.

Sivakoff, F.S. & Gardiner, M.M. 2017. Soil lead contamination decreases bee visit duration at sunflowers. Urban Ecosystems 20:1221-1228.

Spaethe, J. & Weidenmüller, A. 2002. Size variation and foraging rate in bumblebees (*Bombus terrestris*). Insectes Sociaux 49:142-146.

Spielman, D., Brook, B.W., & Frankham, R. 2004. Most species are not driven to extinction before genetic factors impact them. Proceedings of the National Academy of Sciences of the United States of America 101:15261-15264.

Strange, J.P. & Tripodi, A.D. 2019. Characterizing bumble bee (*Bombus*) communities in the United States and assessing a conservation monitoring method. Ecology and Evolution 9:1061-1069.

Stolle, E., Rohde, M., Vautrin, D., Solignac, M., Schmid-Hempel, P., Schmid-Hempel, R., & Moritz, R.F.A. 2009. Novel microsatellite DNA loci for *Bombus terrestris* (Linnaeus, 1758). Molecular Ecology Resources 9:1345-1352.

Sutcliffe, G.H. & Plowright, R.C. 1988. The effects of food supply on adult size in the bumble bee *Bombus terricola* Kirby (Hymenoptera: Apidae). The Canadian Entomologist 120:1051-1058.

Thogmartin, W.E., Wiederholt, R., Oberhauser, K., Drum, R.G., Diffendorfer, J.E., Altizer, S., Taylor, O.R., Pleasants, J., Semmens, D., Semmens, B., & Erickson, R. 2017. Monarch butterfly population decline in North America: identifying the threatening processes. Royal Society Open Science 4:170760.

Thompson, K. & Jones, A. 1999. Human population density and prediction of local plant extinction in Britain. Conservation Biology 13:185-189.

Vaudo, A.D., Tooker, J.F., Grozinger, C.M., & Patch, H.M. 2015. Bee nutrition and floral resource restoration. Current Opinion in Insect Science 10:133-141.

Violle, C., Navas, M.-L., Vile, D., Kazakou, E., Fortunel, C., Hummel, I., & Garnier, E. 2007. Let the concept of trait be functional! Oikos 116:882-892.

Vitousek, P.M. 1997. Human domination of Earth's ecosystems. Science 277:494-499.

Wagner, D.L. 2020. Insect declines in the Anthropocene. Annual Review of Entomology 65.

Walsh, P.S., Metzger, D.A., & Higuchi, R. 1991. Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. Biotechniques 10:506-513.

Wang, J. 2004. Sibship reconstruction from genetic data with typing errors. Genetics 166:1963-1979.

Warzecha, D., Diekötter, T., Wolters, V., & Jauker, F. 2016. Intraspecific body size increases with habitat fragmentation in wild bee pollinators. Landscape Ecology 31:1449-1455.

Watters, J.V, Lema, S.C., & Nevitt, G.A. 2003. Phenotype management: a new approach to habitat restoration. Biological Conservation 112:435-445.

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

Whitehorn, P.R., Norville, G., Gilburn, A., & Goulson, D. 2018. Larval exposure to neonicotinoid imidacloprid impacts adult size in the farmland butterfly *Pieris brassicae*. PeerJ 6:e4772.

Williams, B.L., Brawn, J.D., & Paige, K.N. 2003. Landscape scale genetic effects of habitat fragmentation on a high gene flow species: *Speyeria idalia* (Nymphalidae). Molecular Ecology 12:11-20.

Wilson, C.J. & Jamieson, M.A. 2019. The effects of urbanization on bee communities depends on floral resource availability and bee functional traits. PLOS ONE 14:e0225852.

Wong, B.B.M. & Candolin, U. 2015. Behavioral responses to changing environments. Behavioral Ecology 26:665-673.

Woodard, S.H., Lozier, J.D., Goulson, D., Williams, P.H., Strange, J.P., & Jha, S. 2015. Molecular tools and bumble bees: Revealing hidden details of ecology and evolution in a model system. Molecular Ecology 24:2916-2936.

Woodard, S.H. & Jha, S. 2017. Wild bee nutritional ecology: predicting pollinator population dynamics, movement, and services from floral resources. 21:83-90.

Zayed, A. & Packer, L. 2001. High levels of diploid male production in a primitively eusocial bee (Hymenoptera: Halictidae). Heredity 87:631-636.

Tables

Table 1. Population genetic statistics and colony densities for bumble bees (*Bombus* spp.) in the greater Saint Louis region. Allelic richness (*AR*), calculated as mean allele number across loci, is calculated per species for each site and combining all sites (i.e., global *AR*). F_{ST} describes population genetic differentiation. F_{IS} is the inbreeding coefficient. N_c is colony density. All populations with <25 successfully genotyped individuals following quality control measures were removed from population genetic analyses. N_c was not calculated for populations with ≤15 successfully genotyped females. SE = standard error, 95% CI = 95% confidence interval, CC = Calvary Cemetery, CW = Castlewood State Park, ED = EarthDance Farms, SNR = Shaw Nature Reserve.

Species	*AR* (SE) per Site				Global *AR* (SE)	F_{ST} (95% CI)	F_{IS} (95% CI)	Colony Density			
	CC	CW	ED	SNR				CC	CW	ED	SNR
B. auricomus	7.24 (1.54)	-	7.05 (1.53)	7.41 (1.67)	8.94 (1.92)	0.002 (-0.004 - 0.010)	0.075 (0.031 - 0.124)	84.0	-	55.4	62.5
B. bimaculatus	-	7.25 (1.21)	7.12 (1.32)	7.24 (1.39)	9.25 (1.63)	0.000 (-0.002 - 0.002)	0.023 (-0.028 - 0.072)	-	98.7	69.0	55.3
B. griseocollis	6.76 (1.57)	-	6.43 (1.58)	6.52 (1.65)	8.61 (2.05)	0.002 (-0.002 - 0.008)	0.151 (0.090 - 0.203)	67.5	-	88.2	51.0
B. impatiens	7.77 (1.80)	8.20 (1.98)	7.76 (1.85)	8.15 (1.69)	10.24 (2.21)	0.001 (-0.001 - 0.004)	0.068 (0.025 - 0.122)	74.5	61.5	96.5	55.2
B. pensylvanicus	5.57 (1.32)	-	-	4.92 (1.13)	6.29 (1.42)	-0.003 (-0.008 - 0.002)	0.070 (0.008 - 0.140)	43.1	-	19.6	67.3

Figures

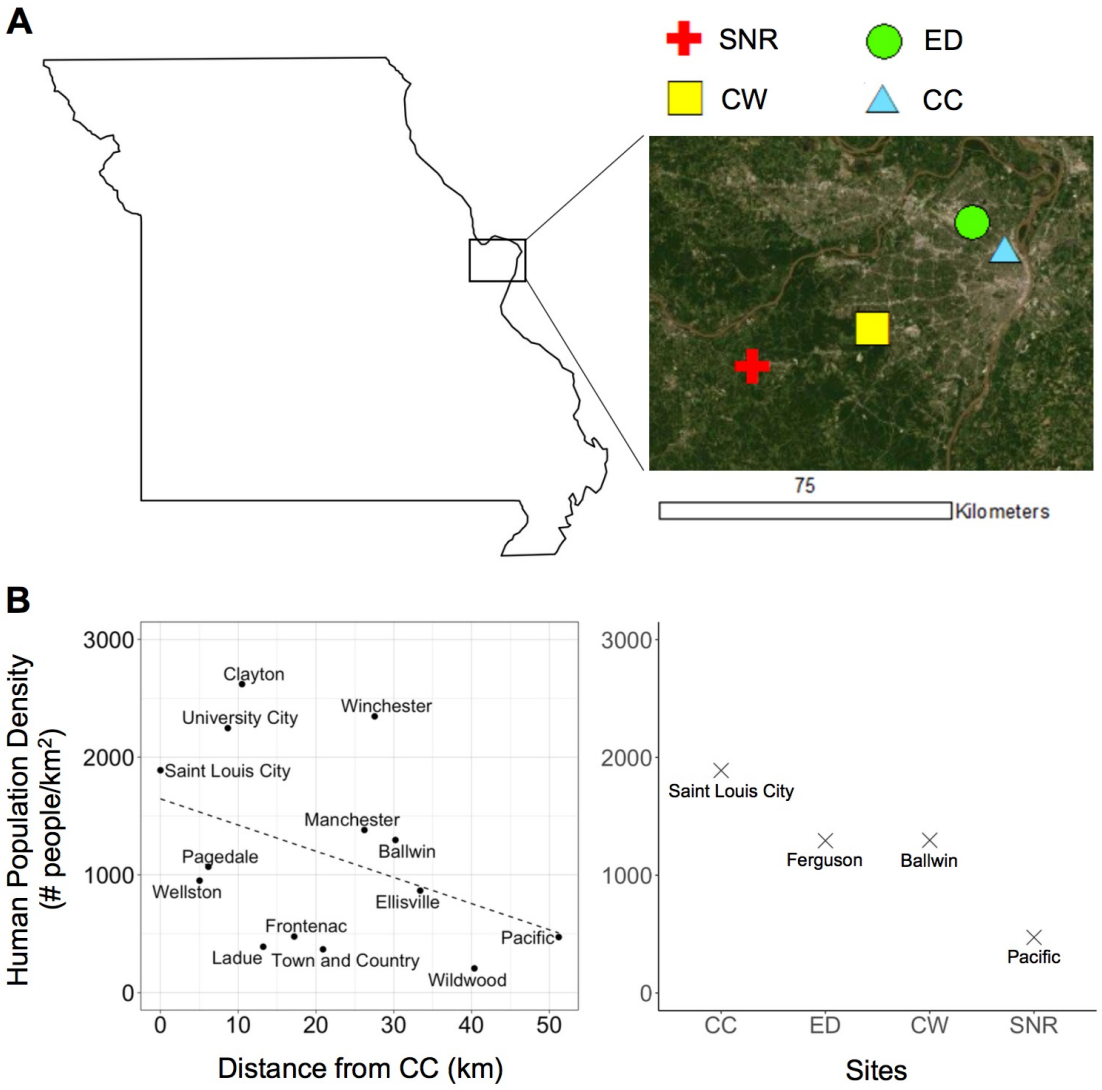

Figure 1. (A) Map of sampling locations. CC = Calvary Cemetery, CW = Castlewood State Park, ED = EarthDance Farms, SNR = Shaw Nature Reserve. (B) Human population density per locality. *Left panel:* Urban gradient depicted by human population density per locality from CC (Saint Louis City, MO) to SNR (Pacific, MO). Distance from CC is the distance from CC to the approximate midpoint of a locality that occurs along the trajectory from CC to SNR. *Right panel:* Human population density of each locality where a site is located.

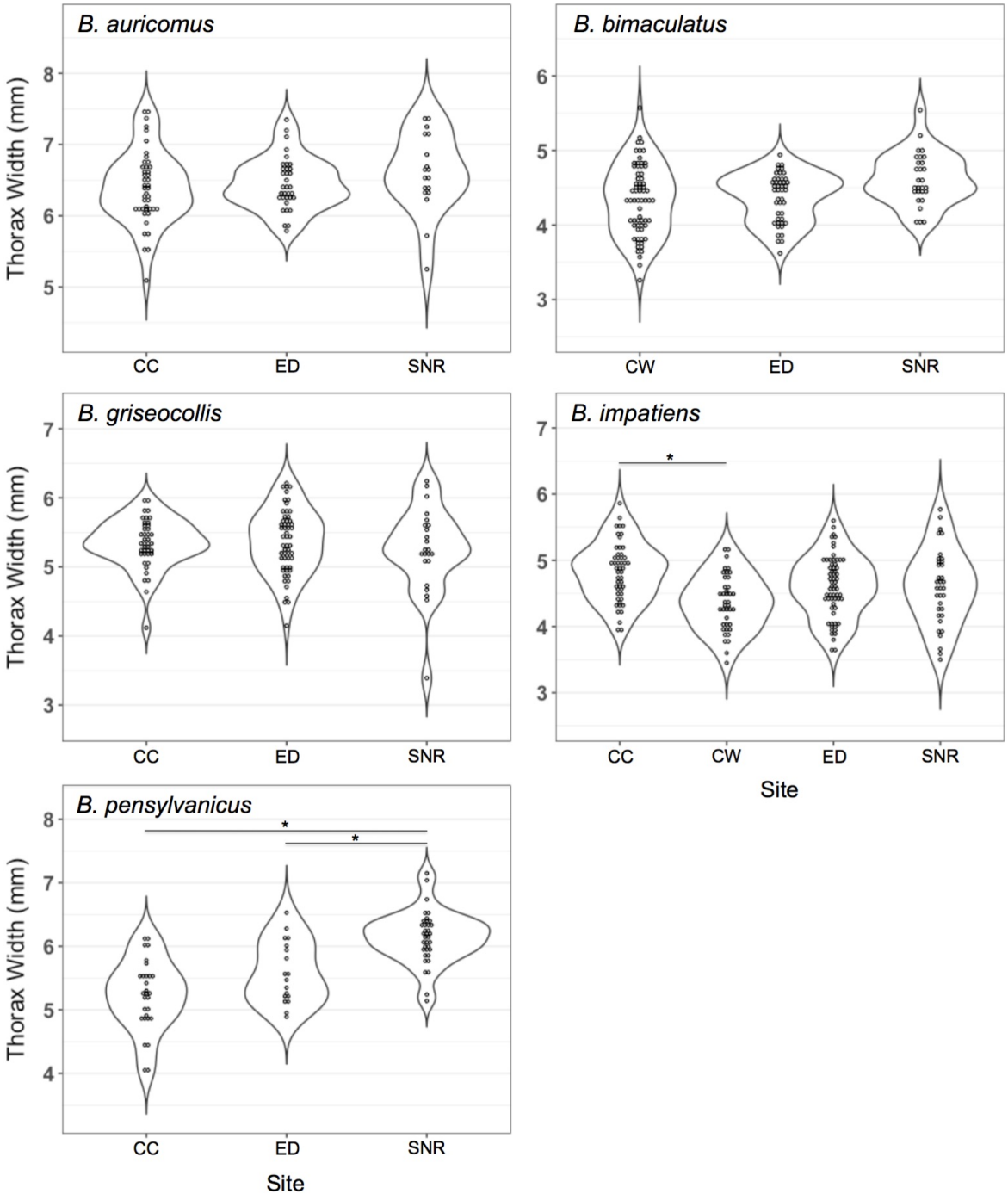

Figure 2. Thorax widths of worker bumble bees (*Bombus* spp.) in the greater Saint Louis region. Asterisks (*) indicate statistically significant differences between means of intraspecific subpopulations following Bonferroni correction (i.e., $p<0.00278$). CC =

Calvary Cemetery, CW = Castlewood State Park, ED = EarthDance Farms, SNR = Shaw Nature Reserve.

Supplementary Material

Human Population Density and Site Descriptions

To calculate human population density, we used data on cities and towns from the United States Census Bureau (United States Census Bureau 2010, 2020). We used population estimates for July 1st, 2018 (United States Census Bureau 2020) as measures of human population size per locality and land area (converted to km^2; United States Census Bureau 2010) as measures of total area per locality that a human population may occupy. We calculated human population density as the average number of people km^{-2}, by dividing population estimates by land area.

Calvary Cemetery (CC) is a Catholic cemetery located in the city of Saint Louis, Missouri (MO) (human population density = 1,889 people km^{-2}), which contains 25 acres of prairie managed by the Missouri Department of Conservation along the cemetery's northwestern edge, for which a conservation plan was implemented in 2005 (Bogan 2018). EarthDance Farms (ED) is an organic farm located in Ferguson, MO (human population density = 1,293 people km^{-2}), comprising 14 acres and a variety of native and agricultural plants, which has been a location of organic food production since 1883 (EarthDance 2019). Castlewood State Park (CW) is a state park adjacent to the Meramac River in Ballwin, MO (human population density = 1,297 people km^{-2}), comprising 1,818 acres of land and was established in 1974 (Wikipedia contributors 2019a). Shaw Nature Reserve (SNR) is a private nature reserve located on the edge of the Missouri Ozarks in Gray Summit, MO - an unincorporated community near Pacific, MO (human population density = 472 people km^{-2}) - comprising 2,500 acres of land and upwards of eight biomes

(Missouri Botanical Garden 2019), which was established in 1925 (Wikipedia contributors 2019b).

Colony Density

Measuring effective population size (N_e) can be problematic in eusocial insects, as non-reproductive worker abundances can inflate N_e, unless colony relationships are controlled for (Chapman & Bourke 2001). Therefore, we used colony density (N_c) (i.e., effective colony number) as a measure of N_e, which estimates the number of colonies at a site after controlling for colony relationships among workers (Chapman & Bourke 2001; Charman et al. 2010; Geib et al. 2015). We calculated N_c solely with female genotypes. Prior to estimating N_c, we removed loci per species that had $\geq$25% null allele frequency following Chakraborty et al. (1992), using the R package PopGenReport version 2.0 (Gruber & Adamack 2014). We estimated N_c per subpopulation by first reconstructing female sibships in Colony 2.0 (Wang 2004) using a 5% genotyping error rate and a 95% probability of females being full siblings. Following sibship reconstructions, we calculated N_c following Geib et al. (2015). To do so, we first determined the number of sampled females (N_i), the number of successfully genotyped females (N_g), and the number of colonies detected by Colony (N_{nr}). We then calculated the number of colonies detected standardized for genotyping success as $N_{ns} = (N_{nr}/N_g)*N_i$. Finally, we calculated N_c according to the Crozier model for effective population size of eusocial haplodiploid species that estimates detected colonies plus colonies not detected by sampling: $N_c = (4.5Nnm)/(1 + 2m)$; N is detected colony number, n is queen number per colony, and m is mating frequency (Crozier 1979). Accordingly, for species like bumble bees, that are

characterized by monogyny and monoandry (Goulson 2010), this calculation simplifies to $N_c = 1.5*N_{ns}$ (Charman et al. 2010). We did not calculate N_c for any subpopulation with 15 or fewer successfully genotyped females (i.e., $N_g \leq 15$).

Power Analysis

To ensure that our data had sufficient statistical power to detect true genetic differentiation, we performed a power simulation per species with the program POWSIM (version 4.1) (Ryman & Palm 2006). POWSIM tests the null hypothesis of no genetic differentiation between subpopulations, given different combinations of samples size, loci, and alleles (Ryman & Palm 2006). Each simulation estimates power via chi-square and Fisher's exact tests, while sampling from populations that diverge following a Wright-Fisher model (Ryman & Palm 2006). For all simulations, we set the expected differentiation between subpopulations to F_{ST}=0.05, which is an appropriate minimum value for true genetic structure (Frankham et al. 2002). This F_{ST} is equivalent to each subpopulation having N_e=100 after 10 generations of drift (Nei 1987). We parameterized each simulation with its respective species' observed sample size, loci number, allele number, and allele frequencies. We ran 1,000 iterations of each simulation with default parameters for dememorizations, batches, and iterations per batch. These simulations indicate the power of our sampling protocol to detect an F_{ST}=0.05 and do not represent the true evolutionary history of our study populations.

Amplification Success, Null Alleles, Linkage Disequilibrium, and Hardy-Weinberg Equilibrium

Prior to performing population genetic analyses, we removed loci from our microsatellite data following various quality control measures. Specifically, we removed loci that had ≥20% amplification failure, noisy amplification (making a locus unreliable to score), ≥25% null allele frequency following Chakraborty et al. (1992), or showed significant linkage disequilibrium (LD) with one or more loci. Analyses were performed in R Statistics. The package PopGenReport version 2.0 (Gruber & Adamack 2014) identified null alleles. The package Genepop '007 (Rousset 2008) identified loci in LD. In the following, we describe the results of these quality control measures per species. See Table S1 for loci retained per species for analyses.

Bombus auricomus

We could not reliably score B124, BTern01, and BTMS0062 in *B. auricomus* due to noisy amplification. BT28 exhibited ≥25% null allele frequency. BTern02 showed significant LD with BTMS0052 and BT30 (both $p<0.05$). Accordingly, we removed B124, BTern01, BTMS0062, BT28, and BTern02 from *B. auricomus*. Following these quality measures, 10 loci remained for the population genetic analyses of *B. auricomus.*

Bombus bimaculatus

We could not reliably score BTern02 in *B. bimaculatus* due to noisy amplification. BTMS0083 exhibited ≥25% null allele frequency. The following loci pairs showed significant LD: BTern01 and B96 ($p<0.01$), BT10 and B126 ($p<0.001$), BTMS0062 and BTMS0044 ($p<0.05$), BT28 and BTMS0059 ($p<0.05$). Accordingly, we removed BTern02, BTMS0083, BTern01, BT10, BTMS0062, and BT28 from *B. bimaculatus*.

Following these quality measures, 12 loci remained for the population genetic analyses of *B. bimaculatus.*

Bombus griseocollis

BTern02 and BT30 exhibited ≥20% amplification failure in *B. griseocollis*. We could not reliably score BL15 due to noisy amplification. BTMS0083 showed significant LD with BTMS0066, BTMS0086, and B126 (all $p<0.05$). Furthermore, the following loci pairs all showed significant LD: BTMS0066 and BTMS0062 ($p<0.05$), BTMS0086 and BT28 ($p<0.05$), and BT10 and B96 ($p<0.05$). BTMS0062 showed significant deviation from Hardy-Weinberg equilibrium across populations ($p<0.05$ in the majority of populations). Accordingly, we removed BTern02, BT30, BL15, BTMS0083, BTMS0066, BTMS0086, BT10, and BTMS0062 from *B. griseocollis*. Following these quality measures, 10 loci remained for the population genetic analyses of *B. griseocollis.*

Bombus impatiens

BTern02 exhibited ≥20% amplification failure in *B. impatiens*. We could not reliably score B126 and BTMS0062 due to noisy amplification. BTMS0066 and BTMS0059 exhibited ≥25% null allele frequency. The following loci pairs showed significant LD: B96 and BTern01 ($p<0.05$) and BT30 and BTMS0044 ($p<0.05$). Accordingly, we removed BTern02, B126, BTMS0062, BTMS0066, BTMS0059, B96, and BT30 from *B. impatiens*. Following these quality measures, 11 loci remained for the population genetic analyses of *B. impatiens.*

Bombus pensylvanicus

We could not reliably score BTMS0062, BTern02, and BTMS0044 in *B. pensylvanicus* due to noisy amplification. BTern01 showed significant LD with BL15 and BTMS0081 (both $p<0.05$). Accordingly, we removed BTMS0062, BTern02, BTMS0044, and BTern01 from *B. pensylvanicus*. Following these quality measures, 14 loci remained for the population genetic analyses of *B. pensylvanicus.*

2017 Population Genetics: Shaw Nature Reserve

Bumble Bee Sampling

In addition to the sampling performed in 2018, in the summer of 2017, we sampled worker bumble bees at Shaw Nature Reserve (SNR). From late-June through mid-August, we sampled foraging workers of *B. impatiens*, *B. griseocollis*, *B. auricomus*, and *B. pensylvanicus* by hand-netting 3-4 days per week. After capture, we immediately transferred bees to individual vials containing 100% ethanol. We did not sample *B. bimaculatus* in 2017, as the onset of our sampling corresponded with the latter half of their seasonal period of foraging activity. Sample sizes collected per species can be found in Table S7.

Microsatellite Genotyping, Colony Density, and Allelic Richness

We genotyped all 2017 *Bombus* samples at the USDA-ARS Pollinating Insect - Biology, Management, Systematics Research Unit in Logan, Utah following the same methods as described in the main text for our 2018 samples, with the following exception. For sequencing of these 2017 samples, we transferred 1.2 μL of each PCR product to a new

well of a 96 well plate, along with 9 μL of a mixture of 975 μL formamide and 25 μL 500 LIZ (internal size standard). Subsequently, an ABI PRISM 3730 DNA Analyzer at Utah State University's Center for Integrated BioSystems sequenced the samples.

After genotyping our 2017 samples, we performed quality control measures (e.g., removing loci with ≥20% amplification failure, noisy amplification, ≥25% null allele frequency, significant linkage disequilibrium) and calculated colony density and allelic richness (*AR*) following the methods described in the main text for our 2018 samples. This resulted in a minimum of seven loci being retained per species for *AR* calculations of 2017 populations. The colony density and *AR* results per species at SNR in 2017 are given in Table S7.

Supplemental Tables

Table S1. Microsatellite loci retained for each species.

Locus	Primer Sequence and Tag	Species				
		Bombus auricomus	***Bombus bimaculatus***	***Bombus griseocollis***	***Bombus impatiens***	***Bombus pensylvanicus***
B124[1]	F: *6FAM*-GCAACAGGCGGGTTAGAG R: CAGGATAGGGTAGGTAAGCAG	-	X	X	X	X
B126[1]	F: *VIC*-GCTTGCTGGTGAATTGTGC R: CGATTCTCTCGTGTACTCC	-	X	X	-	X
B96[2]	F: *PET*-GGGAGAGAAAGACCAAG R: GATCGTAATGACTCGATATG	X	X	X	-	X
BL11[3]	F: *PET*-AAGGGTACGAAATGCGCGAG R: TGACGAGTGCGGCCTTTTTC	-	-	-	X	-
BL13[3]	F: *PET*-CGAATGTTGGGATTTTCGTG R: GCGAGTACGTGTACGTGTTCTATG	X	X	X	X	X
BL15[3]	F: *6FAM*-CGAACGAAAACGAAAAAGAGC R: TCTTCTGCTCCTTTCTCCATTC	X	X	-	-	X
BT10[3]	F: *NED*-TCTTGCTATCCACCACCCGC R: GGACAGAAGCATAGACGCACCG	X	-	-	X	X
BT28[3]	F: *VIC*-TTGCTGACGTTGCTGTGACTGAGG R: TCCTCTGTGTGTTCTCTTACTTGGC	-	-	X	X	X
BT30[3]	F: *PET*-ATCGTATTATTGCCACCAACCG R: CAGCAACAGTCACAACAAACGC	X	X	-	-	X
BTern01[3]	F: *VIC*-CGTGTTTAGGGTACTGGTGGTC R: GGAGCAAGAGGGCTAGACAAAAG	-	-	X	X	-
BTern02[3]	F: *NED*-TTTCCACCCTTCACGCATACAC R: GATTTTATCCTCCGACCGTTCC	-	-	-	-	-
BTMS0044[4]	F: *PET*-AGGATCGAGAGAACGAGCTG R: AGGCCTTGGGAGAGTTCG	X	X	X	X	-
BTMS0052[4]	F: *PET*-AAATCCTTCGCTTCCGGTCT R: TGGGGGTAGCAACACTCAAA	X	X	X	X	X
BTMS0059[4]	F: *PET*-GGCTAGGAAAGATTAGCACTACC R: AGTTCGACAGACCAAGCTGT	-	X	X	-	X
BTMS0062[4]	F: *VIC*-CTGTCGCATTATTCGCGGTT R: CTGGGCGTGATTCGATGAAC	-	-	-	-	-
BTMS0066[4]	F: *6FAM*-CATGATGACACCACCCAACG R: TTAACGCCCAATGCCTTTCC	X	X	-	-	X
BTMS0081[4]	F: *PET*-ACGCGCGCCTTCTACTATC R: AGGGACACGCGAACAGAC	X	X	X	X	X
BTMS0083[4]	F: *6FAM*-CGACTCGTTCGAGCGAAATTA R: GTTTTTGCCAGGCTCCGAAT	-	-	-	X	X
BTMS0086[4]	F: *NED*-AGAGAAATTGCATGCGGTCG R: CTCGCGCTTGTCGAATCAAT	X	X	-	X	X

[1]Estoup et al. 1995; [2]Estoup et al. 1996; [3]Funk et al. 2006; [4]Stolle et al. 2009; X = locus retained, - = locus removed

Table S2. Colony density estimates for bumble bee (*Bombus* spp.) subpopulations throughout the greater Saint Louis region in 2018. N_i is the total number of sampled females, N_g is the number of successfully genotyped females, N_{nr} is the number of colonies detected from genotyping, N_{ns} is the number of colonies standardized for genotyping success, and N_c is colony density. Colony numbers were not calculated for populations with 15 or fewer successfully genotyped females. CC = Calvary Cemetery, CW = Castlewood State Park, ED = EarthDance Farms, SNR = Shaw Nature Reserve.

Species and Colony Estimates	Sites			
	CC	CW	ED	SNR
B. auricomus				
N_i	56	-	39	44
N_g	54	-	38	38
N_{nr}	54	-	36	36
N_{ns}	56.0	-	36.9	41.7
N_c	84.0	-	55.4	62.5
B. bimaculatus				
N_i	1	72	49	38
N_g	1	70	49	34
N_{nr}	-	64	46	33
N_{ns}	-	65.8	46.0	36.9
N_c	-	98.7	69.0	55.3
B. griseocollis				
N_i	45	12	61	34
N_g	45	12	56	32
N_{nr}	45	-	54	32
N_{ns}	45.0	-	58.8	34.0
N_c	67.5	-	88.2	51.0
B. impatiens				
N_i	53	42	71	41
N_g	48	42	64	39
N_{nr}	45	41	58	35
N_{ns}	49.7	41.0	64.3	36.8
N_c	74.5	61.5	96.5	55.2
B. pensylvanicus				
N_i	39	5	19	53
N_g	38	5	16	52
N_{nr}	28	-	11	44
N_{ns}	28.7	-	13.1	44.8
N_c	43.1	-	19.6	67.3

Table S3. Sample sizes and diploidy of male bumble bees (*Bombus* spp.) in the greater Saint Louis region in 2018. N_i is the total number of sampled males, N_g is the number of successfully genotyped males, N_d is the number of diploid males (i.e., number of males with ≥3 heterozygous loci). Percent diploid males is N_d/N_g. Each value is calculated per species by site and globally (i.e., combining all sites). CC = Calvary Cemetery, CW = Castlewood State Park, ED = EarthDance Farms, SNR = Shaw Nature Reserve.

Species and Statistics	Sites				Global Values
	CC	CW	ED	SNR	
B. auricomus					
N_i	0	0	0	0	0
N_g	-	-	-	-	-
N_d	-	-	-	-	-
% Diploid	-	-	-	-	-
B. bimaculatus					
N_i	0	10	8	5	23
N_g	-	10	8	5	23
N_d	-	0	0	0	0
% Diploid	-	0.00%	0.00%	0.00%	0.00%
B. griseocollis					
N_i	1	9	13	2	25
N_g	1	9	13	2	25
N_d	1	9	10	1	21
% Diploid	100.00%	100.00%	76.92%	50.00%	84.00%
B. impatiens					
N_i	0	9	5	0	14
N_g	-	8	3	-	11
N_d	-	0	0	-	0
% Diploid	-	0.00%	0.00%	-	0.00%
B. pensylvanicus					
N_i	2	0	1	0	3
N_g	2	-	1	-	3
N_d	0	-	0	-	0
% Diploid	0.00%	-	0.00%	-	0.00%

Table S4. Contrasts of intraspecific site least-squares means comparisons. These contrasts derive from a full analysis of variance (ANOVA) regressing worker body size against bumble bee species (*Bombus* spp.) and site. Statistical significance of contrasts was determined using a Bonferroni corrected α-value (i.e., $p<0.00278$) and is denoted by an asterisk (*) and italicized *p*-value.

Species and Comparison	Site Contrast *p*-values					
	CC - CW	**CC - ED**	**CC - SNR**	**CW - ED**	**CW - SNR**	**ED - SNR**
B. auricomus	-	0.6244	0.1279	-	-	0.2808
B. bimaculatus	-	-	-	0.7609	0.0067	0.0215
B. griseocollis	-	0.8305	0.442	-	-	0.3241
B. impatiens	*<.0001**	0.0084	0.0256	0.0114	0.0248	0.9722
B. pensylvanicus	-	0.0105	*<.0001**	-	-	*<.0001**

Table S5. Body size statistics for bumble bee (*Bombus* spp.) workers in the greater Saint Louis region. *N* gives the number of workers included in calculations of body size means. CC = Calvary Cemetery, CW = Castlewood State Park, ED = EarthDance Farms, SNR = Shaw Nature Reserve.

Species and Statistics	Sites			
	CC	CW	ED	SNR
B. auricomus				
N	44	-	34	18
Mean (95% CI)	6.40 (6.25-6.55)	-	6.46 (6.34-6.58)	6.60 (6.34-6.86)
B. bimaculatus				
N	-	66	45	28
Mean (95% CI)	-	4.34 (4.22-4.46)	4.37 (4.27-4.47)	4.62 (4.49-4.75)
B. griseocollis				
N	41	-	56	23
Mean (95% CI)	5.34 (5.23-5.45)	-	5.36 (5.24-5.48)	5.25 (5.00-5.50)
B. impatiens				
N	50	40	63	35
Mean (95% CI)	4.83 (4.71-4.95)	4.36 (4.23-4.49)	4.60 (4.48-4.72)	4.60 (4.41-4.79)
B. pensylvanicus				
N	29	-	18	36
Mean (95% CI)	5.23 (5.03-5.43)	-	5.59 (5.36-5.82)	6.14 (6.00-6.28)

Table S6. Floral genera visited by bumble bee (*Bombus* spp.) females in the greater Saint Louis region in 2018. The percent of bees visiting each floral genus per species and site are given in parentheses. *n* = number of female bees collected visiting flowers, CC = Calvary Cemetery, CW = Castlewood State Park, ED = EarthDance Farms, SNR = Shaw Nature Reserve.

Species	Floral Genera (Visitation Percent) per Site			
	CC	**CW**	**ED**	**SNR**
B. auricomus	*n*=49 *Calystegia* (14.3%) *Carduus* (6.1%) *Dipsacus* (8.2%) *Penstemon* (8.2%) *Rumex* (2.0%) *Trifolium* (18.4%) *Vicia* (42.9%)	*n*=0	*n*=39 *Agastache* (5.1%) *Monarda* (2.6%) *Trifolium* (53.8%) *Vicia* (38.5%)	*n*=43 *Baptisia* (11.6%) *Dasistoma* (4.7%) *Iris* (2.3%) *Monarda* (58.1%) *Penstemon* (20.9%) *Pycnanthemum* (2.3%)
B. bimaculatus	*n*=1 *Ipomoea* (100.0%)	*n*=70 *Blephilia* (21.4%) *Glechoma* (4.3%) *Hydrophyllum* (4.3%) *Teucrium* (11.4%) *Trifolium* (58.6%)	*n*=46 *Agastache* (4.3%) *Borago* (2.2%) *Lavandula* (2.2%) *Salvia* (2.2%) *Symphytum* (28.3%) *Trifolium* (17.4%) *Vicia* (43.5%)	*n*=34 *Amorpha* (5.9%) *Asclepias* (2.9%) *Baptisia* (2.9%) *Monarda* (11.8%) *Pedicularis* (5.9%) *Penstemon* (55.9%) *Pycnanthemum* (11.8%) *Trifolium* (2.9%)
B. griseocollis	*n*=42 *Apocynum* (4.8%) *Calystegia* (28.6%) *Carduus* (4.8%) *Dipsacus* (4.8%) *Monarda* (2.4%) *Securigera* (35.7%) *Trifolium* (4.8%) *Vernonia* (2.4%) *Vicia* (11.9%)	*n*=12 *Blephilia* (8.3%) *Teucrium* (16.7%) *Trifolium* (75.0%)	*n*=57 *Agastache* (3.5%) *Asclepias* (21.1%) *Calystegia* (1.8%) *Echinacea* (14.0%) *Monarda* (3.5%) *Teucrium* (1.8%) *Trifolium* (28.1%) *Vicia* (26.3%)	*n*=32 *Amorpha* (6.3%) *Asclepias* (12.5%) *Baptisia* (6.3%) *Echinacea* (3.1%) *Iris* (6.3%) *Monarda* (12.5%) *Penstemon* (15.6%) *Pycnanthemum* (18.8%) *Senecio* (3.1%) *Veronicastrum* (15.6%)
B. impatiens	*n*=51 *Calystegia* (23.5%) *Cirsium* (31.4%) *Dipsacus* (13.7%) *Helianthus* (29.4%) *Securigera* (2.0%)	*n*=40 *Teucrium* (35.0%) *Trifolium* (7.5%) *Verbesina* (57.5%)	*n*=54 *Agastache* (48.1%) *Allium* (5.6%) *Cichorium* (1.9%) *Convolvulus* (1.9%) *Ipomoea* (1.9%) *Symphyotrichum* (22.2%) *Symphytum* (1.9%) *Teucrium* (1.9%) *Trifolium* (14.8%)	*n*=40 *Agastache* (2.5%) *Amorpha* (2.5%) *Baptisia* (2.5%) *Chamaecrista* (7.5%) *Dasistoma* (7.5%) *Lactuca* (5.0%) *Penstemon* (12.5%) *Silphium* (2.5%) *Solidago* (17.5%) *Verbesina* (17.5%) *Veronicastrum* (22.5%)
B. pensylvanicus	*n*=38 *Calystegia* (2.6%) *Carduus* (2.6%) *Cirsium* (2.6%) *Dipsacus* (68.4%) *Trifolium* (18.4%) *Vicia* (5.3%)	*n*=5 *Solanum* (20.0%) *Teucrium* (20.0%) *Trifolium* (40.0%) *Verbesina* (20.0%)	*n*=17 *Trifolium* (82.4%) *Vicia* (17.6%)	*n*=53 *Agastache* (3.8%) *Baptisia* (5.7%) *Chamaecrista* (5.7%) *Coreopsis* (3.8%) *Dasistoma* (26.4%) *Iris* (3.8%) *Lespedeza* (3.8%) *Monarda* (5.7%) *Penstemon* (5.7%) *Scutellaria* (3.8%) *Silphium* (26.4%) *Vernonia* (1.9%) *Veronicastrum* (3.8%)

Table S7. Sample sizes, colony estimates, and allelic richness (*AR*) per bumble bee species (*Bombus* spp.) at Shaw Nature Reserve (SNR) in the summer of 2017. N_i is the total number of sampled females, N_g is the number of successfully genotyped females, N_{nr} is the number of colonies detected from genotyping, N_{ns} is the number of colonies standardized for genotyping success, and N_c is colony density. SE = standard error.

Variable		Species			
		B. auricomus	***B. griseocollis***	***B. impatiens***	***B. pensylvanicus***
Colony Estimates	N_i	30	37	47	48
	N_g	29	31	37	46
	N_{nr}	29	31	37	46
	N_{ns}	30.0	37.0	47.0	48.0
	N_c	45.0	55.5	70.5	72.0
***AR* (SE)**		6.40 (1.38)	7.55 (2.03)	7.62 (2.43)	5.56 (2.18)

Supplemental Figure

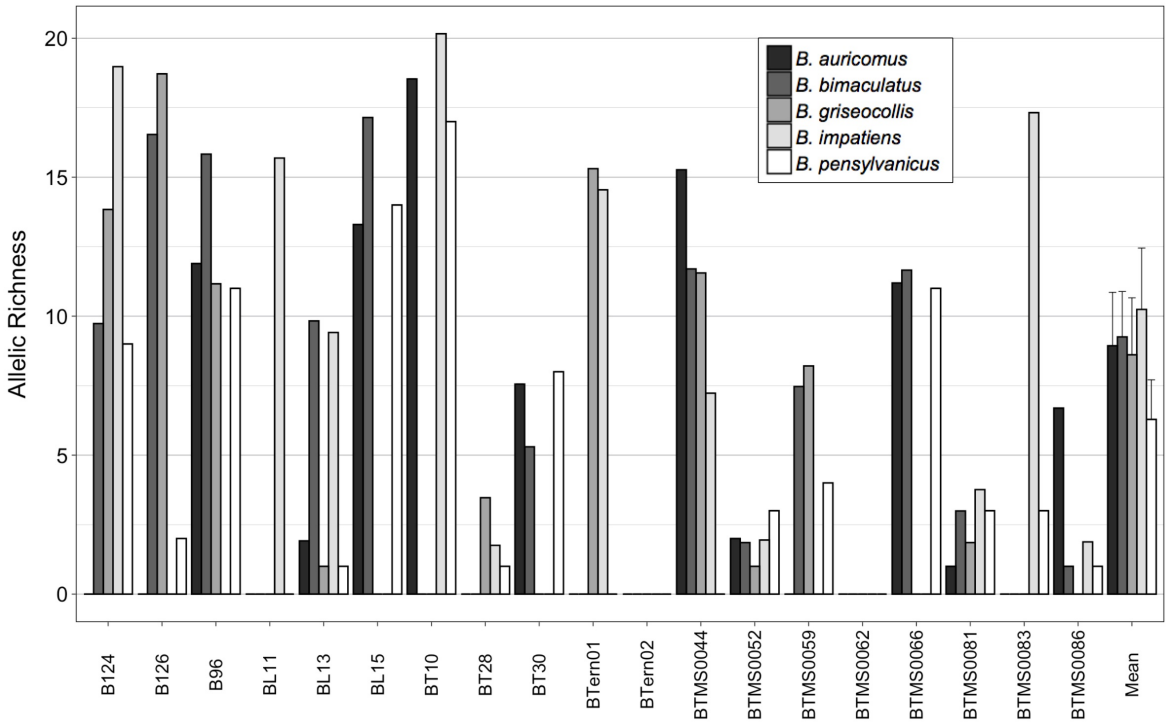

Figure S1. Global allelic richness at each locus after sample size rarefaction (*n*=176 alleles/species). Loci with zero values were either unamplified or dropped from analyses. Means (+SE) are computed for intraspecific loci with nonzero values.

References for Supplemental Materials

Bogan, J. 2018. Prairie restoration project is in bloom in north St. Louis. St. Louis Post-Dispatch. Accessed 18 Dec. 2019. https://graphics.stltoday.com/apps/prairie/

Chakraborty, R., Andrade, M.D.E., Daiger, S.P., & Budowle, B. 1992. Apparent heterozygote deficiencies observed in DNA typing data and their implications in forensic applications. Annals of Human Genetics 56: 45–57.

Chapman, R.E. & Bourke, A.F.G. 2001. The influence of sociality on the conservation biology of social insects. Ecology Letters 4: 650-662.

Charman, T.G., Sears, J., Green, R.E., & Bourke, A.F.G. 2010. Conservation genetics, foraging distance and nest density of the scarce Great Yellow Bumblebee (*Bombus distinguendus*). Molecular Ecology 19: 2661–2674.

Crozier, R.H. 1979. Genetics of sociality. In: Social Insects, Vol. I (ed. Hermann HR), pp. 223–286. Academic Press, New York.

EarthDance. 2019. About. Accessed 18 Dec. 2019. https://earthdancefarms.org/about/

Estoup, A., Scholl, A., Pouvreau, A., & Solignac, M. 1995. Monoandry and polyandry in bumble bees (Hymenoptera; Bombinae) as evidenced by highly variable microsatellites. Molecular Ecology 4: 89–94.

Estoup, A., Solignac, M., Cornuet, J.M., Goudet, J., & Scholl, A. 1996. Genetic differentiation of continental and island populations of *Bombus terrestris* (Hymenoptera: Apidae) in Europe. Molecular Ecology 5: 19–31.

Frankham, R., Ballou, J.D., & Briscoe, D.A. 2002. Introduction to Conservation Genetics. Cambridge University Press, Cambridge.

Funk, C.R., Schmid-Hempel, R., & Schmid-Hempel, P. 2006. Microsatellite loci for *Bombus* spp. Molecular Ecology Notes 6: 83–86.

Geib, J.C., Strange, J.P., & Galen, A. 2015. Bumble bee nest abundance, foraging distance, and host-plant reproduction: Implications for management and conservation. Ecological Applications 25: 768–778.

Goulson, D. 2010. Bumblebees: Behaviour, Ecology, and Conservation. Oxford University Press.

Gruber, B. & Adamack, A.T. 2014. PopGenReport: simplifying basic population genetic analyses in R. Methods in Ecology and Evolution 5:384-387.

Missouri Botanical Garden. 2019. Plant and Animal Habitats. Accessed 18 Dec. 2019. http://www.missouribotanicalgarden.org/visit/family-of-attractions/shaw-nature-reserve/conservation-at-shaw-nature-reserve/plants-animals-habitats.aspx

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

Rousset, F. 2008. genepop'007: a complete re-implementation of the genepop software for Windows and Linux. Molecular Ecology Resources 8: 103-106.

Ryman, N. & Palm, S. 2006. POWSIM: a computer program for assessing statistical power when testing for genetic differentiation. Molecular Ecology Notes 6:600-602.

Stolle, E., Rohde, M., Vautrin, D., Solignac, M., Schmid-Hempel, P., Schmid-Hempel, R., & Moritz, R.F.A. 2009. Novel microsatellite DNA loci for *Bombus terrestris* (Linnaeus, 1758). Molecular Ecology Resources 9: 1345–1352.

United States Census Bureau. 2010. Land Area and Persons Per Square Mile. Accessed 26 March 2020. https://www.census.gov/quickfacts/fact/note/US/LND110210

United States Census Bureau. 2020. City and Town Population Totals: 2010-2018. Accessed 26 March 2020. https://www.census.gov/data/tables/time-series/demo/popest/2010s-total-cities-and-towns.html

Wang, J. 2004. Sibship reconstruction from genetic data with typing errors. Genetics 166: 1963-1979.

Wikipedia contributors. 2019a. Castlewood State Park. Wikipedia, The Free Encyclopedia. Accessed 18 Dec. 2019. https://en.wikipedia.org/w/index.php?title=Castlewood_State_Park&oldid=905158919

Wikipedia contributors. 2019b. Shaw Nature Reserve. Wikipedia, The Free Encyclopedia. Accessed 18 Dec. 2019. https://en.wikipedia.org/w/index.php?title=Shaw_Nature_Reserve&oldid=914313203

Chapter III: Phenology of cognition: seasonal trends in average population-level learning

Matthew W. Austin[a,b,*] and Aimee S. Dunlap[a,b]

[a]Department of Biology, University of Missouri – St. Louis, St. Louis, MO, USA

[b]Whitney R. Harris World Ecology Center, St. Louis, MO, USA

[*] Corresponding author

AustinMattW@gmail.com

Abstract

Patterns of animal behavior can mirror spatiotemporal environmental variation, such as when behavioral events synchronize with resource phenology. Less known is whether cognitive abilities *per se* can also mirror patterns of environmental variation. Here, we test the hypothesis that changes to population-level cognition can occur phenologically, in response to individuals produced at different time points being provisioned with resources of different nutritional quality. We test this hypothesis in bumble bees (Apidae: *Bombus*), a clade of annual pollinating eusocial insects that produce individuals at different time points across their reproductive season and exhibit organ developmental plasticity in response to variance in nutritional quality. To accomplish this we (1) take direct measurements of learning ability across a reproductive season of five bumble bee species and (2) develop a simulation model that depicts how known dynamics of bumble bee life history and foraging ecology, coupled with developmental plasticity of cognition, may affect average colony-level cognition across a season. We find that two of our focal species - *Bombus auricomus* and *Bombus pensylvanicus* - exhibit seasonal trends in cognition, with the proportion of workers successfully completing a learning test increasing as the season progresses. Additionally, our simulation model finds that bumble bees can increase average colony-level learning across a season, due to increased provisioning of larvae across colony development. The exception to this occurs in environments with high resource quality early in colony development, where high average colony-level learning across the season is promoted. Collectively these results support our hypothesis and suggest that population-level phenological changes to cognition is a biologically plausible phenomenon.

Keywords: *cognition, ecological modeling, environmental variation, nutritional ecology, phenology, seasonality, temporal fragmentation*

Introduction

Environmental change is ubiquitous across ecosystems. Spatiotemporal variation of abiotic factors (e.g. temperature, photoperiod, H_2O) drives seasonality and cues phenology across taxa (Chmura et al 2019; Visser & Both 2005). As the expression of behavior depends on environmental factors (Shettleworth 2010), variation in animal behavior can mirror spatiotemporal environmental variation. For example, behavioral events (e.g. foraging, migration, emergence) synchronize with resource phenology (McGrath et al. 2009; García-Navas & José Sanz; Minckley et al. 1994); asynchronous timing between behavioral events and resource phenology can have large fitness costs (e.g. Shoji et al. 2015; van Asch & Visser 2007). Furthermore, in addition to behavior, cognition itself is partially mediated by the environment (e.g. Shettleworth 2010). In certain taxa, adult cognition is determined by developmental plasticity during juvenile ontogeny, whereby certain environmental conditions promote enhanced cognitive abilities (e.g. Lanet & Maurange 2014; Cheng et al. 2011). In systems where environmental conditions vary across time and individuals of a given generation are produced at different time points throughout a reproductive season (e.g. Szigeti et al. 2019), such plasticity of cognition may result in changes to average cognitive ability, at the population-level, across a season. However, while changes in cognitive ability have been explored across individual lifetimes (e.g. Shettleworth 2010) and across evolutionary time (e.g. Dunlap & Stephens 2009; Mery & Kawecki 2002, 2004; Stephens 1991; Dunlap et al. 2009), whether populations living in seasonally variable environments exhibit "*phenology of cognition*" is a contemporarily unexplored topic.

Phenological changes to average cognitive ability, at the population-level, may be expected in systems that exhibit seasonality in nutritional availability, as nutrition received throughout ontogeny is a primary contributor to adult cognition in many species (e.g. Lanet & Maurange 2014; Cheng et al. 2011). For example, consumption of higher quantities of food (Steijven et al 2017) and specific pollen fatty acids (Arien et al. 2018; Muth et al. 2018) promote enhanced associative learning abilities in bees and taurine supplements promote greater spatial learning abilities in birds (Arnold et al. 2007; see also Brust et al. 2014). Additionally, differential feeding regimes during ontogeny can affect neurogenesis (Moda et al. 2013), with nutrient restriction leading to reduced size of certain brain sections (Barbeito-Andrés et al. 2019) and lower brain volume overall (Steijven et al. 2017). Such reduced brain growth under nutrient restriction may result from resources being preferentially allocated to other vital organs (Barbeito-Andrés et al. 2019). While in certain species, central nervous system (CNS) development is spared under nutrient restriction relative to other organs, adult CNS volume overall is lower when food is limited during juvenile ontogeny, compared to when food is provided *ad libitum* (Lanet & Maurange 2014; Cheng et al. 2011). As neuroanatomy (e.g. Julian and Gronenberg 2002; Farris and Roberts 2005; Lefebvre et al. 1997) and relative brain size (e.g. Sol et al. 2005; Collado et al. 2020) are linked to cognitive complexity, these effects of nutrition on brain development likely have implications for adult learning abilities. In rapidly changing environments, an organism's ability to plastically change associations between stimuli, as mediated by their cognitive complexity (Mikhalevich et al. 2017), can increase relative fitness (Fryxell et al. 2005, Snell-Rood 2013; Wong & Candolin 2015;

Tuomainen & Candolin 2011). Therefore, phenological changes in cognitive abilities may have cascading effects on individual fitness and long-term population success.

Pollination systems are an ideal model for the study of cognitive phenology. Due to ephemerality and temporal partitioning of floral resources, pollinators must contend with an environment that rapidly changes in resource composition and abundance across a season (e.g. Szigeti et al. 2019; Ogilvie & Forrest 2017), including intermittent periods of food dearth (Timberlake et al. 2019a). In temperate climates, the reproductive season of pollinating insects is often synchronized with flowering (e.g. Minckley et al. 1994; Bartomeus et al. 2011) and many pollinating insects produce individuals at different time points throughout their reproductive season (e.g. Szigeti et al. 2019). Collectively, this floral turnover and succession of developmental periods results in individuals that develop at different time points being provisioned with resources from different floral species. Consequently, populations of pollinating insects may change phenotypic composition across time, as the larval stage is a critical period of insect development, with resources consumed during larval development having lasting effects on adult phenotype (e.g. Koyama et al. 2013), including nervous system functionality (e.g. Lanet & Maurange 2014). Here, we use bumble bees (Apidae: *Bombus*) - a clade of eusocial insects producing annual colonies across the Northern Hemisphere and South America (Goulson 2010) - as a model system for exploring the concept of a cognitive phenology. Bumble bees are an ideal system for this work; bumble bees are primary pollinators in many temperate ecosystems, have well-described demographic histories (e.g. Pereboom et al. 2003; O'Donnell et al. 2000), where workers are successively produced at different time points across a reproductive season (e.g. Goulson 2010), and exhibit organ

developmental plasticity, with greater organ development resulting from consumption of greater nutritional value during larval development (e.g. Couvillon & Dornhaus 2009)

In this study, we explore the concept of a cognitive phenology by (1) taking direct measurements of learning ability across a reproductive season and (2) presenting a simulation model that depicts how developmental plasticity of cognition, coupled with bumble bee colony demography and foraging dynamics, can produce changes in cognitive abilities, at the colony-level, across a season. Our simulation model is parameterized with observed data on bumble bee life history (e.g. Cnaani et al. 2002; O'Donnell et al. 2000; Goldblatt & Fell 1986; da Silva-Matos & Garófalo 2000) and is run for colonies in different resource environments (e.g. Timberlake et al. 2019a,b). We hypothesize that changes to larval nutritional consumption across colony development can result in different levels of cognitive ability among individuals produced at different time points. Accordingly, we predict field populations and simulated colonies will exhibit seasonal trends in cognition. The causes and consequences of animal cognition is a subject that has received considerable empirical and theoretical attention (e.g. Shettleworth 2010; Dunlap & Stephens 2009; Mery & Kawecki 2002, 2004; Stephens 1991; Dunlap et al. 2009). This study builds upon this work by investigating how seasonal environmental change may promote phenological trends in population-level cognition.

Methods

Study System and Sampling

We sampled bumble bee workers weekly from each of four sites across the greater St. Louis, Missouri area in 2018. To ensure that our sampling period occurred throughout the entire period of foraging worker activity, we began sampling in early-May, before we observed worker bumble bees at our study sites, and concluded sampling in late-September, after we no longer observed workers at these sites. Five bumble bee species can be reliably found throughout the St. Louis area, all of which we included in this study: *Bombus auricomus*, *Bombus bimaculatus*, *Bombus griseocollis*, *Bombus impatiens*, and *Bombus pensylvanicus* (Camilo et al. 2018). These species have partitioned phenologies, with *B. bimaculatus* emerging first in the spring and *B. impatiens* and *B. pensylvanicus* emerging the latest in mid-summer. Furthermore, these species constitute a mix of stable and declining species, with *B. bimaculatus*, *B. griseocollis*, and *B. impatiens* having increased relative abundance in North America over the past century, while *B. auricomus* and *B. pensylvanicus* have decreased relative abundance (Hatfield et al. 2015). We sampled by hand netting free-foraging workers and immediately transferred bees to individual test vials (Fig S1), where they were kept to acclimate prior to the learning test (see below). Following the learning test, we took a mid-leg tarsal clipping per individual (used in a separate genetic study, see Austin et al. *in-prep*) and released bees in their area of capture. We ensured no bees were tested more than once by not testing any bees that were captured with a missing tarsus.

Field Learning Tests

To assess learning ability, we utilized a differential conditioning procedure with a technique called the Free-Moving Proboscis Extension Response (FMPER) (Muth et al.

2017). FMPER involves presenting a bee in a test vial with strips of paper that are inserted ~1 cm into either of two holes located on one end of the vial (Fig S1). Prior to presentation, we soaked the end of each paper strip in a solution of either 50% sucrose (weight/weight), 5% NaCl, or deionized (DI) H_2O. Respectively, these solutions are unconditioned stimuli (US) that are either positively reinforcing (US+), negatively reinforcing (US-), or unrewarding. When a paper strip is presented, the bee extends her proboscis to the strip and drinks from it for 3 sec before the strip is removed.

Our differential conditioning procedure first involved testing bees for their initial preference between a blue and a yellow strip of paper, by pairing both paper strips with 50% sucrose and simultaneously inserting them into the vial. We recorded the bee's color preference as the paper strip the bee first extended her proboscis to. After the bee drank from this preferred paper strip for 3 sec, both paper strips were removed before the bee could drink from both paper strips. Subsequently, we performed five trials, with each trial pairing the color strip that was initially preferred with 5% NaCl and the color strip that was initially not preferred with 50% sucrose. In each trial, these strips were presented one after another and the bee was allowed to drink from each for 3 sec. In each trial, these strips were presented in the same hole and in between trials we alternated the hole that was used and the order the colors were presented in. Finally, we performed a test phase, in which a blue and yellow strip are both paired with DI H_2O and simultaneously inserted into the vial. As pairing of stimuli in the five trials matched the initially preferred color with an aversive stimulus (i.e. 5% NaCl), and the initially non-preferred color with a positive stimulus (i.e. 50% sucrose), if the bee chose the initially non-preferred color in the unrewarded test phase, we recorded the bee as successfully completing the learning

test. In other words, choosing the initially non-preferred color in the unrewarding test phase is evidence that the bee was trained against her preference, learning the initially non-preferred color as a positively conditioned stimulus (CS+) and the initially preferred color as a negatively conditioned stimulus (CS-). Accordingly, this learning test results in binary data; 1 = success in the learning test, 0 = failure in the learning test. We included a 4 min interval between each of the five trials and the unrewarded test phase. Across bees, we randomized the hole in which each color was presented for both the initial preference phase and the unrewarded test phase.

Statistical Analysis of Field Data

To assess whether each species exhibits a temporal trend in worker cognition, we ran a logistic regression per species. For these regressions, we assigned each date a number ranging from 1 for the first date of testing (May 24th) to 113 for the last date of testing (September 12th). In each of these regressions, we used date as the predictor variable and success in the learning test as the response variable.

As our sampling protocol resulted in each species having an uneven temporal distribution of data points, to ensure that the results of our logistic regressions were not an artifact of this uneven sampling distribution, we also performed a randomization test for each species. Specifically, for each species we performed 1,000 logistic regression simulations that contained the observed sample size and probability of success in the learning test, but with each data point randomly assigned to a date from throughout that species' range of testing dates. For each species, we then compared the *z*-value from the logistic regression of observed data to the *z*-value distribution constructed from these

1,000 simulations. If a species' observed *z*-value fell within the top 2.5% or bottom 2.5% of their *z*-value distribution (assuming a two-tailed distribution), we denoted the results of that species' observed logistic regression as not being an artifact of an uneven temporal data distribution.

Simulation Model

To determine whether a phenological trend in cognition is a biologically plausible phenomenon, we developed a simulation model that simulates the average learning ability of workers within a colony across repeated time steps. This model (1) parameterizes colony growth and foraging from observed data on bumble bee colonies (e.g. Cnaani et al. 2002; O'Donnell et al. 2000; Goldblatt & Fell 1986; da Silva-Matos & Garófalo 2000) and (2) assumes a positive relationship between the value of resources consumed during larval development and adult cognition. To evaluate the effect of colony life history and floral environment on seasonal changes to average colony-level learning, we simulate our model for colonies of different size in various resource environments.

Bumble Bee Demography

Bumble bees are an annual, eusocial species of Hymenoptera, that predominantly occur in temperate and subpolar environments (Goulson 2010). In late winter or spring, foundress queens emerge from diapause and initiate a colony. While the duration of bumble bee colonies varies interspecifically, they typically last for several months after founding (Goulson 2010). To incorporate seasonality into our model, we simulate our

model across 200 repeated time steps, t, where each time step is analogous to one day (i.e. $t = 1, 2, 3, \ldots, 200$).

Like many annual eusocial Hymenoptera, bumble bee colonies follow a bang-bang strategy of colony growth, whereby the colony is divided into two phases: (1) In the first phase, all reproductive effort is allocated to worker production and no reproductives are produced. (2) After a switching point (i.e. critical time), t_s, the second phase commences, in which workers cease to be produced and all reproductive effort is allocated to the production of reproductives (Macevicz & Oster 1976). We parameterize our model following this bang-bang strategy, where the number of worker eggs laid before the switching point (i.e. $t \leq t_s$) is given as

$$B_t = \theta \quad (1)$$

Here, B_t - the number of worker eggs laid by the queen at the current time step - is given as a constant rate of colony growth, θ. While the production of reproductives is not explicitly built into this model, as the bang-bang strategy results in no worker eggs being laid after the reproductive phase of the colony commences, after the switching point (i.e. $t > t_s$), the number of workers eggs laid is given as

$$B_t = 0 \quad (2)$$

From oviposition to eclosion, bumble bee developmental periods are subdivided into egg, larva, and pupa stages (Cnaani et al. 2002). We parameterize developmental periods in

our model following data on *Bombus impatiens* from Cnaani et al. (2002), whereby each worker's developmental period (i.e. from oviposition to eclosion) lasts for 24 time steps: the egg stage occurs for 5 days (time steps 1-5 post-oviposition), the larva stage occurs for 9 days (time steps 6-14 post-oviposition), and the pupa stage occurs for 10 days (time steps 15-24 post-oviposition). For each worker, j, the time step of oviposition is denoted as t_B and the time step of death is denoted as t_D. To parameterize worker death, at the end of each time step a given percent of adult workers die, following observed bumble bee life table mortality schedules (see *Colony Size and Phenology* below). Prior to death, each worker's t_D = NA; upon death, a worker's t_D is updated to the current time step (i.e. $t_D = t$).

In each time step, all individuals in the colony are stored in a matrix, $\mathbf{A}_t$, of the corresponding variables A_{jk}. In this matrix, rows, j, are separate individuals and columns, k, are parameters. J is the total number of individuals in $\mathbf{A}_t$. Specifically, $k = 1$ is D_t, $k = 2$ is L_t (see below for description of D_t and L_t), $k = 3$ is t_B, and $k = 4$ is t_D. $\mathbf{A}_t$ has the following structure

$$\mathbf{A}_t = \begin{bmatrix} A_{11} & A_{12} & A_{13} & A_{14} \\ A_{21} & A_{22} & A_{23} & A_{24} \\ \vdots & \vdots & \vdots & \vdots \\ A_{J1} & A_{J2} & A_{J3} & A_{J4} \end{bmatrix} \quad (3)$$

Foraging and Resources

Unlike other Hymenopteran groups, bumble bees do not exhibit strict task specialization; bumble bee workers often switch between various tasks throughout their lifetime (Goulson 2010). However, data from colonies established by field caught bumble bee

queens suggest that the majority of workers in a colony - between 88-94% - will be designated as foragers (O'Donnell et al. 2000). We parameterize our model with these data, whereby, in each time step, between 88-94% of living adult workers are randomly designated as foragers. In other words, in each time step, let Z_t be a random variable from the continuous uniform distribution of {0.88, 0.94}. Then, create a new matrix of foraging workers, $\mathbf{F}_t$, by subsetting $\mathbf{A}_t$ with Z_t proportion of randomly selected workers (j) who are adults ($t - t_B \geq 25$) and are alive (t_D = NA). Furthermore, prior to the switching point (i.e. $t \leq t_s$), the queen acts as a forager and is added to $\mathbf{F}_t$. $\mathbf{F}_t$, with the corresponding variables F_{hi}, has the following structure

$$\mathbf{F}_t = \begin{bmatrix} F_{11} & F_{12} & F_{13} & F_{14} \\ F_{21} & F_{22} & F_{23} & F_{24} \\ \vdots & \vdots & \vdots & \vdots \\ F_{H1} & F_{H2} & F_{H3} & F_{H4} \end{bmatrix} \quad (4)$$

As $\mathbf{F}_t$ is a subset of $\mathbf{A}_t$, i corresponds to similar parameters as k between each matrix, where $i = 1$ is d_t, $i = 2$ is l_t (see below for a description of d_t and l_t), $i = 3$ is t_B, and $i = 4$ is t_D; however, in $\mathbf{F}_t$, h, denotes individual foragers. H is the total number of foragers in $\mathbf{F}_t$.

To assess the effect of floral resource phenology on seasonal changes to colony-level cognition, we simulate our model in multiple resource environments. These resource environments are divided into the following: (1) an *observed resource environment*, based on a real community-level nectar phenology dataset (Timberlake et al. 2019a,b), (2) *stable resource environments*, consisting of a single resource value across the entire season, and (3) *pulsed resource environments*, in which periods of low and high resource values vacillate across the season (Fig 1). Our stable resource

environments are further subdivided into a *stable low resource environment*, providing a single low resource value across the season, and a *stable high resource environment*, providing a single high resource value across the season. Similarly, our pulsed resource environments are subdivided into a pulsed environment with the low pulse occurring first (i.e. *pulsed low 1st resource environment*) and a pulsed environment with the high pulse occurring first (i.e. *pulsed high 1st resource environment*). In our pulsed environments, the resource values provided in the low and high pulses are equivalent to the values provided in the stable low and stable high resource environments respectively, with each pulse lasting for 25 time steps.

The observed resource environment is based on a high-resolution dataset on nectar phenology from Timberlake et al. (2019a,b). To compile this dataset, in 2017, Timberlake et al. (2019a) quantified flowering phenology of every floral species at three farms in Somerset, England from late-February through mid-October using a transect sampling approach. In addition to recording the date that each flowering plant was observed on, Timberlake et al. (2019a) (1) estimated flowering density (i.e. mean number of flowers meter^{-2}) of each species and (2) used previously published data on nectar content of English flora from Baude et al. (2016) to estimate sugar content per flower (i.e. mean sugar flower^{-1} day^{-1}) of each encountered species. This resulted in a dataset of every floral species encountered per transect on each sampling date, with coinciding data on the mean number of flowers per square-meter and mean sugar content flower^{-1} day^{-1} (Timberlake et al. 2019b).

We use this high-resolution dataset on nectar phenology from Timberlake et al. (2019a,b) to parameterize the observed resource environment in our model. Specifically,

for every date that Timberlake et al. (2019a,b) sampled, we calculated community-wide values of the mean, standard deviation, minimum, and maximum sugar content per flower (mean sugar flower^{-1} day^{-1}) that account for flowering density (mean number of flowers meter^{-2}) per species. Subsequently, we approximated these community-wide sugar content values across time steps by (1) treating the first date of sampling by Timberlake et al. (2019a,b) (February 28th, 2017) as $t = 1$, (2) assigning every subsequent sampling date to its corresponding time step, and (3) approximating sugar content values for each unsampled time step by imputing linearly fit values of the mean, standard deviation, minimum, and maximum between every two consecutively sampled time steps. We then assembled a matrix of resource values in the environment, **S**, of the corresponding variables S_{tg}. In this matrix, rows, t, are time steps and columns, g, correspond to the following: $g = 1$ is mean community-wide sugar content, $g = 2$ is the standard deviation of community-wide sugar content, $g = 3$ is the minimum community-wide sugar content value, and $g = 4$ is the maximum community-wide sugar content value. **S** has the following structure

$$\mathbf{S} = \begin{bmatrix} S_{1,1} & S_{1,2} & S_{1,3} & S_{1,4} \\ S_{2,1} & S_{2,2} & S_{2,3} & S_{2,4} \\ \vdots & \vdots & \vdots & \vdots \\ S_{200,1} & S_{200,2} & S_{200,3} & S_{200,4} \end{bmatrix} \quad (5)$$

In every time step, the mean, standard deviation, minimum, and maximum community-wide sugar content values are used to create a normal distribution of resource values in the environment. Each forager encounters a resource value randomly drawn from this normal distribution, which is stored as d_t in $\mathbf{F}_t$. In other words, in each time step, let

$N_t(S_{t1}, S_{t2}, S_{t3}, S_{t4})$ be a normal distribution with mean S_{t1}, standard deviation S_{t2}, minimum S_{t3}, and maximum S_{t4}. Then, assign every F_{h1} a random value drawn from N_t.

We determine the resource values of our stable and pulsed environments from the community-wide floral sugar content values used to parameterize our observed resource environment. Specifically, we took the maximum community-wide sugar content values per time step, and used the lowest of these values as the resource value in the stable low environment and the highest of these values as the resource value in the stable high environment. To ensure that each forager in each time step of a stable environment encounters only the exact resource value provided by that environment, in our stable environments, we set this resource value in **S** as equivalent to the mean (S_{t1}), minimum (S_{t3}), and maximum (S_{t4}) and arbitrarily set the standard deviation to $S_{t2} = 1$. This results in each forager's resource value in each time step (d_t) of a stable environment equaling the exact resource value provided by that environment. For our pulsed environments, we alternate the stable low and stable high environments for 25 time steps each across the 200 time steps the model is run for.

The resources returned to the colony in each time step, R_t, is a function of d_t, the resource value encountered by each forager and each forager's learning score, l_t. Accordingly, the value of resources returned to the colony by foragers in each time step is given as

$$R_t = \sum_{h=1}^{N_{Ft}} d_{th}\, l_{th} \quad (6)$$

where N_{Ft} is the total number of foragers (h) at the current time step.

As resources are returned to the colony, they are evenly divided among all developing larvae (i.e. $t - t_B > 5$ & $t - t_B < 15$). Before the switching point (i.e. $t \leq t_s$), the resource value fed to each developing worker larva per time step is given as

$$P_t = \frac{R_t}{N_{Lt}} \quad (7)$$

Here, N_{Lt} denotes the total number of developing worker larvae at the current time step (i.e. $t - t_B > 5$ & $t - t_B < 15$). Note that N_{Lt} does not include workers developing in the egg (i.e. $t - t_B \leq 5$) and pupal stages (i.e. $t - t_B \geq 15$ to $t - t_B < 25$), as individuals in the egg and pupal stages do not feed (Cnaani et al. 2002 and Couvillon & Dornhaus 2009). After the switching point, reproductive eggs begin to be laid and the queen ceases to lay worker eggs. However, some developing worker larvae are still present in the colony. Similar to workers, the egg stage of developing reproductives occurs for 5 days (Cnaani et al. 2002). Thus, after $t_s + 5$, when both worker and reproductive larvae are present in the colony, our model assumes the resources returned to the colony are evenly divided among developing worker larvae and developing reproductive larvae. Accordingly, after $t_s + 5$ (i.e. $t > t_s + 5$), the equation for P_t is now given as

$$P_t = \frac{R_t/2}{N_{Lt}} \quad (8)$$

After resources are divided among developing larvae, P_t is fed to each worker larva by adding P_t to each larva's D_t. Accordingly, D_t is a variable tracking the total value of

resources consumed by a worker during larval development, and is calculated separately per worker. In each time step, D_t is updated according to the following equation

$$D_t = D_{t-1} + P_t \quad (9)$$

Cognition

Upon eclosion, the model assumes that worker cognition is fixed and is based upon the value of resources consumed during larval development. Research on insect central nervous system (CNS) development suggests that adult CNS growth is bounded, with a minimum CNS size resulting from reduced growth under nutrient restriction and a maximum CNS size resulting from unrestricted growth under *ad libitum* feeding (Lanet & Maurange 2014; Cheng et al. 2011). We accordingly bound adult learning scores, L_t, between 1 and 2 following a logistic function relating adult learning scores to resources consumed during larval development. Specifically, L_t is determined from a logistic function (Fig S2) given as

$$L_t = \frac{\alpha}{1 + e^{-\kappa(D_t - \delta_0)}} + 1 \quad (10)$$

Here, α gives the logistic curve's maximum value, κ is the logistic curve growth rate, and δ_0 is the value of D_t that gives a corresponding value of $L_t = 1.5$. To bound L_t between 1 and 2, we set $\alpha = 1$ and add 1 to the logistic function. To determine the value of δ_0, we calculated the sum of all consecutive seven-day community-wide sugar content means in the observed environment (i.e. the total value of resources consumed during a larva's

development, assuming she is fed by only one forager), took the median of these values, and divided by two (i.e. as δ_0 is half of the maximum D_t value in the logistic function). We then simulated different values of κ and used in our model the minimum value of κ that resulted in $L_t \geq 1.95$ for δ_0*2.

Finally, to explore how average learning, at the colony-level, may vary across a season, at each time step the average learning score of living adult workers is given as

$$C_t = \frac{\sum_{w=1}^{N_{Wt}} L_{tw}}{N_{Wt}} \quad (11)$$

where C_t is the average learning score of living adult workers (w) in the colony at time t. N_{Wt} gives the total number of living adult workers at the current time step.

We use two statistics to assess changes across the season in C_t: (1) whether C_t is greater or less than 1.5, with $C_t < 1.5$ indicating low colony-level cognition and $C_t > 1.5$ indicating high colony-level cognition, and (2) the slope of the regression line from a linear regression of C_t regressed against t, with higher slope values indicating a more rapid increase in C_t across the season.

Colony Size and Phenology

While all bumble bee colonies are annual, they exhibit marked interspecific variation in size - i.e. the number of adult workers - and phenology (Goulson 2010). To explore how seasonal trends in cognition are affected by colony size, we run our model in each resource environment for a representative small colony and a representative large colony. Furthermore, in our observed resource environment, we explore how seasonal trends in

cognition are affected by colony phenology, by employing a factorial design, whereby a colony may be either small or large and either early- or late-emerging (Fig 2). To parameterize our model for emergence time, we set early-emerging colonies to beginning simulation at $t = 1$ and late-emerging colonies to beginning simulation at $t = 100$. Note that we do not run our model with different emergence times in our stable and pulsed resource environments, as different emergence times would not change the value of resources foragers encounter in these artificial environments. See Fig 2 for a visual depiction of the colony types our model was simulated for in each resource environment.

Across the lifespan of a colony, bumble bee species producing small colonies typically produce around a few hundred workers, while species producing large colonies can produce over a thousand workers (Macfarlane et al. 1994). We thus use observed colony demography data from *Bombus pensylvanicus* (Goldblatt & Fell 1986) and *Bombus atratus* (da Silva-Matos & Garófalo 2000) to parameterize our model for small and large colonies respectively, as observational studies suggest these species' colony sizes fall within these ranges (Macfarlane et al. 1994).

To parameterize our model for colony size, we first created a Leslie matrix per species, parameterized with observed life table data from *B. pensylvanicus* and *B. atratus* (Goldblatt & Fell 1986; da Silva-Matos & Garófalo 2000). Leslie matrices are a common approach to modeling age-structured population growth and decline, which incorporate unique survival and birth rates per age cohort. In our Leslie matrices, colony growth is projected as

$$n_{T+1} = \mathbf{M} n_T \quad (12)$$

In each time step, T, n_T is a population state vector of the number of adult workers, ω_x, in each age cohort, x. **M** is the Leslie matrix. c is the total number of age cohorts. Note that because age cohorts in our observed life tables are divided into five day intervals (Goldblatt & Fell 1986; da Silva-Matos & Garófalo 2000), each time step in our Leslie matrices is equivalent to five days and every x is a five day age cohort (i.e. $x = 1$ is 1-5 days age, $x = 2$ is 6-10 days age, …). Equation 12, the product of **M** and n_T, can also be written as

$$\mathbf{M}n_T = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 & E_T \\ (1-q_1) & 0 & 0 & \dots & 0 & 0 \\ 0 & (1-q_2) & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & 0 \\ 0 & 0 & 0 & \dots & (1-q_c) & 0 \\ 0 & 0 & 0 & \dots & 0 & 1 \end{bmatrix} \begin{bmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \\ \vdots \\ \omega_c \\ Q \end{bmatrix} \quad (13)$$

Q - the number of queens - is set to $Q = 1$ in our Leslie matrices. Mortality of each age cohort, q_x, is taken directly from our observed life tables (Goldblatt & Fell 1986; da Silva-Matos & Garófalo 2000). It is assumed that the queen produces a constant number of workers, given as E_T workers born per time step.

To fit our Leslie matrices to colonies representative of *B. pensylvanicus* and *B. atratus*, for each species we ran the Leslie matrix with their species-specific mortality data (Goldblatt & Fell 1986; da Silva-Matos & Garófalo 2000) and a value of E_T that yields a total worker number that falls within the range of small and large colonies (Macfarlane et al. 1994). Data on *B. pensylvanicus* from Goldblatt & Fell (1986) suggests that the switching point for *B. pensylvanicus* colonies occurs 45 days after the

commencement of worker production. Thus, we set the switching point for our Leslie matrices, T_s, to $T_s = 9$ (i.e. 45 days/5 day intervals), and the switching point for our simulation model to $t_s = 45$. Given a lack of data on large colony switching points, we assume that switching points are consistent between small and large colonies. After values of E_T were determined for both small and large colonies, we set the rate of colony growth, θ, in our simulation model to $\theta = E_T/5$, as time steps in our Leslie matrices (T) are five days intervals, whereas time steps in our simulation model (t) are one day intervals.

Additionally, we parameterize our simulation model so that at the end of each time step, a percentage of living adult workers die according to the five day age cohort mortality rates (q_x) from our observed life tables (Goldblatt & Fell 1986; da Silva-Matos & Garófalo 2000). Specifically, at the end of each time step, each living adult worker (i.e. $t - t_B \geq 25$ and t_D = NA) is subject to mortality, such that a percentage of each five day age cohort (i.e. age cohort 1 = workers with $t - t_B \geq 25$ & $t - t_B \leq 29$, age cohort 2 = workers with $t - t_B \geq 30$ & $t - t_B \leq 34$, …), rounded up to the nearest integer, has t_D assigned to t (i.e. $t_D = t$).

To verify that the demography modeling in our simulation model produces colony growth consistent with our Leslie matrices, we compared colony growth between the two. The best fit between simulation model and Leslie matrix colony growth occurred when we weighted mortality rates in our simulation model by dividing q_x by a constant integer γ. The fit between colony growth in our simulation model and Leslie matrices is shown in Fig 3. For *B. pensylvanicus*, we also compare the fit of our small colony simulation model and Leslie matrix to data on colony growth extracted from Goldblatt & Fell (1986) (Fig 3).

Model Simulation

We ran 1,000 iterations of our model for each colony type in each resource environment (Fig 2) and report results compiled across these iterations. Our model was coded in RStudio (version 0.99.902) and utilized the package MCMCglmm (version 2.29). See table 1 for definitions of all parameters used in the simulation model.

Results

Field Sampling and Learning Tests

Across our five focal species, we sampled a total of 160 worker bumble bees. Sample size varies per species, with a minimum of 24 for *Bombus auricomus* to a maximum of 39 for *Bombus impatiens*. While these species have staggered phenological timing, this sampling occurred from throughout the entire period of foraging worker activity for each species; across all species, the first worker was sampled on May 24th and the last worker was sampled on September 12th.

Two of these five species - *B. auricomus* and *Bombus pensylvancius* - show a significant increase in learning test success across the season (both $p < 0.05$), with average success in the learning test being lower at the beginning of the season than at the end of the season (Fig 4; Table 2). The other three species - *Bombus bimaculatus*, *Bombus griseocollis*, and *B. impatiens* - do not show a significant trend in learning test success across the season (all $p > 0.05$) (Fig 4; Table 2). Our randomization tests suggest that the significant trends identified for *B. auricomus* and *B. pensylvancius* are not a result of uneven temporal sampling across the season. The observed *z*-values of both *B.*

auricomus and *B. pensylvancius* fall within the top 2.5% tail of their respective randomized *z*-value distributions (Fig S3). None of the observed *z*-values of *B. bimaculatus*, *B. griseocollis*, and *B. impatiens* fall within the top or bottom 2.5% tails of their randomized *z*-value distributions (Fig S3).

Simulation Model

Bumble Bee Demography

Fitting our model to small and large colonies using demographic data on *Bombus pensylvanicus* and *Bombus atratus*, following Leslie matrix simulations, we set $\theta = 7$ (i.e. $E_t = 35$) and $\theta = 25$ (i.e. $E_t = 125$) for small and large colonies respectively. With $t_s = 45$, these growth rates yield a total of 315 workers produced by small colonies and 1,125 workers produced by large colonies, which fall within known worker number ranges for small and large bumble bee colonies (Macfarlane et al. 1994). Across 1,000 simulations, this resulted in a max of ~204 workers and ~681 workers being alive in the colony at the peak of colony growth (i.e. $t_s + 25$; the first time step after t_s when all workers have eclosed) for small and large colonies respectively. Colony growth in our simulation models was fit to Leslie matrix colony growth projections with $\gamma = 2.5$ for small colonies and $\gamma = 5.0$ for large colonies (Fig 3). The small colony growth projections roughly match the *B. pensylvanicus* colony growth reported by Goldblatt & Fell (1986) (Fig 3).

Foraging and Resources

As foraging worker number (N_F) is always 88-94% of living adult workers in the colony (O'Donnell 2000), N_F follows overall colony growth (N_W), with the number of foragers in

the colony steadily increasing up to the first time step after t_s when all workers have eclosed (i.e. t_s + 25) and decreasing thereafter. Across 1,000 simulations, a max of ~184 and ~618 foragers, not including the queen, were in the colony at the peak of colony growth (i.e. t_s + 25) for small and large colonies respectively. Resources returned to the colony (R_t) follow this pattern of N_F across time, with R_t generally increasing up to the peak of colony growth and decreasing thereafter (Fig S4). This general trend in R_t across time is observed for all colonies; however, the exact shape of R_t across the season is dependent on the resource environment (Fig S4). Additionally, R_t is always higher for large colonies than for small colonies, regardless of resource environment. The resource value fed to each worker larva (P_t) generally increases up to the last time step worker larvae are fed (i.e. t_s + 15) (Fig S5); however, the exact shape of this increase in P_t across time is dependent on the resource environment (Fig S5). In pulsed environments with the low pulse occurring first, a decrease in P_t occurs prior to the last time step worker larvae are fed, at the time step at which R_t is first divided among developing worker and developing reproductive larvae (i.e. t_s + 5; Equation 8); however, this is not observed for the pulsed environment with the high pulse occurring first (i.e. pulsed high 1st environment), as this time step (t_s + 5) occurs immediately before the beginning of the second high resource pulse (i.e. t = 51). While R_t is greater for large colonies compared to small colonies, P_t is not similarly affected by colony size.

Cognition

We use two statistics to assess changes across the season in colony average learning (C_t): (1) whether C_t is greater or less than 1.5, and (2) the slope of the regression line from a

linear regression of C_t regressed against t. As small colonies have protracted colony lifespans relative to large colonies - owing to differences between *B. pensylvanicus* and *B. atratus* demography (Goldblatt & Fell 1986; da Silva-Matos & Garófalo 2000) - prior to running these regressions we truncated t for small colonies to match the range of t observed in large colonies. This resulted in each regression being constricted to $t = 26$ through $t = 115$, thus ensuring that differences between small and large colony regression line slopes are not attributed to the protracted lifespan of small colonies.

Colonies simulated with the observed resource environment (Timberlake et al. 2019b) show an increase in colony average learning (C_t) across colony development (Fig 5). This increase in C_t across the season occurs regardless of colony size or emergence time, with $C_t < 1.5$ at the beginning of colony development (i.e. early-emerging $t \leq 53$; late-emerging $t \leq 50$) and $C_t > 1.5$ at the end (i.e. early-emerging $t \geq 54$; early-emerging $t \geq 51$). The increase in C_t across the season is more rapid for large colonies than for small colonies, with the regression line slopes for large colonies being greater than for small colonies (i.e. small ≈ 0.009; large ≈ 0.013). This difference in the rate of C_t increase is due to different mortality schedules between small and large colonies, rather than different rates of colony growth (θ) (see Supplemental Materials for details; Fig S6).

In the stable low resource environment, colonies similarly show an increase in C_t across colony development, regardless of colony size (Fig 5); $C_t < 1.5$ at the beginning of colony development (i.e. small $t \leq 66$; large $t \leq 63$) and $C_t > 1.5$ at the end (i.e. small $t \geq 67$; large $t \geq 64$). Similar to the observed resource environment, C_t increases more rapidly for large colonies than for small colonies, as evidenced by a steeper regression line slope for large colonies (i.e. small ≈ 0.008; large ≈ 0.014). In the stable high resource

environment, C_t is high across the entire season; across all time steps, $C_t > 1.5$ for both small and large colonies. Accordingly, the regression line slopes for colonies in the stable high environment are relatively shallow (i.e. small ≈ 0.000; large ≈ 0.002). Notably, for large colonies in the stable high environment, there is a slight decrease in C_t across time for the time steps in which adult workers received food during larval development from only the queen (i.e. $t = 26$ through $t = 36$). After this period, when eclosing workers were also fed by foraging workers during larval development (i.e. $t \geq 37$), C_t increases across time. Despite this decrease in C_t at the beginning of colony development, C_t remains above 1.5 across the entire season in the stable high environment.

In the pulsed resource environments, the trends in C_t across the season mirror the trends observed in the stable resource environments (Fig 5). Specifically, (1) colonies simulated in the pulsed environment with the low pulse occurring first (i.e. pulsed low 1st environment) show an increase in C_t across colony development, regardless of colony size, similar to stable low environments; $C_t < 1.5$ at the beginning of colony development (i.e. small $t \leq 45$; large $t \leq 46$) and $C_t > 1.5$ at the end (i.e. small $t \geq 46$; large $t \geq 47$). Analogous to the observed and stable low environments, large colonies simulated in the pulsed low 1st environment exhibit a more rapid increase in C_t across the season compared to small colonies; i.e. the regression line slope is steeper for large colonies than for small colonies (i.e. small ≈ 0.009; large ≈ 0.012). (2) Colonies simulated in the pulsed environment with the high pulse occurring first (i.e. pulsed high 1st environment) show C_t high across the entire season, similar to the stable high environment (Fig 5); across all time steps, $C_t > 1.5$ for both small and large colonies. The regression line slopes for colonies in the pulsed high 1st environment are relatively shallow (i.e. small ≈ 0.000;

large ≈ 0.004). Notably, comparing large colonies between the stable high environment and the pulsed high 1st environment, there is similarly a decrease in C_t across time at the beginning of colony development; however, in the pulsed high 1st environment, after the first time step at which eclosing workers were fed by foraging workers during larval development (i.e. t = 36), C_t continues to decrease for another 10 time steps before beginning to rebound. Despite this decrease, C_t remains above 1.5 across the entire season in the pulsed high 1st environment.

Discussion

We find evidence that field bumble bee populations and simulated colonies can exhibit seasonal trends in cognition, thus supporting our hypothesis. Specifically, we find (1) two of our focal species - *Bombus auricomus* and *Bombus pensylvanicus* - significantly increased in learning test success across the season (Fig 4), and (2) our simulated colonies increased average colony level learning (C_t) across the season, except in environments providing high resource values early in colony development, which yielded persistently high C_t values across the season (Fig 5). The increase in C_t across the season resulted from larvae being provisioned with higher value resources as the season progressed, following known dynamics of bumble bee demography (Goldblatt & Fell 1986; da Silva-Matos & Garófalo 2000; Cnaani et al. 2002) and foraging ecology (O'Donnell et al. 2000). Collectively, these results support the idea that phenological trends in cognition may exist in certain populations, particularly those in which individuals are produced at different time points and cognition is developmentally plastic. Prior studies have suggested that changes to cognitive abilities can occur across

individual lifetimes (e.g. Shettleworth 2010) and across evolutionary time (e.g. Dunlap & Stephens 2009; Mery & Kawecki 2002, 2004; Stephens 1991; Dunlap et al. 2009). This study builds upon this literature by suggesting that changes to cognitive ability can also occur across seasons, at the population-level.

A wealth of literature has documented how behavior can exhibit phenological patterns, due to behavioral events synchronizing with resource phenology (e.g. McGrath et al. 2009; García-Navas & José Sanz; Minckley et al. 1994). This is the first study, to the best of our knowledge, to suggest that cognitive abilities *per se* may also exhibit seasonal changes at the population-level. This is notable as cognitive abilities of a population's individual constituents can affect population dynamics and community interactions (Fryxell et al. 2005, Snell-Rood 2013; Wong & Candolin 2015; Tuomainen & Candolin 2011). For example, increased cognitive complexity is associated with a greater ability to plastically match behavior to current environmental conditions (Mikhalevich et al. 2017), which helps ensure population success in rapidly altered environments (e.g. Sol et al. 2008; Sol et al. 2002; Amiel et al. 2011). This is particularly pertinent in plant-pollinator communities, as pollinator populations must contend with changing associations between environmental stimuli due to floral turnover across a season, with additional interannual variance induced by anthropogenic global change (e.g. Cleland et al. 2007; Hegland et al. 2009). The literature has increasingly appreciated the value of incorporating analysis of intraspecific trait variation into considerations of how populations respond to environmental variability (e.g. Austin & Dunlap 2019; Forsam 2014; González-Suárez et al. 2015). This study suggests that how intraspecific

cognitive abilities change temporally across seasons should be included in these considerations.

The finding that pulsed resource environments yield C_t trends that are analogous to stable resource environments suggests that seasonal trends in cognition are more dependent on the resource environment early in colony development, as opposed to resource environments that occur later in a colony's life. In other words, C_t values increase across the season in environments providing either a low pulse first or a consistently low resource value across the season, while C_t values are high across the entire season in environments providing either a high pulse first or a consistently high resource value (Fig 5). Thus, our simulation model suggests that seasonal trends in population-level cognition are dependent on whether a colony first encounters a low- or high-quality resource environment. This is notable as many flowering plant communities produce nectar in pulses across a season, with a typical period of early-spring dearth followed by alternating periods of bloom and nectar scarcity (e.g. Timberlake et al. 2019a; Hemberger et al. 2020). Furthermore, emerging literature suggests that early in colony development, when provisioning of larvae is dependent on only the queen and several foragers, low resource availability can have immediate and persistent effects on colony fitness (e.g. Woodard et al. 2019). By decreasing the ability of individuals to plastically match behavior to changing environmental conditions, low colony-level cognitive abilities early in colony development may add further stress to colonies during this sensitive period in colony development. It is notable that the two species we find evidence for a cognitive phenology in - *B. auricomus* and *B. pensylvanicus* - are the two out of our five focal species that have declined in relative abundance across North

American over the past century (Hatfield et al. 2015). Future research should explore whether these species encounter low quality resource environments early in colony development, and, if so, whether such resource dearth produces low cognitive abilities in early worker cohorts that enacts a fitness cost on the colony.

While our field data suggest that populations can exhibit seasonal trends in learning ability, these data are limited by (1) constituting only one population in one season, for each species, and (2) not assessing the mechanism that underlies seasonal trends in learning. Coupling these data with a simulation model, our simulation model suggests that such seasonal trends in learning are driven by larvae being provisioned with different resource values across the season. Whenever novel theoretical models are developed in ecology, a tension exists between making models grounded in natural history and making models generalizable across species (Dunlap et al. 2019). We have approached this tension by parameterizing our model with observed data on representative bumble bee species (Goldblatt & Fell 1986; da Silva-Matos & Garófalo 2000; Cnaani et al. 2002; O'Donnell et al. 2000), while making several simplified assumptions throughout the model. Notably, our simulation model assumes a summative relationship between resources consumed during larval development (D_t) and adult cognition (L_t). While the relationship between resources consumed during juvenile ontogeny and adult cognition is not merely this simple across taxa, research suggests that a positive relationship between nutritional consumption during ontogeny and cognitive development is indeed real in many species (e.g. Lanet & Maurange 2014; Cheng et al. 2011). For example, Cheng et al. (2011) demonstrate, while central nervous system (CNS) development in insects is spared relative to other organs under nutrient restriction,

overall CNS volume is lower when nutrients are restricted during development relative to *ad libitum* feeding - a phenomenon analogous to brain-sparing in the last third of mammalian pregnancy (Lanet & Maurange 2014). Due to the positive relationship between nutritional consumption during ontogeny and development of adult cognition in various species (e.g. Lanet & Maurange 2014; Cheng et al. 2011), we argue that seasonal trends in population-level learning are likely to occur in diverse species, as opposed to our study system alone.

This study provides evidence for seasonal changes in population-level cognition, by taking direct measurements of bumble bee learning abilities and developing a simulation model of temporal changes to colony-level learning. While our results do not rule out the possibility that seasonal changes in population-level cognition only manifest in certain populations under specific environmental conditions, our study suggests a novel level of analysis for variation in cognitive abilities: the population-level across seasons. Such seasonal changes in population-level cognition may be particularly prone to manifesting in systems where cognition is developmentally plastic and individuals of a given generation are produced at different time points throughout a reproductive season (e.g. Szigeti et al. 2019). Future research should explore whether phenological trends in cognition occur in diverse taxa, while directly quantifying the environmental conditions that drive cognitive development, and may thus underlie phenological trends in cognition. To fully understand the eco-evolutionary implications of animal cognition, our study suggests that phenological changes to cognition be considered.

Acknowledgements

We thank Shaw Nature Reserve, EarthDance Farms, Calvary Cemetery, Castlewood State Park, and their employees for access to these sites, which made this work possible. In particular, we acknowledge James Trager (Shaw Nature Reserve) and Eric Stevens (EarthDance) who provided helpful insight to learning about the biota of their respective sites. We thank the Missouri Department of Natural Resources for granting us permission to sample at Castlewood State Park. We thank Nathan Muchhala and various members of ASD's laboratory for their valuable feedback and discussion on this project. This study was funded by grants from the Whitney R. Harris World Ecology Center, Trans World Airlines, and the Webster Groves Nature Study Society to MWA. MWA was partially supported by the Raven Fellowship and the Graduate Dissertation Fellowship through the University of Missouri - St. Louis.

Funding

This work was funded by grants to MWA from the Whitney R. Harris World Ecology Center, Trans World Airlines, and the Webster Groves Nature Study Society. MWA was partially funded by the Raven Fellowship (spring 2019) and the Graduate Dissertation Fellowship (fall 2019, spring 2020) through the University of Missouri - St. Louis.

References

Amiel, J.J., Tingley, R., & Shine, R. 2011. Smart moves: effects of relative brain size on establishment success of invasive amphibians and reptiles. PLoS ONE 6:e18277.

Arien, Y., Dag, A., & Shafir, S. 2018. Omega-6:3 ratio more than absolute lipid level in diet affects associative learning in honey bees. Frontiers in Psychology 9:1001.

Arnold, K.E., Ramsay, S.L., Donaldson, C., & Adam, A. 2007. Parental prey selection affects risk-taking behaviour and spatial learning in avian offspring. Proceedings of the Royal Society B 274:2563-2569.

Austin, M.W. & Dunlap, A.S. 2019. Intraspecific variation in worker body size makes North American bumble bees (*Bombus* spp.) less susceptible to decline. The American Naturalist 194:381-394.

Austin, M.W., Tripodi, A.D., Strange, J.P., & Dunlap, A.S. *in-review*. Bumble bees exhibit intraspecific body size spatial structuring despite low genetic differentiation.

Barbeito-Andrés, J., Castro-Fonseca, E., Qiu, L.R., Bernal, V., Lent, R., Henkelman, M., Lukowiak, K., Gleiser, P.M., Hallgrimsson, B., & Gonzalez, P.N. 2019. Region-specific changes in *Mus musculus* brain size and cell composition under chronic nutrient restriction. Journal of Experimental Biology 222:jeb204651

Bartomeus, I., Ascher, J. S., Wagner, D., Danforth, B. N., Colla, S., Kornbluth, S., & Winfree, R. 2011. Climate-associated phenological advances in bee pollinators and bee-pollinated plants. Proceedings of the National Academy of Sciences of the United States of America 108:20645–20649.

Baude, M., Kunin, W.E., Boatman, N.D., Conyers, S., Davies, N., Gillespie, M.A., Morton, R.D., Smart, S.M., & Memmott, J. 2016. Historical nectar assessment reveals the fall and rise of floral resources in Britain. Nature 530:85–88.

Brust, V., Krüger, O., Naguib, M., & Krause E.T. 2014. Lifelong consequences of early nutritional conditions on learning performance in zebra finches (*Taeniopygia guttata*). Behavioural Processes 103:320-326.

Camilo, G.R., Muñiz, P.A., Arduser, M.S., & Spevak, E.M. 2018. A checklist of the bees (Hymenoptera: Apoidea) of St. Louis, Missouri, USA. Journal of the Kansas Entomological Society 90:175-188.

Cheng, L.Y., Bailey, A.P., Leevers, S.J., Ragan, T.J., Driscoll, P.C., & Gould, A.P. 2011. Anaplastic lymphoma kinase spares organ growth during nutrient restriction in Drosophila. *Cell* 146:435–447.

Chmura, H.E., Kharouba, H.M., Ashander, J., Ehlman, S.M., Rivest, E.B., & Yang, L.H. 2019. The mechanisms of phenology: the patterns and processes of phenological shifts. Ecological Monographs 89:e01337.

Cleland, E.E., Chuine, I., Menzel, A., Mooney, H. A., & Schwartz, M.D. 2007. Shifting plant phenology in response to global change. Trends in Ecology and Evolution 22:357–365.

Cnaani, J., Schmid-Hempel, R., & Schmidt, J.O. 2002 Colony development, larval development and worker reproduction in *Bombus impatiens* Cresson. Insectes Soc. 49:164–170.

Collado, M.Á., Montaner, C.M., Molina, F.P., Sol, D., & Bartomeus, I. 2020. Brain size predicts learning abilities in bees. bioRxiv 2020.01.27.921114.

Couvillon, M.J. & Dornhaus, A. 2009. Location, location, location: Larvae position inside the nest is correlated with adult body size in worker bumble-bees (*Bombus impatiens*). Proceedings of the Royal Society B: Biological Sciences 276:2411-2418.

da Silva-Matos, E.V. & Garófalo, C.A. 2000. Worker life tables, survivorship, and longevity in colonies of *Bombus* (*Fervidobombus*) *atratus* (Hymenoptera: Apidae). Rev. Biol. Trop. 48:657-664.

Dunlap, A.S. & Stephens, D.W. 2009. Components of change in the evolution of learning and unlearned preference. Proceedings of the Royal Society B: Biological Sciences 276:3201–3208.

Dunlap, A.S., Austin, M.W., & Figueiredo, A. 2019. Components of change and the evolution of learning in theory and experiment. Animal Behaviour 147:157-166.

Dunlap, A.S., McLinn, C.M., MacCormick, H.A., Scott, M.E., & Kerr, B. 2009. Why some memories do not last a lifetime: dynamic long-term retrieval in changing environments. Behavioral Ecology 20:1096–1105.

Farris, S.M. & Roberts, N.S. 2005. Coevolution of generalist feeding ecologies and gyrencephalic mushroom bodies in insects. Proceedings of the National Academy of Sciences of the United States of America 102:17394-17399.

Forsam, A. 2014. Effects of genotypic and phenotypic variation on establishment are important for conservation, invasion, and infection biology. Proceedings of the National Academy of Sciences of the United States of America 111:302-307.

Fryxell, J.M., Wilmshurst, J.F., Sinclair, A.R.E., Haydon, D.T., Holt, R.D., & Abrams, P.A. 2005. Landscape scale, heterogeneity, and the viability of Serengeti grazers. Ecology Letters 8:328-335.

García-Navas, V. & Sanz, J.J. 2011. The importance of a main dish: nestling diet and foraging behaviour in Mediterranean blue tits in relation to prey phenology. Oecologia 165:639-649.

Goldblatt, J.W. & Fell, R.D. 1987. Adult longevity of workers of the bumble bees *Bombus fervidus* (F.) and *Bombus pensylvanicus* (De Geer) (Hymenoptera: Apidae). Canadian Journal of Zoology 65:2349-2353.

González-Suárez, M., Bacher, S., & Jeschke, J.M. 2015. Intraspecific trait variation is correlated with establishment success of alien mammals. The American Naturalist 185:737-746.

Goulson, D. 2010. Bumblebees: Behaviour, Ecology, and Conservation. Oxford University Press.

Hatfield, R., Colla, S., Jepsen, S., Richardson, L., Thorp, R., & Jordan, S. 2015. IUCN assessments of North American *Bombus* spp. http://www.xerces.org/.

Hegland, S.J., Nielsen, A., Lázaro, A., Bjerknes, A.L., & Totland, Ø. 2009. How does climate warming affect plant-pollinator interactions? Ecology Letters 12:184–195.

Hemberger, J., Frappa, A., Witynski, G., & Gratton, C. 2020. Saved by the pulse? Separating the effects of total and temporal food abundance on the growth and reproduction of bumble bee microcolonies. Basic and Applied Ecology 45:1-11.

Julian, G.E. & Gronenberg, W. 2002. Reduction of brain volume correlates with behavioral changes in queen ants. Brain, Behavior and Evolution 60: 152-164.

Koyama, T., Mendes, C.C., & Mirth, C.K. 2013. Mechanisms regulating nutrition-dependent developmental plasticity through organ-specific effects in insects. Frontiers in Physiology 4:263.

Lanet, E. & Maurange, C. 2014. Building a brain under nutritional restriction: insights on sparing and plasticity from *Drosophila* studies. Frontiers in Physiology 5:117.

Lefebvre, L., Whittle, P., Lascaris, E., & Finkelstein, A. 1997. Feeding innovations and forebrain size in birds. Animal Behaviour 53:549-560.

Macevicz, S. & Oster, G. 1976. Modeling social insect populations. 2. Optimal reproductive strategies in annual eusocial insect colonies. Behavioral Ecology and Sociobiology 1:265–282.

Macfarlane, R.P., Patten, K.D., Royce, L.A., Wyatt, B.K.W., & Mayer, D.F. 1994. Management potential of sixteen North American bumble bee species. Melanderia 50:1–12.

McGrath, L.J., van Riper III, C., & Fontaine, J.J. 2009. Flower power: tree flowering phenology as a settlement cue for migrating birds. Journal of Animal Ecology 78:22-30.

Mery, F. & Kawecki, T.J. 2002. Experimental evolution of learning ability in fruit flies. PNAS 99:14274–14279.

Mery, F. & Kawecki, T.J. 2004. An operating cost of learning in Drosophila melanogaster. Animal Behaviour 68:589–598.

Mikhalevich, I., Powell, R., & Logan, C. 2017. Is behavioural flexibility evidence of cognitive complexity? How evolution can inform comparative cognition. Interface Focus DOI: 10.1098/rsfs.2016.0121.

Moda, L.M., Vieira, J., Guimarães Freire, A.C., Bonatti, V., Bomtorin, A.D., Barchuk, A.R., & Simoes, Z.L.P. 2013. PLoS One 8:e64815.

Muth, F., Breslow, P.R., Masek, P., & Leonard, A.S. 2018. A pollen fatty acid enhances learning and survival in bumblebees. Behavioral Ecology 29:1371–1379.

Muth, F., Cooper, T.R., Bonilla, R.F. & Leonard, A.S. 2017. A novel protocol for studying bee cognition in the wild. Methods in Ecology and Evolution 2017:1-10.

O'Donnell, S., Reichardt, M., & Foster, R. 2000. Individual and colony factors in bumble bee division of labor (*Bombus bifarius nearcticus* Handl; Hymenoptera, Apidae). Insectes Sociaux 47:164-170.

Ogilvie, J.E. & Forrest, J.R. 2017. Interactions between bee foraging and floral resource phenology shape bee populations and communities. Current Opinion in Insect Science 21:75-82.

Pereboom, J.J.M., Velthuis, H.H.W., & Duchateau, M.J. 2003. The organisation of larval feeding in bumblebees (Hymenoptera, Apidae) and its significance to caste differentiation. Insectes Sociaux 50:127-133.

Shettleworth, S.J. 2010. Cognition, Evolution, and Behavior, Oxford University Press, New York.

Shoji, A., Aris-Brosou, S., Culina, A., Fayet, A., Kirk, H., Padget, O., Juarez-Martines, I., Boyle, D., Nakata, T., Perrins, C.M., & Guilford, T. 2015. Breeding phenology and winter activity predict subsequent breeding success in a trans-global migratory seabird. Biology Letters 11:20150671.

Snell-Rood, E.C. 2013. An overview of the evolutionary causes and consequences of behavioural plasticity. Animal Behaviour 85:1004–1011.

Sol, D., Bacher, S., Reader, S.M., & Lefebvre, L. 2008. Brain size predicts the success of mammal species introduced into novel environments. The American Naturalist 172:S63-S71.

Sol, D., Duncan, R.P., Blackburn, T.M., Cassey, P., & Lefebvre, L. 2005. Big brains, enhanced cognition, and response of birds to novel environments. Proceedings of the National Academy of Sciences of the United States of America 102:5460-5465.

Sol, D., Timmermans, S., & Lefebvre, L. 2002. Behavioural flexibility and invasion success in birds. Animal Behaviour 63:495-502.

Steijven, K., Spaethe, J., Steffan-Dewenter, I., Härtel, S. 2017. Learning performance and brain structure of artificially-reared honey bees fed with different quantities of food. PeerJ 5:e3858.

Stephens, D.W. 1991. Change, regularity, and value in the evolution of animal learning. Behavioral Ecology 2:77–89.

Szigeti, V., Kőrösi, Á., Harnos, A., & Kis, J. 2019. Lifelong foraging and individual specialisation are influenced by temporal changes of resource availability. Oikos 128:649–658.

Timberlake, T.P., Vaughan, I.P., & Memmott, J. 2019a. Phenology of farmland floral resources reveals seasonal gaps in nectar availability for bumblebees. Journal of Applied Ecology 56:1585-1596.

Timberlake, T.P., Vaughan, I.P., & Memmott, J. 2019b. Data from: Phenology of farmland floral resources reveals seasonal gaps in nectar availability for bumblebees. Dryad Digital Repository.

Tuomainen, U. & Candolin, U. 2011. Behavioural responses to human-induced environmental change. Biological Reviews 86:640–657.

van Asch, M. & Visser, M.E. 2007. Phenology of Forest Caterpillars and their host trees: the importance of synchrony. Annual Reviews in Entomology 52:37-55.

Visser, M.E. & Both, C. 2005. Shifts in phenology due to global climate change: The need for a yardstick. Proceedings of the Royal Society B: Biological Sciences 272:2561–2569.

Wong, B.B.M. & Candolin, U. 2015. Behavioral responses to changing environments. Behavioral Ecology 26:665–673.

Woodard, H.S., Duennes, M.A., Watrous, K.M., & Jha, S. 2019. Diet and nutritional status during early adult life have immediate and persistent effects on queen bumble bees. Conservation Physiology 7:coz048.

Tables

Table 1. Parameters used in simulation model.

Parameter	Meaning
Simulation Model	
t	time step in simulation model
B_t	number of worker eggs laid
θ	colony growth rate
$\mathbf{A}_t$	matrix of all individuals in the colony
A_{jk}	variables in $\mathbf{A}_t$ matrix
k	parameters in $\mathbf{A}_t$ matrix
j	individuals in $\mathbf{A}_t$ matrix
J	total number of individuals in $\mathbf{A}_t$ matrix
D_t	total resource value consumed during larval development; each worker has a different D_t
L_t	adult worker learning score; each worker has a different L_t
t_B	time step of birth; i.e. time step a worker egg is laid
t_D	time step of death
Z_t	proportion of adult workers designated as foragers
$\mathbf{F}_t$	matrix of all foraging workers
F_{hi}	variables in $\mathbf{F}_t$ matrix
i	parameters in $\mathbf{F}_t$ matrix
h	individual foragers in $\mathbf{F}_t$ matrix
H	total number of foragers in $\mathbf{F}_t$ matrix
d_t	resource value encountered by a forager; each forager has a different d_t
l_t	forager learning score; each forager has a different l_t
N_{Ft}	total number of foragers
N_{Lt}	total number of worker larvae
N_{Wt}	total number of workers
$\mathbf{S}$	matrix of resource values in the environment
R_t	resources returned to colony
P_t	resource value fed to each worker larvae
C_t	colony average learning
q_x	mortality rate per age cohort
γ	mortality weighting integer
w	living adult workers
Logistic Function	
α	logistic curve's maximum value
κ	logistic curve growth rate
δ_0	value of D_t giving a corresponding value of $L_t = 1.5$
Leslie Matrices	
T	time step in Leslie matrices
$\mathbf{M}$	Leslie matrix
n_T	population state vector of the number of adult workers
E_T	workers born per time step (T)
x	five day age cohorts
Q	number of queens in Leslie matrix
ω_x	number of adult workers per age cohort
c	number of age cohorts

Table 2. Logistic regression results from field learning tests. Significant *p*-values are indicated in bold. SE = standard error

Species	**SE**	***z***	***p***
Bombus auricomus	0.022	2.397	**<0.05**
Bombus bimaculatus	0.028	-0.745	0.456
Bombus griseocollis	0.021	1.243	0.214
Bombus impatiens	0.014	1.534	0.125
Bombus pensylvanicus	0.028	2.298	**<0.05**

Figures

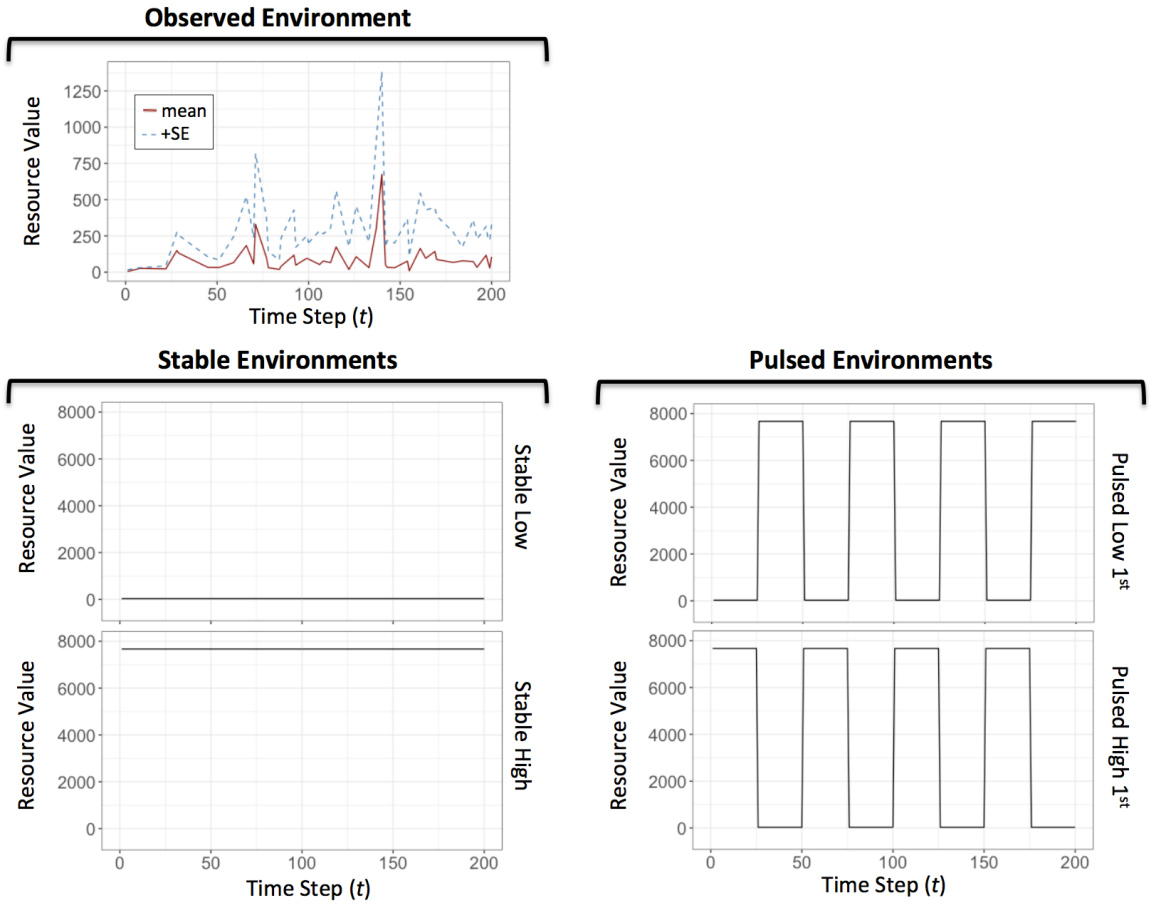

Figure 1. Resource environments (**S**) in the simulation model. In the observed environment, the solid red line gives the mean community-wide sugar content values from Timberlake et al. (2019b) and the dotted blue line gives the mean +SE.

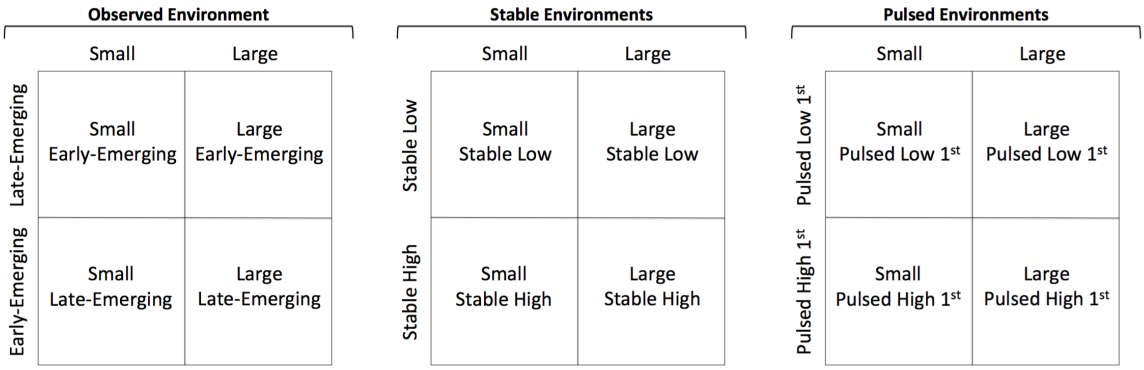

Figure 2. Depiction of the 12 colony types the simulation model was run for across each resource environment.

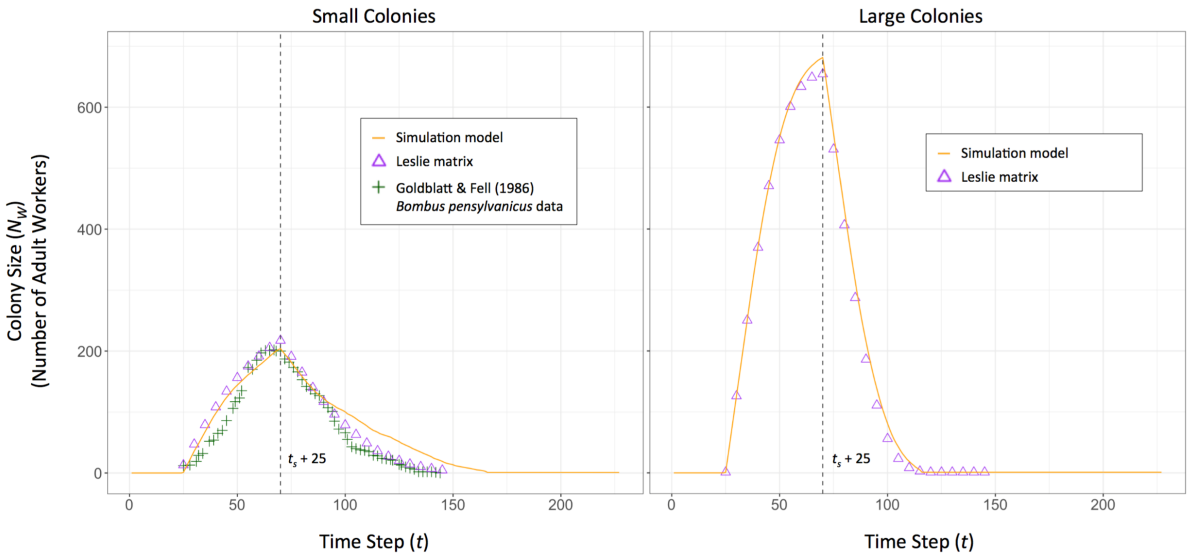

Figure 3. Colony growth (N_W) comparisons between our simulation model and Leslie matrices for small (left panel) and large (right panel) colonies. For small colonies, colony growth is also plotted against *Bombus pensylvanicus* colony growth data from Goldblatt & Fell (1986). Vertical dashed lines give the time point when the last workers in the colony eclose; i.e. the switching point (t_s) plus the period of worker development.

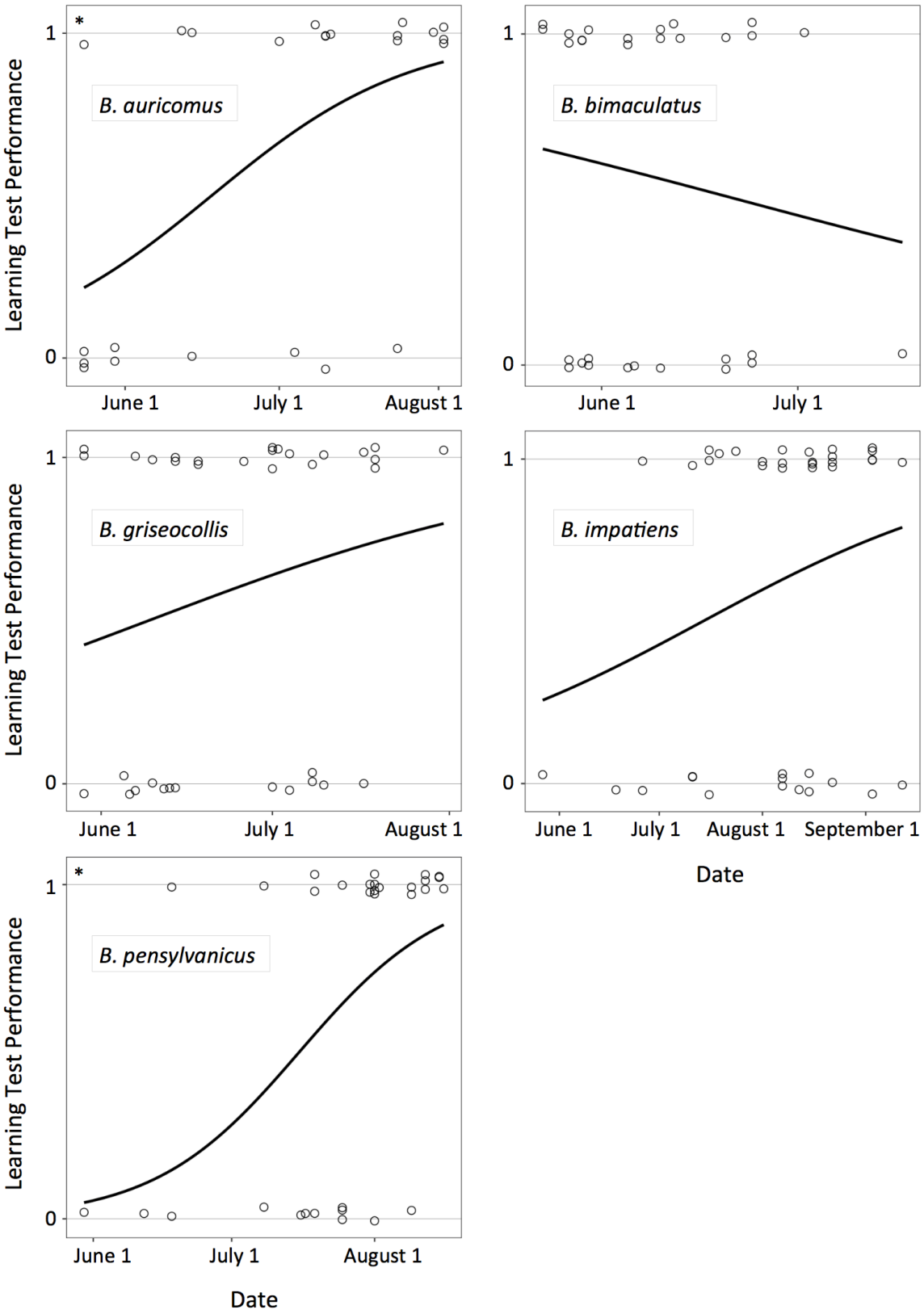
*
1
0
B. auricomus
Learning Test Performance
June 1
July 1
August 1
1
0
B. bimaculatus
June 1
July 1
1
0
B. griseocollis
Learning Test Performance
June 1
July 1
August 1
1
0
B. impatiens
June 1
July 1
August 1
September 1
Date
*
1
0
B. pensylvanicus
Learning Test Performance
June 1
July 1
August 1
Date

Figure 4. Seasonal trends in field learning test performance. Each point is one worker bumble bee (*Bombus* spp.); points are vertically offset to avoid complete overlap of bees tested on the same date. 1 = success in the learning test, 0 = failure in the learning test. Asterisks (*) indicate statistically significant ($p<0.05$) logistic regressions.

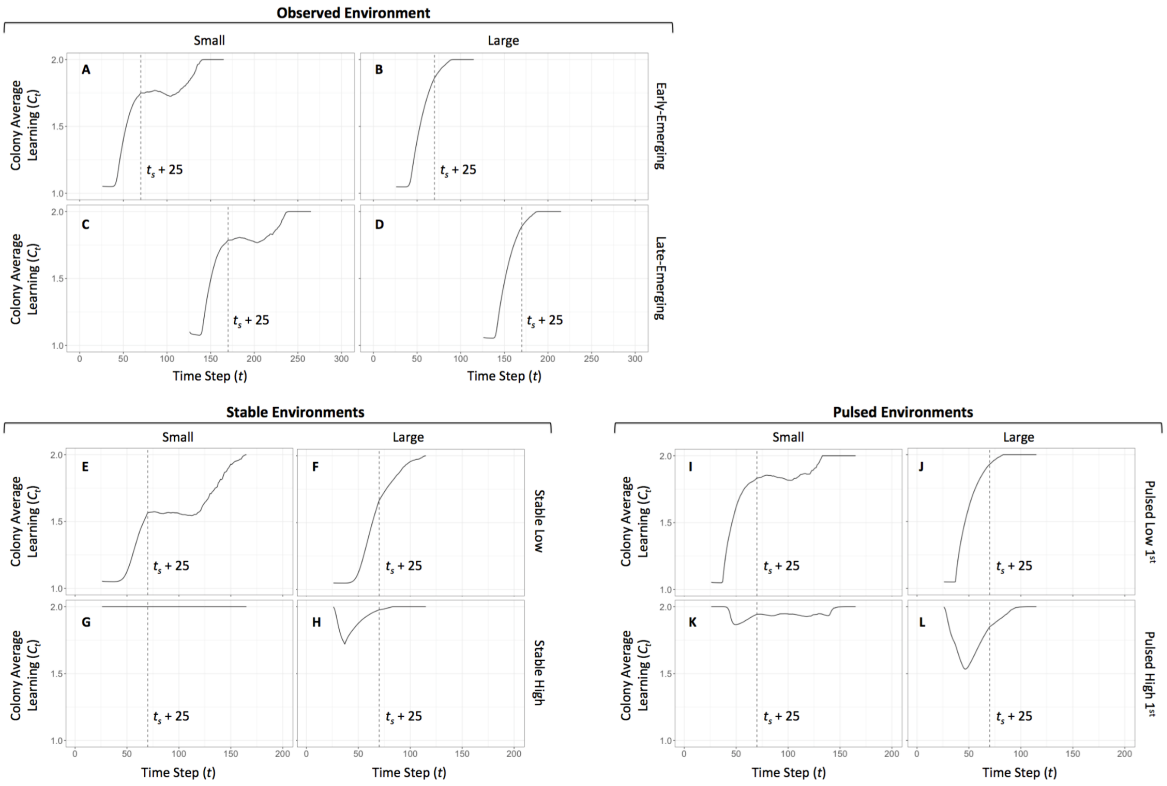

Figure 5. Average colony-level learning (C_t) across time steps (t) for each resource environment: (A-D) observed environment, (E-H) stable environments, and (I-L) pulsed environments. Vertical dashed lines give the time point when the last worker cohort ecloses; i.e. the switching point (t_s) plus the period of worker development. Across all panels, solid lines give the average of 1,000 simulations.

Supplemental Materials

Differences in rate of C_t increase between small and large colonies

In the observed resource environment, the increase in colony average learning (C_t) across the season is more rapid for large colonies than for small colonies, as evidenced by the regression line slopes for large colonies being greater than for small colonies (i.e. small ≈ 0.009; large ≈ 0.013). Small and large colonies differ based on (1) different colony growth rates (θ) and (2) different mortality schedules. Specifically, small colonies have θ = 7 and mortality parameterized following the *Bombus pensylvanicus* mortality schedule, while large colonies have θ = 25 and mortality parameterized following the *Bombus atratus* mortality schedule.

To determine whether the difference in rate of C_t increase across the season between small and large colonies is due to either different colony growth rates or different mortality schedules, we ran our model in the observed environment for the alternate unique pairings of these variables. In other words, we ran our model for (1) colonies parameterized with the *B. pensylvanicus* mortality schedule and θ = 25, and (2) colonies parameterized with the *B. atratus* mortality schedule and θ = 7. We ran both of these colony types for both early- and late-emerging colonies. From 1000 iterations of each of these models, we find that the mortality schedule, as opposed to θ, drives the difference in rate of C_t increase across the season between small and large colonies. This difference is the rate of C_t increase is not affected by colony emergence.

Figures

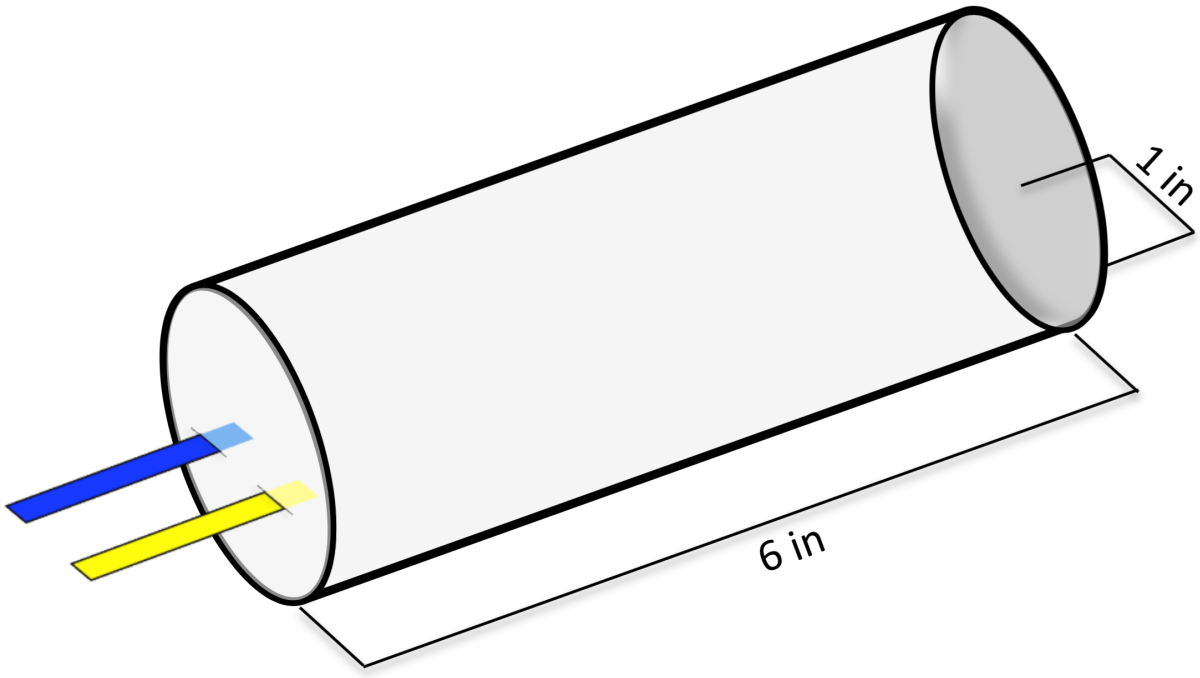

Figure S1. Diagram of vial used for field learning tests, utilizing the Free-Moving Proboscis Extension Response. For testing, a single bee is placed in the vial and allowed to drink from a blue and/or yellow strip of paper inserted into the vial's anterior end. Vial design adapted from Muth et al. (2017).

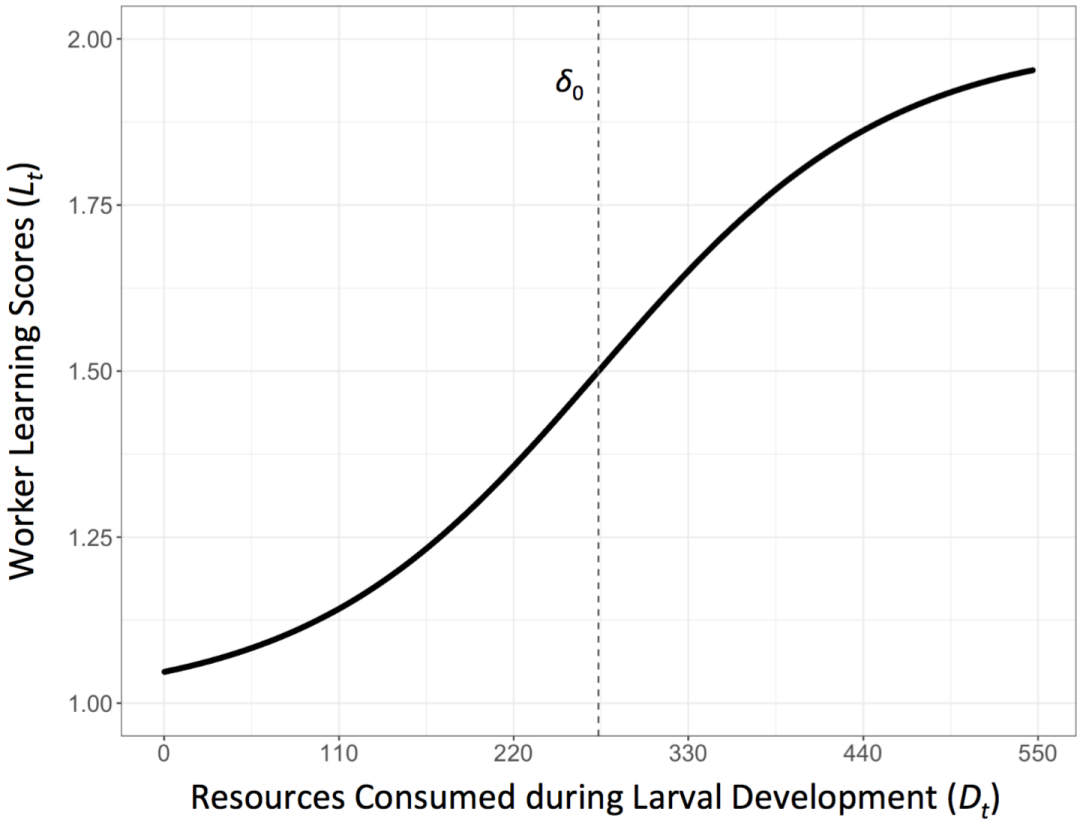

Figure S2. Logistic curve relating adult worker learning scores (L_t) to resources consumed during larval development (D_t). The logistic function producing this curve follows equation 10, with $\alpha = 1$, $\kappa = 0.011$, and $\delta_0 = 273.4$ (i.e. 546.8/2).

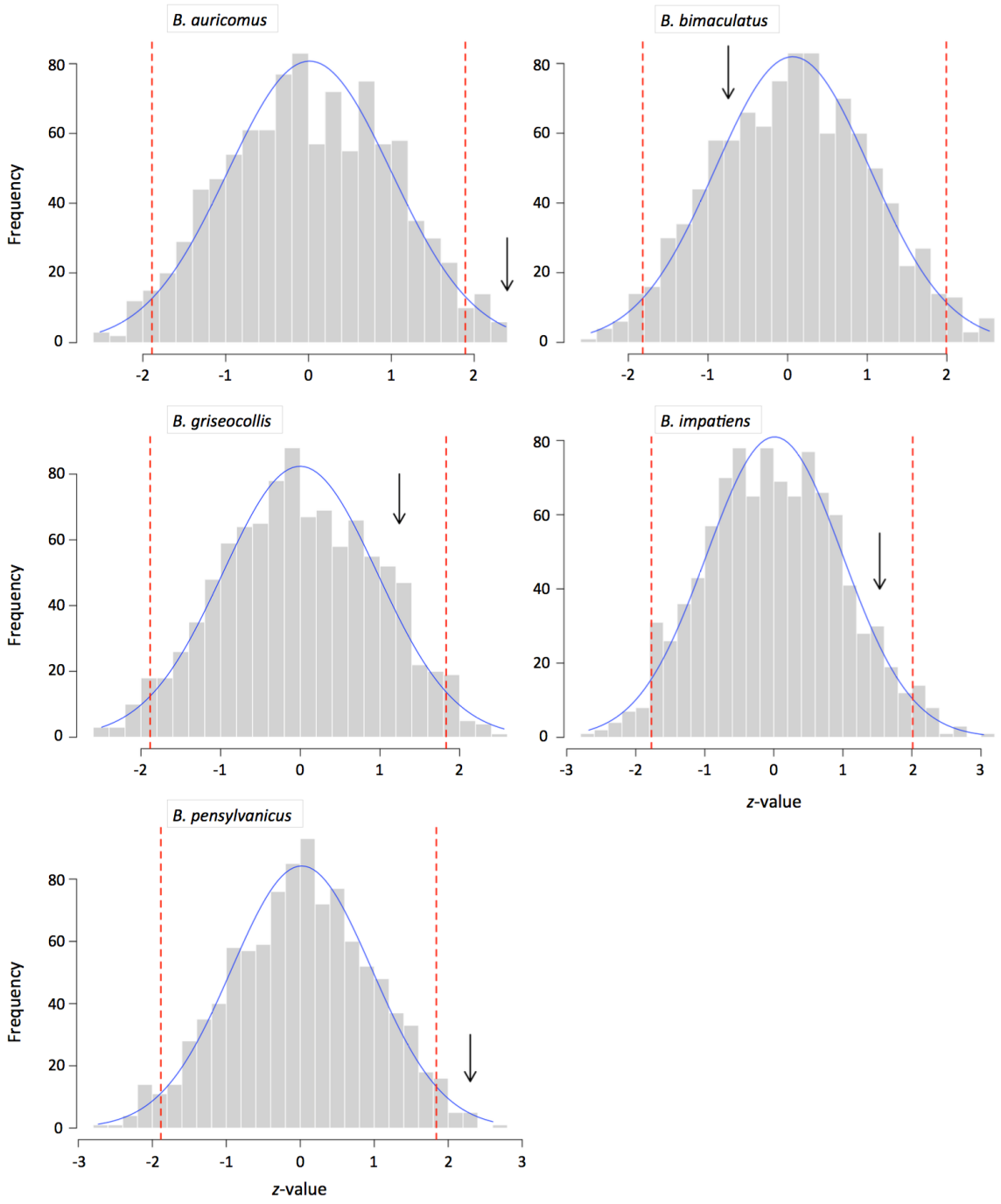

Figure S3. Histograms of *z*-values from 1,000 randomized logistic regressions per bumble bee (*Bombus* spp.) species. *z*-values are binned in gray bars, with blue lines representing the normal curve of the data. Dotted red lines indicate the boundary of each distribution's 2.5% tails, assuming two-tailed distributions (i.e. collectively 5% per plot).

Arrows represent where the *z*-value from each species' observed logistic regression falls within the distribution of randomized *z*-values.

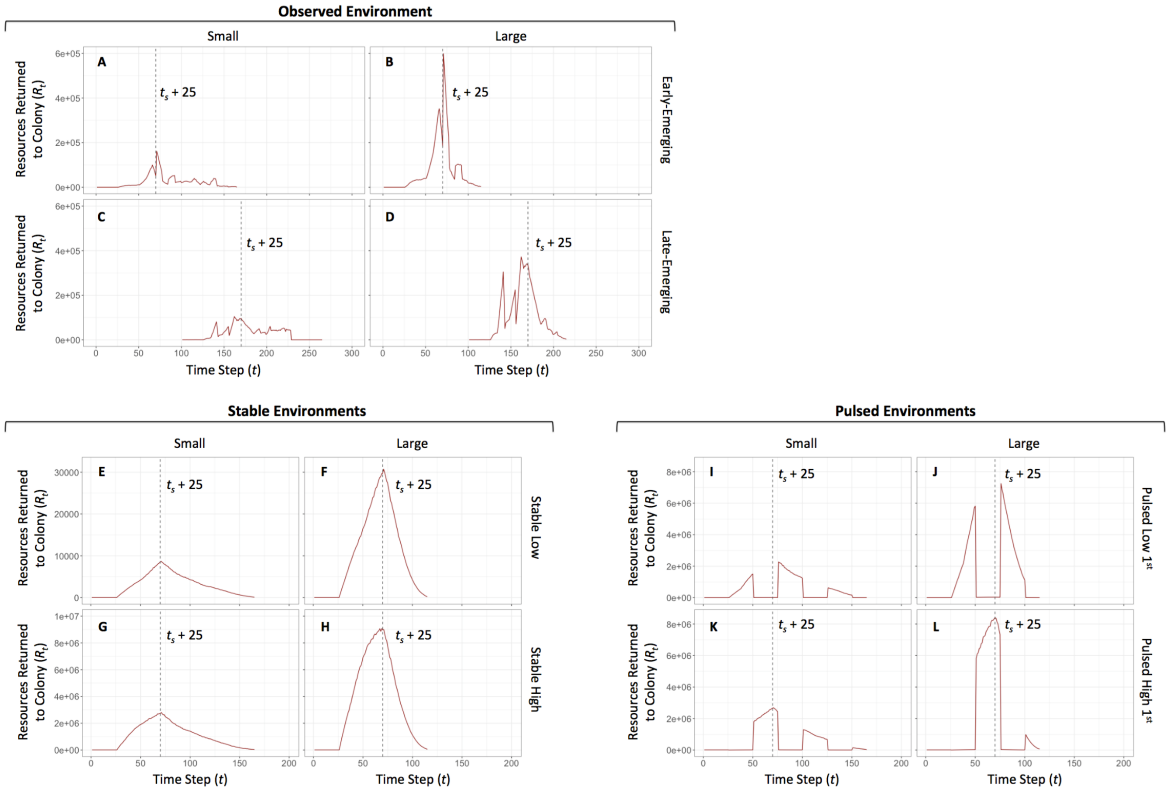

Figure S4. Resources returned to colony (R_t) across time steps (t) for each resource environment: (A-D) observed environment, (E-H) stable environments, and (I-L) pulsed environments. Vertical dashed lines give the time point when the last worker cohort ecloses; i.e. the switching point (t_s) plus the period of worker development.

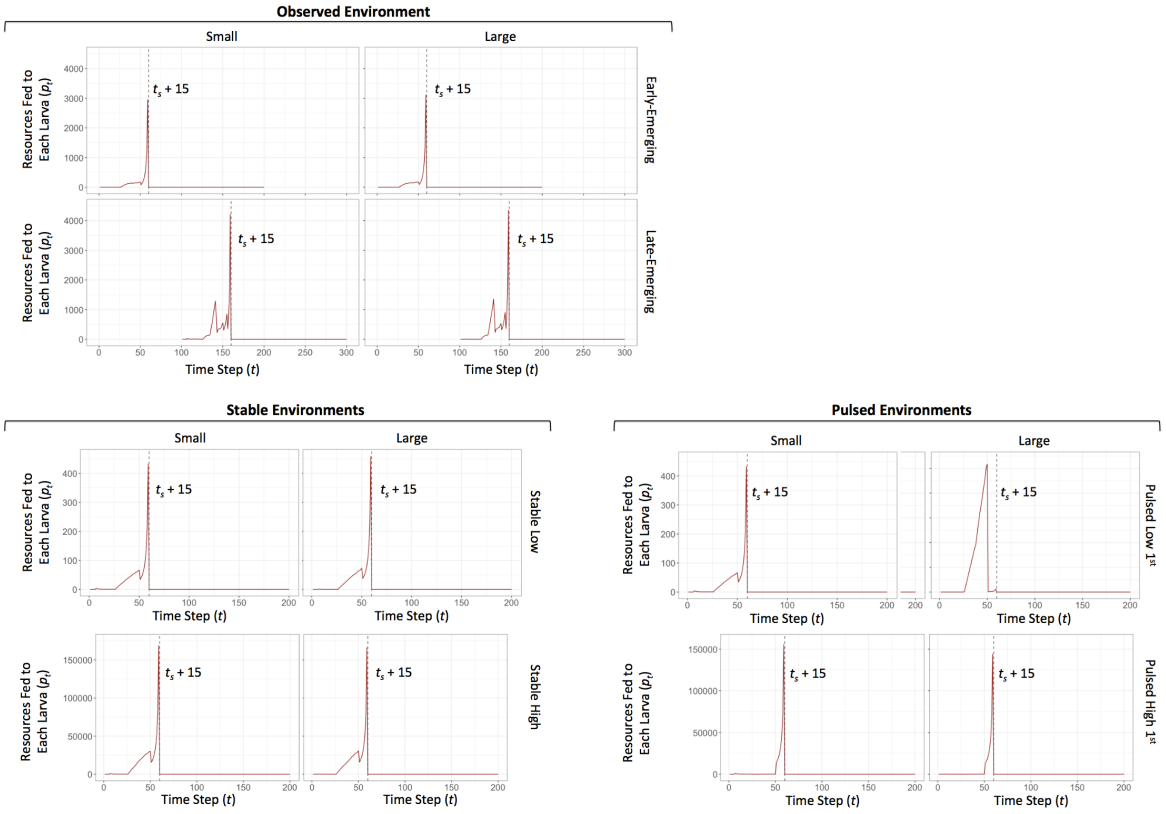

Figure S5. Resources fed to each larva (p_t) across time steps (t) for each resource environment: (A-D) observed environment, (E-H) stable environments, and (I-L) pulsed environments. Vertical dashed lines give the time point when the last larvae in the colony are fed; i.e. the switching point (t_s) plus the period of egg and larval development.

Chapter IV: Choice in a floral marketplace: the role of complexity in bumble bee decision-making

Matthew W. Austin*, Patricia Horack, and Aimee S. Dunlap

Department of Biology, University of Missouri – St. Louis, St. Louis, MO, USA

* Corresponding author

AustinMattW@gmail.com

Published in the *Behavioral Ecology*:

Austin, M.W., Horack P. & Dunlap, A.S. 2019. Choice in a floral marketplace: the role of complexity in bumble bee decision-making. Behavioral Ecology 30: 500-508. DOI: 10.1093/beheco/ary190

Abstract

Animals have evolved in complex, heterogeneous environments. Thus, decision-making behavior is likely affected by a diversity of co-occurring community-level traits. Here, we investigate how three co-occurring traits of floral communities - the number of flower types, reliability that flowers are associated with a reward, and signal complexity of flowers - affect bumble bee (*Bombus impatiens*) decision-making. We used arrays of artificial flowers in a full factorial experimental design to assess floral selectivity (preference and constancy), foraging efficiency, and decision latency in foraging bumble bees. We find that our environmental traits uniquely affect each of these behavioral variables, revealing the intricate, yet biologically significant ways that co-occurring environmental traits can affect behavior. Floral selectivity, but not foraging efficiency, is increased by a greater number of choices. Decision latency is greatest when bees are inexperienced foraging in environments with high choice number. Collectively taken, we argue that these results suggest a cost to deciding among many choices, which promotes choice fidelity when many options are present. We suggest that these results have implications for theory on decision-making and selection in biological markets, while demonstrating the importance of studying interactions between naturally co-occurring traits.

Keywords*:* Bombus*, constancy, decision-making, floral selectivity, foraging theory, rationality theory*

Introduction

Animals live in complex environments. Throughout their lives, they must attend to many aspects of environmental variation, which can affect relative fitness. However, most studies and models of animal behavior overlook this complexity in favor of isolating variables of interest (Fawcett et al., 2014). While this approach of testing one independent variable at a time is robust for controlling against compromising effects of extraneous variables, it is also important to explore interactions between naturally co-occurring traits that may have significantly impacted species throughout their evolution (Fawcett et al., 2014). Such trait interactions likely have a significant effect on animal decision-making. In nature, animals are often confronted with choice environments that exhibit heterogeneity in time and space, with additional variance added through changes in perception. Furthermore, decision-making is fundamentally intertwined with learning (Dukas and Ratcliffe, 2009), and significant interactions between ecological traits and individual experience likely affect decision-making in a myriad of ways. A necessary step toward understanding how real-world complexity affects animal decision-making is studying how naturally co-occurring traits interact with each other and with individual experience to affect behavior.

Biological markets are useful systems for the experimental study of these interactions. Defined as biological systems comprised of two trader classes that exchange mutually beneficial commodities (Noë and Hammerstein, 1995), biological markets naturally contain a diversity of co-occurring traits that may interact to affect consumer decision-making. Biological market theory is well described theoretically due to its analogy with human economic markets and is empirically tractable due to the model use

of various biological markets in the field of animal behavior (Noë and Hammerstein, 1995). Pollination systems are a classic example of biological markets, where forager decision-making results in the trading of pollination services to angiosperms in exchange for nutritional rewards of nectar and pollen. While in a floral marketplace, foragers are confronted with various aspects of heterogeneity that may influence what choices they make. Floral composition within and between habitats may change due to heterospecific differences in angiosperm phenology and the ephemeral nature of most flowers. Nectar availability may also exhibit variation between co-occurring flowers between years, over the course of a season, daily, and even hourly (Pleasants and Zimmerman, 1979; Pleasants, 1981; Real and Rathcke, 1988). Forager decision-making is likely affected by interactions between a variety of ecological traits that are relevant to optimizing energetic gain.

One ecological trait that should be considered in any decision-making scenario is the framing of the choice, i.e. the number of choices offered and the way they are presented. Choice framing is an important aspect in considerations of rational choice behavior, for instance a rational forager should not alter their preference between two items when the context changes - such as when an irrelevant alternative is added to a choice set (Fawcett et al., 2014). However, countless studies have documented violations of rationality theory in both human economic decision scenarios and ecological decision scenarios (e.g. Huber et al., 1982; Bateson et al., 2002; Shafir et al., 2002; Latty and Beekman, 2011). Collectively, these studies indicate that deciding between two options is fundamentally different from deciding between more than two options, due to costs associated with deciding among a high number of choices. For example, within human

economics, numerous studies have found that decision-making is impaired when multiple choices are available to choose from - an effect termed the "paradox of choice" (Iyengar and Lepper, 2000; Schwartz, 2004; Kinjo and Ebina, 2015). But while these choice behaviors may appear irrational in simplified experimental settings, they are ecologically rational in the wild (Stephens et al., 2004; Fawcett et al., 2014). Despite the importance that choice framing has on decision-making, it is not well studied how choice framing affects behavior in environments that mimic real-world complexity, as opposed to simplified experimental settings. If there are costs associated with deciding among a high number of choices in environments mimicking real-world complexity, we predict that foragers will attempt to avoid these costs by being selective on one of the available options.

Another ecological trait that is ubiquitous across decision-making scenarios is the reliability that a given cue is associated with a reward. Reliability of stimuli plays a large role across behavior, from animal communication (e.g. Bradbury and Vehrencamp, 2011) to the evolution of learning (reviewed in Dunlap and Stephens, 2016 and Dunlap et al., 2018). Within the function of learning, operantly conditioned behaviors can be strengthened by intermittent reward, thus making a learned behavior less likely to disappear (e.g. Mackintosh, 1974). Reliability of reward also interacts with memory and forgetting (e.g. McNamara and Houston, 1987a; Dunlap et al., 2009; Dunlap and Stephens, 2012). Collectively taken, changing rewards can absolutely promote strong and persistent learning, while at the same time too much change can promote constancy of choice. However, most decision-making studies incorporate only perfect reliability or random reliability into their experimental design. Few studies incorporate moderate

levels of reliability, despite these moderate levels producing intriguing results, such as bees not learning about social information in moderately reliable environments (Dunlap et al., 2016) and bees tracking resources suboptimally, except when resource persistence and reward quality are high (Dunlap et al., 2017). In an environment with moderate reward reliability, we predict that individuals will exhibit less foraging selectivity than in an environment with perfect reliability.

In nectar foragers, such as bees, the stimuli that are being associated with rewards are floral cues. If a bee reliably encounters a given floral cue paired with reward over time, learning this association between cue and reward should increase the bee's foraging performance. A rich empirical history exists on how bees use floral signals to guide their foraging behavior (e.g. Chittka et al., 1999; Gumbert, 2000; Gegear & Laverty, 2005; Dunlap et al., 2017; Kulahci et al., 2008; Katzenberger et al., 2013; Chittka, 2017). Bees are known to use a variety of floral signals when locating rewarding flowers, such color (e.g. Spaethe et al., 2001; Morawetz et al., 2013), pattern (e.g. Giurfa et al., 1996; Horridge, 1996; Plowright et al., 2011), morphology (e.g. Stout et al., 1998; Dohzono et al., 2011; Krishna & Keasar, 2018), odor (Raguso, 2008), and electric fields (Clarke et al., 2013). There is a growing recognition within the field of animal communication that multiple signals may function together in a composite multimodal/multicomponent signal (i.e. signal varying along multiple trait parameters), which may increase saliency of the signal to the receiver (e.g. Hebets and Papaj, 2005; Leonard et al., 2012). Such multimodal signals are ubiquitous in nature and are likely selected upon as functional units (Hebets and Papaj, 2005). In pollination systems, multimodal floral signals are predicted to increase floral selectivity (Gegear and Laverty, 2005) and foraging

performance (e.g. Kulahci et al., 2008; Leonard and Papaj, 2011). Potential mechanistic explanations for such effects include multimodal floral signals increase the speed of floral detection and enhance a pollinator's ability to learn about and remember rewarding flowers (Chittka et al., 1999; Leonard et al., 2011a; Leonard et al., 2012), as well as act upon cognitive constraints. While an empirical history exists on how multicomponent floral signals affect pollinator foraging behavior, a gap in the literature exists on how signal complexity affects foraging behavior in non-simplified environments that exhibit naturally co-occurring traits.

Here, we examine how three ecologically relevant traits - choice number, reward reliability, and signal complexity - may interact to affect bumble bee foraging behavior. By making a controlled behavioral test more similar to real-world environments that exhibit co-occurring traits, we are able to identify potential ways in which these factors may interact to affect behavior in ways not captured by tests that isolate one environmental trait at a time. To accomplish this, we observe bumble bee foraging behavior in floral marketplaces that vary in choice number (two or four flower types), reward reliability (completely or moderately reliable), and the signal complexity of flowers (flowers differing in one or two traits) according to a full factorial design. To measure bumble bee foraging, we quantify two measures of floral selectively - preference (i.e. the flower type a bee visits most often) and constancy (i.e. how often a bee makes consecutive visits to the same flower type), - and two measures of foraging performance - foraging efficiency (i.e. energetic gain per unit time) and decision latency (i.e. the time elapsed between floral visits). Our measures of floral selectivity provide direct quantification of the choices made by bumble bees, while our measures of foraging

performance provide an indication of costs imposed by different treatments (i.e. higher costs are associated with lower foraging efficiency and a higher latency between decisions). We hypothesize that choice number, reward reliability, and signal complexity interact with each other and with individual experience to significantly affect bumble bee foraging behavior. By simultaneously testing the effects of these environmental traits on the foraging behavior of bumble bees, we take a necessary step toward understanding how co-occurring environmental traits interact to affect animal decision-making.

Materials and Methods

Bumble Bee Husbandry

We obtained commercial colonies of bumble bees (*Bombus impatiens,* Hymenoptera: Apidae; Cresson, 1863) from Koppert Biological Systems. Upon arrival, we transferred bumble bee colonies to individual nest boxes (43 cm x 23 cm x 10 cm; wood frame, mesh ventilation holes, Plexiglas lid) attached to a foraging arena (1.2 m x 0.3 m x 0.4 m; wood frame, mesh ventilation holes, Plexiglas lid) (Fig. S1) and illuminated with full-spectrum LED lights (CH Lighting T5 13-watt, 6500K) on a 12:12 h light-dark cycle, with light beginning at 8:00 AM. Outside of training and experimentation, we provided bumble bees with a 20% (weight/weight) sucrose solution (nectar equivalent) *ad libitum* from wick feeders within the foraging arena and administered pollen to the hive approximately three times per week.

Training

Prior to experimentation we trained worker bees to brown artificial 'training flowers' that were structurally similar to the artificial flowers in testing. We constructed artificial flowers by attaching a small plastic cup made from a mircotiter well (i.e. the nectary) surrounded with laminated paper (~4.5 cm diameter) to the top of a 9.5 cm tall metal stalk. In order to familiarize bees with the artificial flower design, we placed four training flowers randomly within the foraging arena, each containing a reward of 100 μl of 60% sucrose solution, and we allowed bees to freely move between the nest box and foraging arena. This concentration of sucrose was chosen given that its higher concentration relative to *ad libitum* feeding should have helped motivate bees to visit the training flowers (Cnaani et al., 2006). We refilled training flowers by pipette immediately upon depletion. After an individual worker bee made two consecutive trips between training flowers and the nest box, we tagged her on the thorax and deemed her ready for experimentation. Following training, we removed all sucrose and olfactory residue from training flowers with water and 70% ethanol.

Experimental Design

For experimentation, we placed 40 artificial flowers inside of a foraging arena (1.2 m x 0.3 m x 0.4 m), arranged in five rows of eight, spaced apart by a distance of 13.3 cm for columns and 10.0 cm for rows (Fig. S1). We randomly assigned each array a 'focal' flower type – either blue or purple flowers – with all other flower types present comprising 'non-focal' flower types.

We designed the experiment as a full factorial, with two factors each of choice number (2 or 4 choices), reward reliability (100% or 80% reliable), and signal complexity

(color alone versus color/pattern/shape). This design resulted in eight total arrays (floral array types shown in Fig. 1). We tested six bees in each array type ($N = 48$ total), from a total of nine colonies. To avoid pseudoreplication of bees from the same colony, we randomized the arrays across colonies, so that no bees from the same colony were tested in the same array. Choice number featured arrays containing two flower types (blue and purple) or four flower types (blue, purple, orange, and pink). In arrays with two flower types, the spatial arrangement of flowers consisted of an alternating, checkered pattern of each flower type. In arrays with four flower types, the spatial arrangement of flowers was randomized within certain parameters (e.g. no more than two of the same flower type placed next to each other). For reliability of reward we created levels of either a 100/0 reward ratio or an 80/20 reward ratio. In 100/0 reward ratios (100% reliable), all of the focal flowers offered a nectar reward of 8 μl of 60% sucrose solution, while none of the non-focal flowers offered a nectar reward. In 80/20 reward ratios (80% reliable), 80% of the focal flowers and 20% of the non-focal flowers offered a nectar reward of 8 μl of 60% sucrose solution. A volume of 8 μl of 60% sucrose was chosen to encourage bees to visit multiple flowers per foraging trip. Finally, signal complexity was either color alone (visually simple) or color, pattern, and shape (visually complex). Visually simple flowers varied from one another in one signal alone (color), while visually complex flowers exhibited variation in two signal types (color and pattern) (Fig. 1). Each color of the experimental flowers was mapped into a color hexagon, a vision model for bees which calculates perceptual differences between colors based on photoreceptor excitations, rooted in *Bombus impatiens* color vision, to determine perceived color contrasts (Chittka,

1992; Skorupski & Chittka, 2010). The spectral reflectance curves and hexagonal color space can be found in the supplemental materials (Figs. S2 and S3).

We exposed individual subjects to a floral array, recording the first 100 foraging choices per bee while allowing the experimental bee to freely forage and move between the foraging arena and nest box. Each foraging choice was defined by at least two legs touching the dorsal side of an experimental flower. We refilled rewarding flowers with 8 µl of 60% sucrose solution immediately after depletion, while the bee was distracted by feeding from an alternative flower. The location of flowers was kept consistent throughout all 100 choices per bee. After experimentation, we euthanized experimental bees below 0° C and removed all sucrose and olfactory residue from experimental flowers with water and 70% ethanol.

Behavioral Variables

A video camera (Sony HDR-CX330) placed above the foraging arena recorded all experimentation. Using video playback in QuickTime Player (version 10.4) we quantified each behavioral variable from these recordings.

Preference

To assess preference, we must control for the number of options a bee has, making comparisons possible between treatments with two and four flower types. Thus, we calculated preference via Jacobs' index (D) (Jacobs, 1974), where preference is a measure of the degree to which an individual bee is biased in their selection for the focal flower type. Accordingly, $D = (r - p) / (r + p - 2rp)$; r is the proportion of focal flowers

selected and p is the proportion of focal flowers available in the array. A value of +1 indicates complete preference for the focal flower type and a value of -1 indicates foraging solely from non-focal flower types (Gegear and Laverty, 2005).

Constancy

We calculated constancy according to Bateman's index (BI) (Bateman, 1951; Gegear and Laverty, 2005), which describes the tendency of foragers to move assortatively between flowers of the same type over what would be expected given a certain degree of preference. For arrays containing two flower types, BI = $((AD)^{1/2} - (BC)^{1/2}) / ((AD)^{1/2} + (BC)^{1/2})$; A is the total number of moves between flowers of color one, B is the total number of moves from flower color one to flower color two, C is the total number of moves from flower color two to flower color one, and D is the total number of moves between flowers of color two. For arrays containing four flower types, BI = $((AFKP)^{1/4} - (BCDEGHIJLMNO)^{1/12}) / ((AFKP)^{1/4} + (BCDEGHIJLMNO)^{1/12})$, where each letter, A, F, K, and P, all represent moves between similar flower types and the remaining letters all represent moves between different flower types. A value of +1 indicates complete constancy and a value of -1 indicates complete inconstancy (i.e. that bees never visited the same flower type two times in a row).

Foraging Efficiency

Foraging efficiency is a measure of energetic gain per unit time. We calculated foraging efficiency as the amount of sucrose solution consumed per unit time spent foraging (i.e. the amount of time a bee spent in the foraging arena). This calculation assumed that all 8

μl of sucrose solution were consumed from rewarding flowers whenever a bee extended her proboscis into a flower's nectary.

Decision Latency

Decision latency is a measure of how quickly bees made foraging choices. We calculated decision latency as the 'time landing on a flower' minus the 'time leaving the previous flower.' When a given landing was the first landing since the bee entered the arena, 'time leaving the previous flower' was replaced with 'time entering the arena.' Smaller decision latency values reflect quicker decision-making.

Statistical Analyses

We divided each subject's total 100 choices into four blocks of 25 consecutive choices and calculated each of the dependent variables for each choice block. As constancy is a measure of moves between flower types, and thus the total number of moves between flowers equals 99, given 100 choices, we calculated the first block of Bateman's index with one move less than all subsequent blocks (i.e. block one = first 24 choices). For each behavioral measure we performed a full factorial ANOVA with main effects of choice number, signal complexity, and reward reliability, with repeated measures on the four choice blocks of each bee. We performed post-hoc tests, Tukey's HSD and contrasts, to examine aspects of significant interactions. All statistical analyses were performed in Statistica 8.

Results

We present the results of each behavioral variable separately. Full ANOVA tables (tables S1-S4) and the proportions of visits to each flower type (Fig. S4) can be found in the supplemental materials.

Preference

Bees showed greater preference for the focal flower with either greater choice number, 100% reliability, or complex signals ($F_{1,40}$=33.297, $F_{1,40}$=5.096, $F_{1,40}$=7.365, respectively, all p<0.03). Furthermore, bees increased preference for the focal flower as they gained experience foraging in their experimental floral array ($F_{3,120}$=20.921, p<0.0001). No statistically significant interactions between the main effects of our environmental traits were found for preference. However, a significant interaction was found between choice number and individual experience (i.e. choice block) for preference ($F_{3,120}$=2.89, p<0.05). This interaction reveals that bees always showed greater preference in treatments with greater choice number, regardless of how experienced bees were foraging in their array (Fig. 2). A significant interaction was also found between signal complexity and choice block for preference ($F_{3,120}$=3.69, p<0.05). This interaction reveals that when bees were experienced foraging in their floral array (i.e. in the last two choice blocks), bees showed greater preference when signals were complex (Fig. 2).

Constancy

Bees showed greater constancy on the focal flower with either greater choice number, 100% reliability, or complex signals ($F_{1,40}$=70.984, $F_{1,40}$=7.544, $F_{1,40}$=9.189, respectively, all p<0.009). Furthermore, bees increased constancy on the focal flower as they gained

experience foraging in their experimental floral array ($F_{3,120}$=4.398, p=0.006). A statistically significant interaction between the main effects of our three environmental traits (i.e. choice number, reward reliability, and signal complexity) was found for constancy ($F_{1,40}$=4.378, p<0.05). This interaction reveals that bees always showed greater constancy in treatments with more flower types, except during treatments with fewer flower types, 100% reliability, and complex floral signals, during which constancy was similarly high (Fig. 3). Additionally, two two-way interactions for constancy are also statistically significant: choice number and signal complexity ($F_{1,40}$=5.251, p<0.05), and reward reliability and signal complexity ($F_{1,40}$=5.285, p<0.05). These interactions reveal a similar trend to the three-way interaction: constancy was greater in treatments with more flower types and that constancy was greater in treatments with 100% reliability and complex floral signals, respectively.

Foraging Efficiency

Foraging efficiency was not significantly affected by either choice number, reward reliability, or signal complexity (all p>0.1). However, bees always increased their foraging efficiency as they gained experience foraging in their experimental floral array ($F_{3,120}$=55.559, p<0.0001; Fig. 2). We find no statistically significant interactions between any of our environmental traits (i.e. choice number, reward reliability, or signal complexity) or individual experience (i.e. choice block) for foraging efficiency (all p>0.07).

Decision Latency

The time elapsed between bees' foraging choices was affected by a statistically significant interaction between our three environmental traits (i.e. choice number, reward reliability, and signal complexity) and individual experience (i.e. choice block) ($F_{3,120}$=3.967, p<0.01). This interaction reveals that bees took longest to make decisions when they were inexperienced (i.e. in the first choice block) in four choice environments (when flowers were either visually simple and 100% reliable or visually complex and 80% reliable) (Fig. 4). Furthermore, regardless of floral array, bees always visited flowers more quickly as they gained experience foraging ($F_{3,120}$=23.924, p<0.0001); i.e. bees decreased their decision latency as they gained foraging experience. Finally, we find a statistically significant interaction between signal complexity and choice block ($F_{3,120}$=2.812, p<0.05). This interaction reveals that this decrease in decision latency was greater between the first and second choice blocks for bees in treatments with simple flowers compared to bees in treatments with complex flowers.

Discussion

We found that signal complexity, reward reliability, and choice number all interacted with one another and with individual experience to affect bees' decision-making behavior, supporting our hypotheses. Each of our behavioral variables was uniquely affected by these environmental traits, revealing the intricate, yet biologically significant ways that co-occurring environmental traits can affect behavior. While the environmental traits tested in this study have a history of being singularly tested in the cognitive sciences, our study provides a novel take on how interactions between these traits affect behavior in ways not captured by tests that isolate only one environmental trait at a time.

Here, we disentangle our results by discussing them in the context of consumer behavior in biological markets.

The finding that a greater number of choices increased floral selectivity is a novel result in the context of pollinator decision-making. It seems clear from numerous studies that making a choice between two options is fundamentally different than making a choice among three or more options (e.g. Bateson et al., 2002; Shafir et al., 2002; Latty and Beekman, 2011). Such option-dependent shifts in behavior have been demonstrated in a wide array of taxonomically diverse species - e.g. mammals (Huber et al., 1982), birds (Bateson et al., 2002; Shafir et al., 2002), insects (Shafir et al., 2002), ameboids (Latty and Beekman, 2011). In our study, we found that four choices significantly increased bees' selectivity relative to two choices (Figs. 2 & 3). This effect was immediate for our measure of preference, with inexperienced bees exhibiting greater preference in four choice environments than in two choice environments (Fig. 2). Additionally, bees' constancy was always increased in four choice environments, irrespective of signal complexity or reward reliability, while constancy was differentially affected by signal complexity and reward reliability in two choice environments (Fig. 3). In other words, constancy was always high in four choice environments and low in two choice environments, except in two choice environments with complex signals and reliable rewards, in which constancy was just as high as in four choice environments.

What do these results reveal about decision processes in bumble bees? We argue that these results suggest a high cost of being inconstant in environments with more than two choices, outweighing the cost of being constant on a moderately reliable resource. Numerous studies have documented impaired decision-making when a high number of

choices are available to choose from (e.g. Iyengar and Lepper, 2000; Schwartz, 2004; Kinjo and Ebina, 2015). Iyengar and Lepper (2000) provide a classic example of this in human economic markets, where individuals in a supermarket encountered either an extensive display of many jam types or a limited display of fewer jam types. Individuals who encountered the extensive display purchased fewer jams than individuals who encountered the limited display. This finding, replicated in other human decision-making scenarios (e.g. Kinjo and Ebina, 2015), is contrary to the idea that 'more choice is better' (Schwartz, 2004). This type of decision-making is often quantified in terms of decreased purchasing or decreased performance on a task, however this always implies a cost to deciding among an extensive set of choices, whether it be through distractors or background noise. A bee searching for a given flower in an environment with more flower types will have a higher number of distractors and background noise against which the flower's signal must be detected.

Decision latency is often analyzed as an indication of cost in animal foraging studies (e.g. Chittka et al, 1999). Higher latency between choices indicates a greater cost to decision-making (Chittka et al., 2007). In the decision latency results, we describe a four-way interaction in which greater latency for choice is found for more inexperienced bees, as they make choices in their first block of trials, in the four choice treatments. This greater decision latency may reflect a cost to decision-making in high choice environments (i.e. search time for a given flower should be greater with a greater number of distractors), especially as experience interacts with both reward reliability and signal complexity (as well as choice number framing). Reliability and signal complexity can both function to reduce uncertainty and a classic prediction of speed-accuracy trade-offs

is that time until a decision should be increased under noisy conditions (Chittka et al., 2007). The decrease in decision latency that occurred after bees gained experience (i.e. choice blocks two through four) in these treatments might reflect that bees were able to reduce the cost associated with noise from high choice number by increasing their selectivity on the focal flower type.

We found that while bees gained experience in their environments, they increased their foraging efficiency regardless of treatment (Fig. 2). This is an intriguing result given that a greater absolute amount of nectar was always available in two choice treatments, compared to four choice treatments, due to two choice treatments offering a greater number of rewarding flowers. Therefore, a null expectation would be that foraging efficiency should be greater in two choice scenarios than in four choice scenarios. Given that floral selectivity was greater in treatments with a higher choice number, we interpret these results as supporting the hypothesis that bees can avoid costs associated with foraging in a high choice environment by being highly selective, even if the flower type they select is only moderately rewarding. In other words, bees may be able to avoid a lower foraging efficiency in environments with a greater number of choices by being highly selective on one flower type.

Our results suggest a cost to being inconstant in environments with more than two choices, however there are several alternative explanations for these results. First, the greater selectivity found in four choice scenario for visually simple flowers might result from an inability of bees to reliably, or quickly discriminate between the blue and purple colors used. In such a case bees would make random choices between simple blue and purple flowers, unless they can rely on a secondary cue such as spatial location, which

bumble bees are well able to do (e.g. Church & Plowright, 2006; Jin et al., 2014). Here one can predict that a spatial cue effect would be likely more pronounced in the four choice scenario where fewer flowers must be learned; and indeed we do observe more selectivity in this case for visually-simple flowers. These colors are well within the range of discriminability found in other studies (e.g. Leonard et al., 2011), but we do not explicitly test discriminability here. If discrimination is possible, there may not have been enough trials for learning to occur in the most difficult scenario of two choices and visually simple flowers without many spatial cues. Aspects of partial preferences may also be at play (e.g. McNamara and Houston, 1987b, Stephens, 1985), as well as bees following a simple matching law in some circumstances (e.g Hernnstein, 1970, Houston et al., 2007). Finally, bees may be making fast, but inaccurate choices in two choice scenarios because the costs of mistakes were lower than in four choice scenarios (i.e. there were more rewarding flowers in two choice scenarios), making this speed-accuracy tradeoff worthwhile (Chittka et al., 2003), especially with a more difficult discrimination (e.g. Ings and Chittka, 2008, Kulachi et al., 2008). Indeed, these alternative explanations are not mutually exclusive of each other and may each be functionally relevant in an ecological setting.

The fitness of flowering plants depends on the reliable transfer of conspecific pollen between flowers (Galen and Gregory, 1989; Chittka et al., 1999; Morales and Traveset, 2008). Accordingly, plants benefit from a high degree of floral selectivity by their pollinators, and it has been hypothesized that complex floral signals have evolved to ensure that pollinators remain constant to conspecific flowers (Chittka et al., 1999; Gegear and Laverty, 2005; Hebets and Papaj, 2005; Leonard et al., 2011b). Our finding

that a greater number of flower types increased floral selectivity leads us to similarly hypothesize that evolutionary pressure to promote pollinator selectivity may have selected for concurrent blooming periods of sympatric angiosperms. Indeed, it may be adaptively beneficial for a plant to bloom in the absence of sympatric interspecific blooms. However, our results suggest that more flower types increasing pollinator floral selectivity may be a mechanism for concurrent blooming periods among species. We found preference was significantly increased by a greater number of flower types even when bees had little to no experience foraging in that environment. If this behavior extends to natural environments, then even mostly naïve bees would transfer less interspecific pollen between flowers in environments with many concurrently blooming species than in environments with fewer concurrently blooming species. Testing this hypothesis through studies on comparative behavior and phylogenetics would help elucidate the evolutionary significance of these findings.

The evolution of decision-making has been fundamentally affected by environmental complexity. In this study, we assessed how some of the natural co-occurring environmental traits that pollinators experience in their floral environments affect their decision-making while foraging. Our finding that a higher number of choices increased floral selectivity, but not foraging efficiency, is novel to the best of our knowledge. Due to the ubiquity of environmental spatiotemporal heterogeneity, many species likely have encountered choice number, reliability, and signal diversity throughout their evolution. Accordingly, the combined effects of these traits likely affect decision-making in a variety of biological markets. Pollination systems are ideal for studying such complex environments: floral communities often exhibit variation in floral

diversity, phenology, and nectar availability, and pollinators must attend to all of this variation to optimize their foraging. Thus, pollination systems offer a rich interdisciplinary approach for studying how biological market dynamics are affected by a diversity of real-world environmental traits, and whose results likely extend to many other types of biological markets.

Funding

This work was not supported by any funding agencies.

Acknowledgements

We thank Kathryn Janssen, Lauren Szyhowski, and Maria Nguyen for assistance with data collection and bumble bee husbandry, Avery Russell and Carla Essenberg for their protocol on mapping hexagonal color space, and Nathan Muchhala, Patricia Parker, Dunlap lab members, and two anonymous reviewers for constructive feedback on this manuscript.

Data Accessibility Statement

Analyses reported in this article can be reproduced using the data provided by Austin et al. (2018).

References

Austin, M. A., Horack, P. and Dunlap, A. S. (2018). Data from: choice in a floral marketplace: the role of complexity in bumble bee decision-making. Dryad Digital Repository. https://doi.org/10.5061/dryad.0rh6sg5

Bateman, A. J. (1951). The taxonomic discrimination of bees. *Heredity*. 5, 271-278.

Bateson, M., Healy, S. D. and Hurly, T. A. (2002). Irrational choices in hummingbird foraging behaviour. *Animal Behaviour*. 63, 587-596.

Bradbury, J. W. and Vehrencamp, S. L. (2011). *Principles of Animal Communication*. Sunderland, MA: Sinauer Associates.

Chittka, L. (1992). The colour hexagon: a chromaticity diagram based on photoreceptor excitations as a generalized representation of colour opponency. *Journal of Comparative Physiology A*. 170, 533-543.

Chittka, L. (2017). Bee cognition. *Current Biology*. 27, R1037-R1059.

Chittka, L., Dyer, A. G., Bock, F. and Dornhaus, A. (2003). Bees trade of foraging speed for accuracy. *Nature*. 424, 388.

Chittka, L., Skorupski, P. and Raine, N. E. (2009). Speed-accuracy tradeoffs in animal decision making. *Trends in Ecology and Evolution*. 24, 400-407.

Chittka, L., Thomson, J. D. and Waser, N. M. (1999). Flower constancy, insect psychology, and plant evolution. *Naturwissenschaften*. 86, 361-377.

Church, D. L. and Plowright, C. M. S. (2006). Spatial encoding by bumblebees (*Bombus impatiens*) of a reward within an artificial flower array. *Animal Cognition*, 9, 131-140.

Clarke, D., Whitney, H., Sutton, G. and Robert, D. (2013). Detection and learning of floral electric fields by bumblebees. *Science*. 340, 66-69.

Cnaani, J., Thomson, J. D. and Papaj, D. R. (2006). Flower choice and learning in foraging bumblebees: effects of variation in nectar volume and concentration. *Ethology*. 112, 278-285.

Dohzono, I., Takami, Y. and Suzuki, K. (2011). Is bumblebee foraging efficieny mediated by morphological correspondence to flowers? *International Journal of Insect Science*. 3, 1-10.

Dukas, R. and Ratcliffe J. M. (2009). Introduction. In *Cognitive Ecology II* (ed. R. Dukas and J.M. Ratcliffe), pp. 1-4. Chicago: University of Chicago Press.

Dunlap, A. S. and Stephens, D. W. (2012). Tracking a changing environment: optimal sampling, adaptive memory and overnight effects. *Behavioural Processes*. 89, 86-94.

Dunlap, A. S. and Stephens, D. W. (2016). Reliability, uncertainty, and costs in the evolution of animal learning. *Current Opinion in Behavioral Sciences*. 12, 73-79.

Dunlap, A. S., Austin, M. W. and Figueiredo, A. (2018). Components of change and the evolution of learning in theory and experiment. *Animal Behaviour*. in-press

Dunlap, A. S., McLinn, C. M., MacCormick, H. A., Scott, M. E. and Kerr, B. (2009). Why some memories do not last a lifetime: dynamic long-term retrieval in changing environments. *Behavioral Ecology*. 20, 1096-1105.

Dunlap, A. S., Nielsen, M. E, Dornhaus, A. and Papaj, D. R. (2016). Foraging bumble bees weigh the reliability of personal and social information. *Current Biology*. 26, 1195-1199.

Dunlap, A. S., Papaj, D. R. and Dornhaus, A. (2017). Sampling and tracking a changing environment: persistence and reward in the foraging decisions of bumblebees. *Interface Focus*. 7, 20160149.

Dyer, A. G. and Chittka, L. (2004a). Biological significance of distinguishing between similar colours in spectrally variable illumination: bumblebees (*Bombus terrestris*) as a case study. *Journal of Comparatice Physiology A*. 190, 105-114.

Dyer, A. G. and Chittka, L. (2004b). Fine colour discrimination requires differential conditioning in bumblebees. *Naturwissenschaften.* 91, 224-227.

Dyer, A. G., Paulk, A. C. and Reser, D. H. (2011). Colour processing in complex environments: insights from the visual system of bees. *Proceedings of the Royal Society of London. Series B: Biological Sciences*. 278, 952-959.

Fawcett, T. M., Fallenstein, B., Higginson, A. D., Houston, A. I., Mallpress, D. E. W. and Trimmer, P. C. (2014). The evolution of decision rules in complex environments. *Trends in Cognitive Sciences*. 18, 153-161.

Galen, C. and Gregory, T. (1989). Interspecific pollen transfer as a mechanism of competition: consequences of foreign pollen contamination for seed set in the alpine wildflower, *Polemonium viscosum*. *Oecologia*. 81, 120-123.

Gegear, R. J. and Laverty, T. M. (2005). Flower constancy in bumblebees: a test of the trait variability hypothesis. *Animal Behaviour*. 69, 939-949.

Giurfa, M., Eichmann, B. and Menzel, R. (1996). Symmetry perception in an insect. *Nature*. 382, 458-461.

Gumbert, A. (2000). Color choices by bumble bees (*Bombus terrestris*): innate preferences and generalization after learning. *Behavioral Ecology and Sociobiology*. 48, 36-43.

Hebets, E. A. and Papaj, D. R. (2005). Complex signal function: developing a framework of testable hypotheses. *Behavioral Ecology and Sociobiology*. 57, 197-214.

Herrnstein, R. J. (1970). On the law of effect. *Journal of the Experimental Analysis of Behavior*. 13, 243–66.

Horridge, G. A. (1996). The honeybee (*Apis mellifera*) detects bilateral symmetry and discriminates its axis. *Journal of Insect Physiology*. 42, 755-764.

Houston, A. I., McNamara, J. M. and Steer, M. D. (2007). Do we expect natural selection to produce rational behaviour? *Philosophical Transactions of the Royal Society B*. 362, 1531-1543.

Huber, J., Payne, J. W. and Puto, C. (1982). Adding asymmetrically dominated alternatives: violations of regularity and the similarity hypothesis. *Journal of Consumer Research*. 9, 90-98.

Ings, T. C. and Chittka, L. (2008). Speed-accuracy tradeoffs and false alarms in bee responses to cryptic predators. *Current Biology*. 18, 1520-1524.

Iyengar, S. S. and Lepper, M. R. (2000). When choice is demotivating: can one desire too much of a good thing? *Journal of Personality and Social Psychology*. 79, 995-1006.

Jacobs, J. (1974). Quantitative measurement of food selection. *Oecologia*. 14, 413–417.

Jin, N., Landgraf, T., Klein, S. and Menzel, R. (2014). Walking bumblebees memorize panorama and local cues in a laboratory test of navigation. *Animal Behaviour*, 97, 13-23.

Katzenberger, T. D., Lunau, K. and Junker, R. R. (2013). Salience of multimodal flower cues manipulates initial responses and facilitates learning performance of bumblebees. *Behavioral Ecology and Sociobiology*. 67, 1587-1599.

Kinjo, K. and Ebina, T. (2015). Paradox of choice and consumer nonpurchase behavior. *AI & Society*. 30, 291-297.

Krishna, S. and Keasar, T. (2018). Morphological complexity as a floral signal: from perception by insect pollinators to co-evolutionary implications. *International Journal of Molecular Sciences*. 19, 1-15.

Kulahci, I. G., Dornhaus, A. and Papaj, D. R. (2008). Multimodal signals enhance decision making in foraging bumble-bees. *Proceedings of the Royal Society of London. Series B: Biological Sciences*. 275, 797-802.

Latty, T. and Beekman, M. (2011). Irrational decision-making in an amoeboid organism: transitivity and context-dependent preferences. *Proceedings of the Royal Society of London. Series B: Biological Sciences*. 278, 307-312.

Leonard, A. S. and Papaj, D. R. (2011). 'X' marks the spot: the possible benefits of nectar guides to bees and plants. *Functional Ecology*. 25, 1293-1301.

Leonard, A. S., Dornhaus, A. and Papaj, D. R. (2011a). Flowers help bees cope with uncertainty: signal detection and the function of floral complexity. *The Journal of Experimental Biology*. 214, 113-121.

Leonard, A. S., Dornhaus, A. and Papaj, D. R. (2011b). Forget-me-not: complex floral displays, inter-signal interactions, and pollinator cognition. *Current Zoology*. 57, 215-224.

Leonard, A. S., Dornhaus, A. and Papaj, D. R. (2012). Why are floral signals complex? An outline of functional hypotheses. In *Evolution of Plant-Pollinator Relationships* (ed. S. Patiny), pp. 261-282. Cambridge: Cambridge University Press.

Mackintosh, N. J. (1974). *The Psychology of Animal Learning*. New York, NY: Academic Press.

McNamara, J. M. and Houston, A. I. (1987a). Memory and the efficient use of information. *Journal of Theoretical Biology*. 125, 385-395.

McNamara, J. M. and Houston, A. I. (1987b). Partial preferences and foraging. *Animal Behaviour*. 35, 1085-1099.

Morales, C. L. and Traveset, A. (2008). Interspecific pollen transfer: magnitude, prevalence and consequences for plant fitness. *Critical Reviews in Plant Sciences*. 27, 221-238.

Morawetz, L., Svoboda, A., Spaethe, J. and Dyer, A. G. (2013). Blue colour preference in honeybees distracts visual attention for learning closed shapes. *Journal of Comparative Physiology*. 199, 817-827.

Noë, R. and Hammerstein, P. (1995). Biological markets. *Trends in Ecology & Evolution*. 10, 336-339.

Pleasants, J. M. (1981). Bumblebee response to variation in nectar availability. *Ecology*. 62, 1648-1661.

Pleasants, J. M. and Zimmerman, M. (1979). Patchiness in the dispersion of nectar resources: evidence for hot and cold spots. *Oecologia*. 41, 283-288.

Plowright, C. M. S., Evans, S. A., Chew Leung, J. and Collin, C. A. (2011). The preference for symmetry in flower-naïve and not-so-naïve bumblebees. *Learning and Motivation*. 42, 76-83.

Raguso, R. A. (2008). Wake up and smell the roses: the ecology and evolution of floral scent. *Annual Review of Ecology, Evolution, and Systematics*. 39, 549-569.

Real, L. and Rathcke, B. J. (1988). Patterns of individual variaibility in floral resources. *Ecology*. 69, 728-735.

Schwartz, B. (2004). *The Paradox of Choice: Why More is Less*. New York, NY: HarperCollins Publishers Inc.

Shafir, S., Waite, T. A. and Smith, B. H. (2002). Context-dependent violations of rational choice in honeybees (*Apis mellifera*) and gray jays (*Perisoreus Canadensis*). *Behavioral Ecology and Sociobiology*. 51, 180-187.

Skorupski, R. and Chittka L. (2010). Photoreceptor spectral sensitivity in the bumblebee, *Bombus impatiens* (Hymenoptera: Apidae). *PLoS ONE*. 5, e12049.

Spaethe, J., Tautz, J. and Chittka, L. (2001). Visual constraints in foraging bumblebees: flower size and color affect search time and flight behavior. *Proceedings of the National Academy of Sciences of the United States of America*. 98, 3898-3903.

Stephens, D. W. (1985). How important are partial preferences? *Animal Behaviour*. 33, 667-669.

Stephens, D. W., Kerr, B. and Fernández-Juricic, E. (2004). Impulsiveness without discounting: the ecological rationality hypothesis. *Proceedings of the Royal Society of London. Series B: Biological Sciences*. 271, 2459-2465.

Stout, J. C., Allen, J. A. and Goulson, D. (1998). The influence of relative plant density and floral morphological complexity on the behaviour of bumblebees. *Oecologia*. 117, 543-550.

van der Kooi, C. J., Dyer, A. G., Kevan, P. G. and Lunau, K. (2018). Functional signicance of the optical properties of fowers for visual signalling. *Annals of Botany*. in-press

Figures

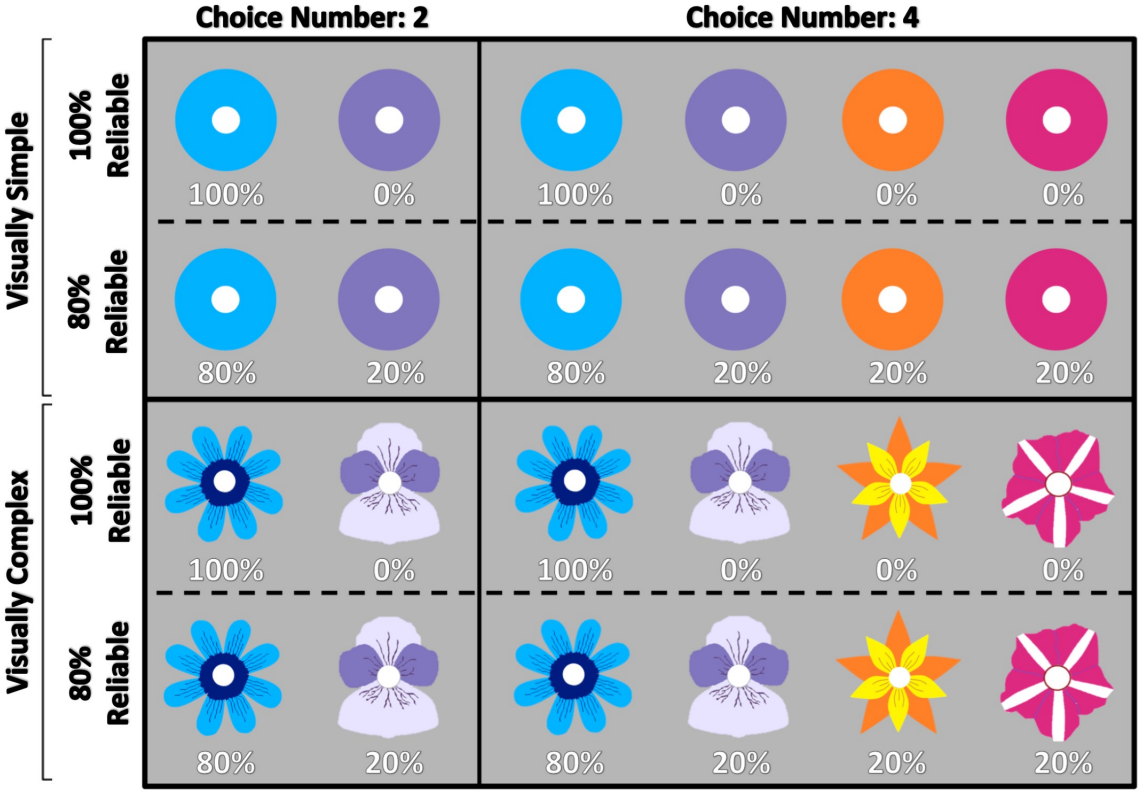

Figure 1. Breakdown of treatments based on a full factorial experimental design. Each unique combination of choice number, signal complexity, and reward reliability was used as an array, resulting in a total of eight array types (n = 6 per array type). Flowers are shown dorsally, with white circles indicating the location of nectar reward. Percentages indicate how many flowers of each flower type were paired with a nectar reward in each array. The percentages shown in this figure are for arrays in which the blue flower type was the focal flower.

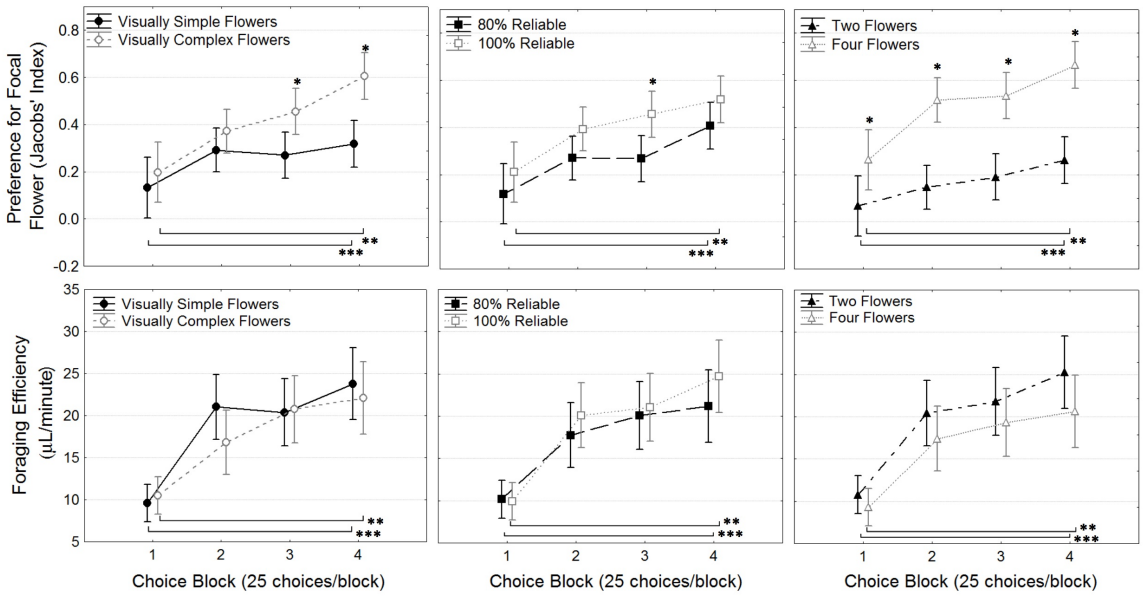

Figure 2. Interactions between each of our main effects with individual experience (i.e. choice blocks) for each preference (Jacobs' index) and foraging efficiency. Statistically significant differences (p<0.05) are denoted by * between treatments within a single block, ** between the first and fourth decision blocks for complex flowers, 100% reliability, or four flower types, and *** between the first and fourth decision blocks for simple flowers, 80% reliability, or two flower types. Significant differences were determined using contrasts for least squares means. Error bars are 95% CIs. n = 24 per treatment.

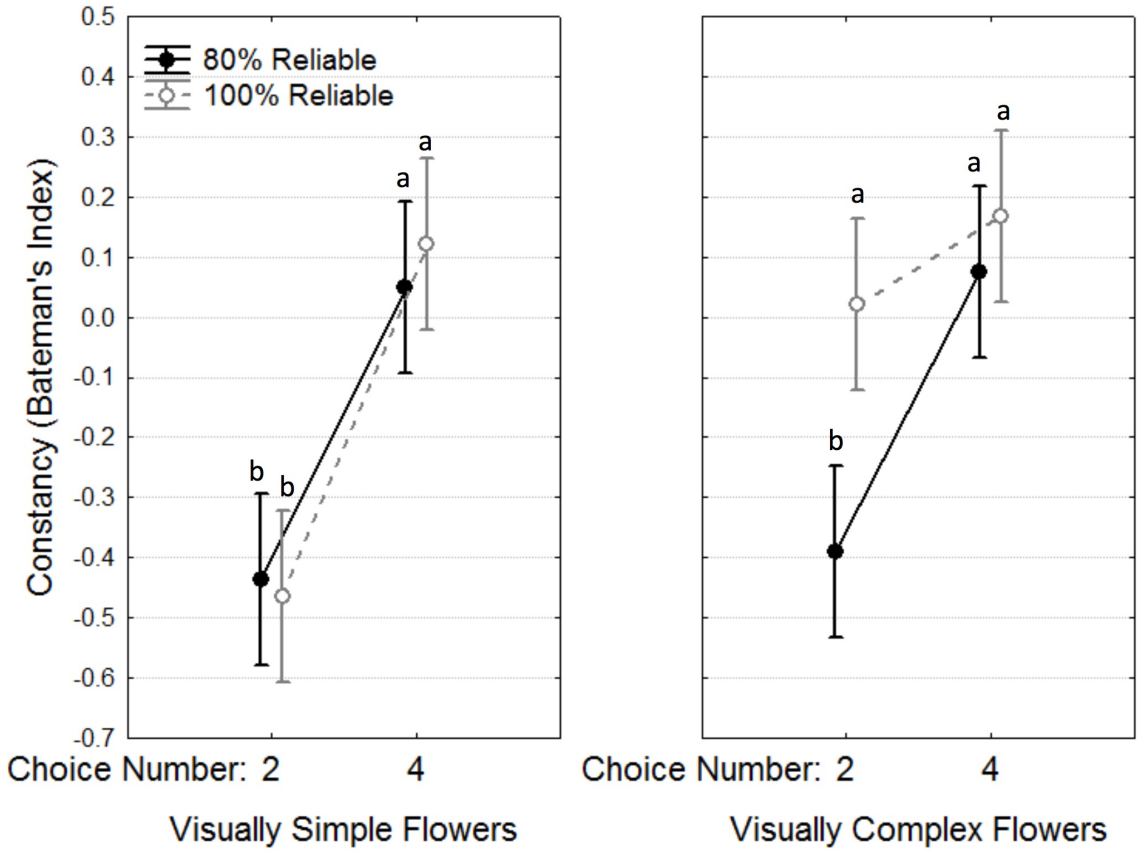

Figure 3. Three-way interaction of signal complexity, reward reliability, and choice number on constancy (Bateman's index). Significant differences (p<0.05), as determined by Tukey's HSD, exist between points labeled with different letters. Error bars are 95% CIs. n = 6 per treatment.

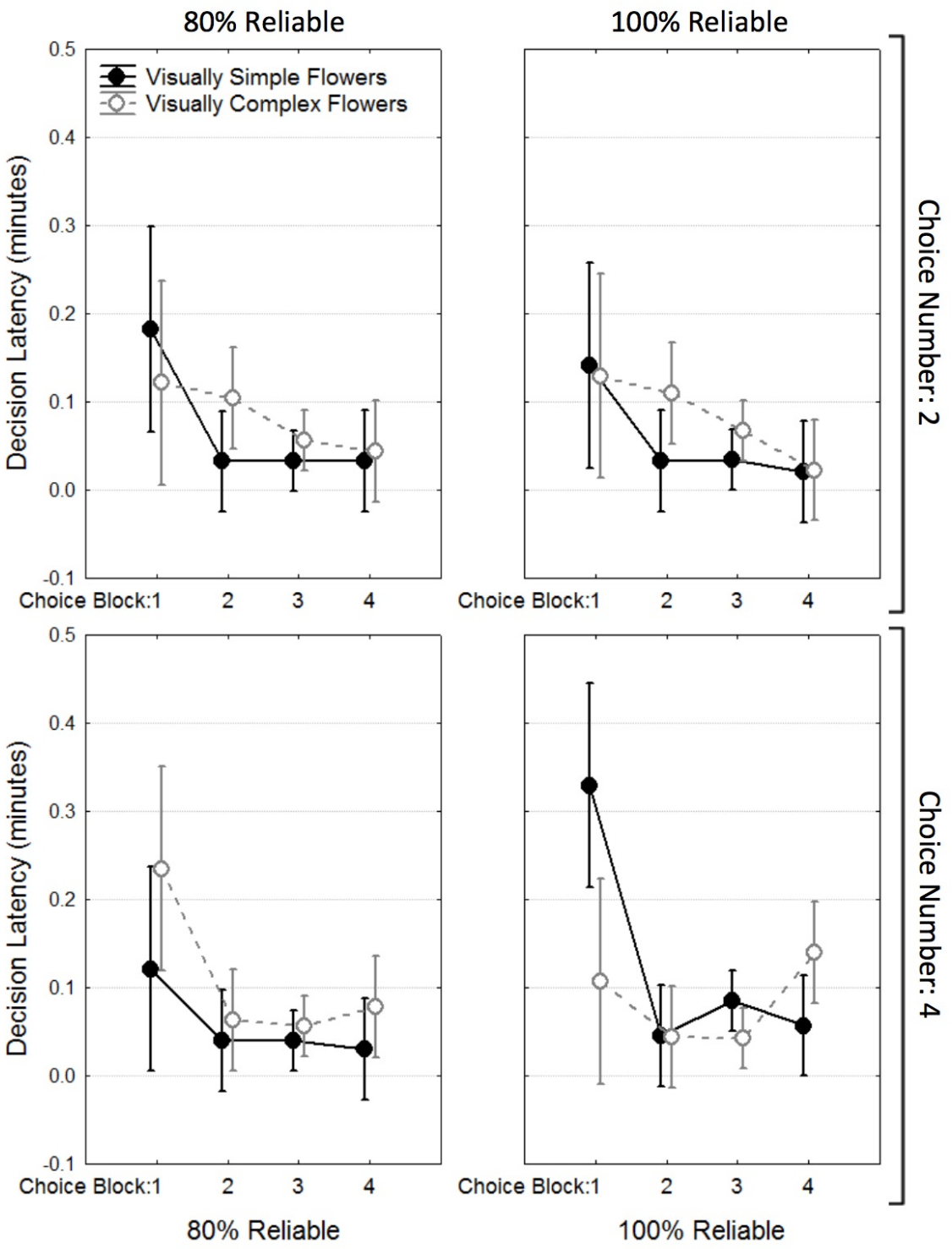

Figure 4. Four-way interaction between the main effects of signal complexity, reward reliability, choice number, and individual experience (i.e. choice block) on decision latency. The most visible differences in this interaction can be seen by two treatments:

the first choice blocks in arrays of four choices, 100% reliability, and visually simple flowers and arrays of four choices, 80% reliability, and visually complex flowers. Neither of these points is significantly different from one another. Additionally, the former is significantly different from every point except (i) the first choice block in arrays with two choices and visually simple flowers, and (ii) the fourth choice block in arrays with four choices, 100% reliability, and visually complex flowers. The latter is only significantly different from choice blocks two, three, and four in (i) arrays with 80% reliability and visually simple flowers and (ii) arrays with two choices, 100% reliability, and visually simple flowers. This point is also significantly different from (iii) the fourth choice block in arrays with two choices, 100% reliability, and visually complex flowers. Error bars are 95% CIs. $n = 6$ per treatment.

Supplemental Materials

Figures

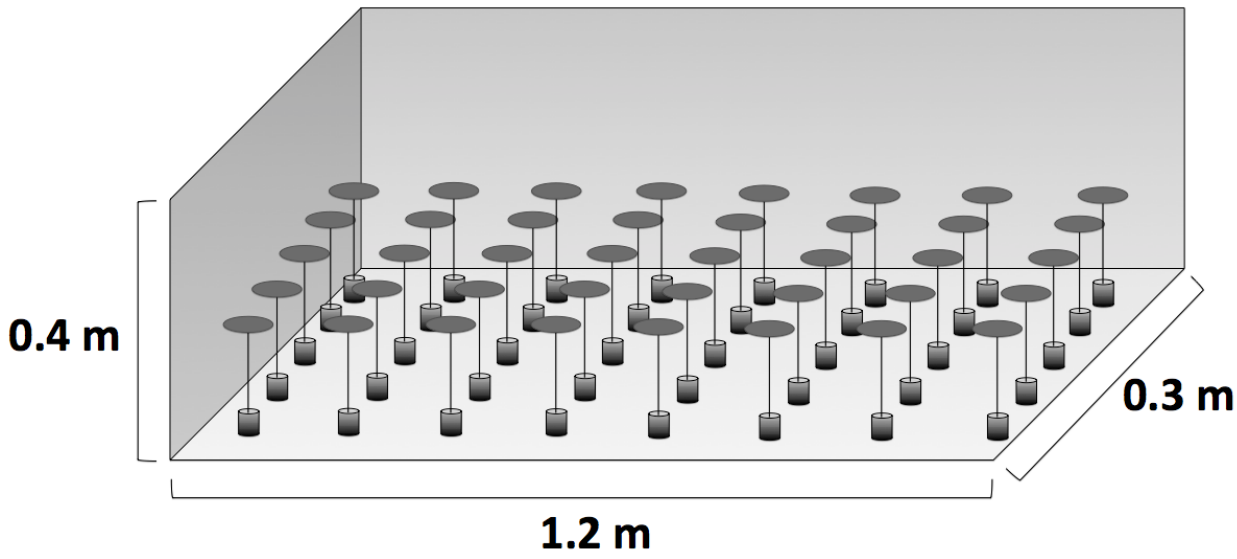

Figure S1. Three-dimensional depiction of foraging arena containing the five-row by eight-column arrangement of artificial flowers.

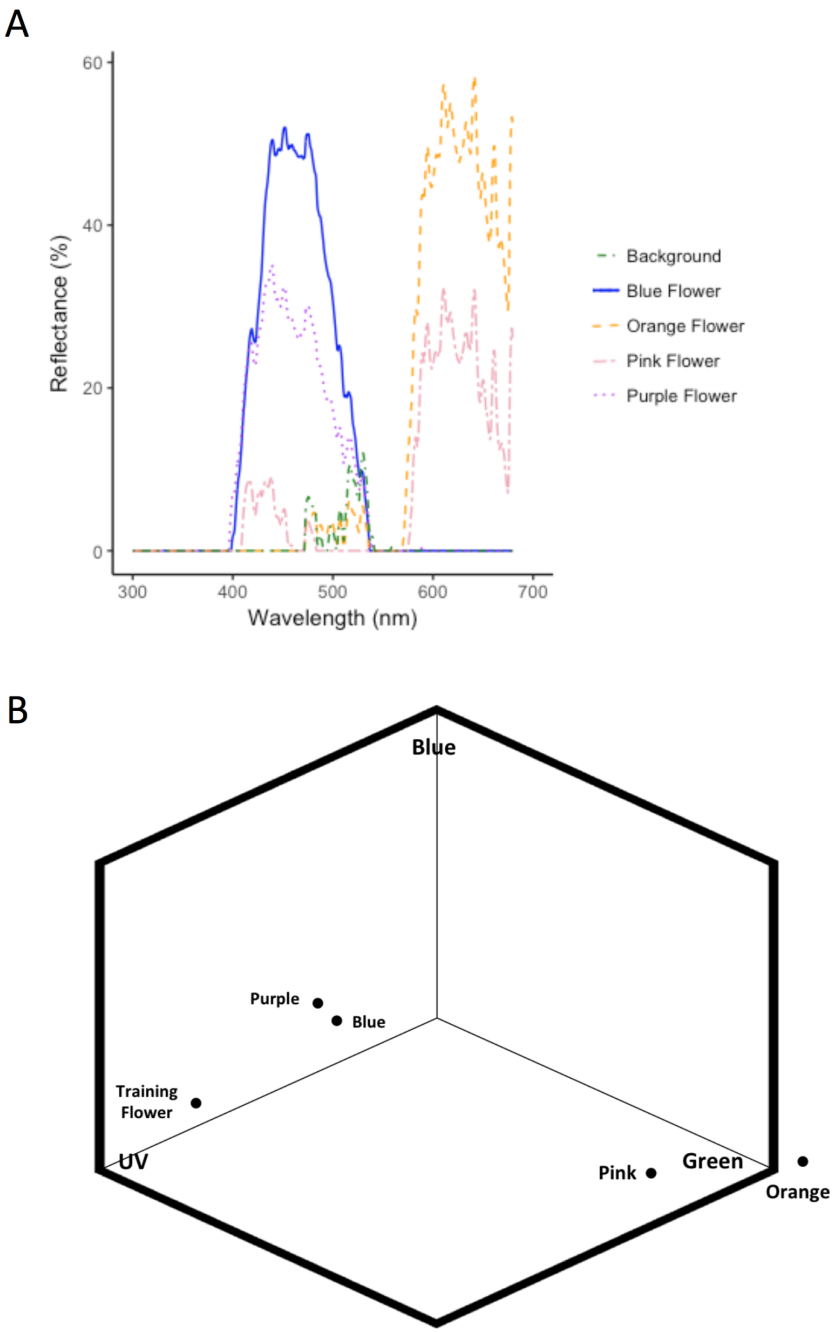

Figure S2. Colors of visually simple flowers used in experiment. (A) Spectral reflectance curves of simple flowers and background. Measurements taken with an Ocean Optics fiber optic spectrometer. (B) Colors of simple flowers and training flower depicted in hexagonal color space for *Bombus impatiens* (Chittka, 1992; Skorupski & Chittka, 2010).

The hexagon is rooted in the green background against which the flowers were presented to bees.

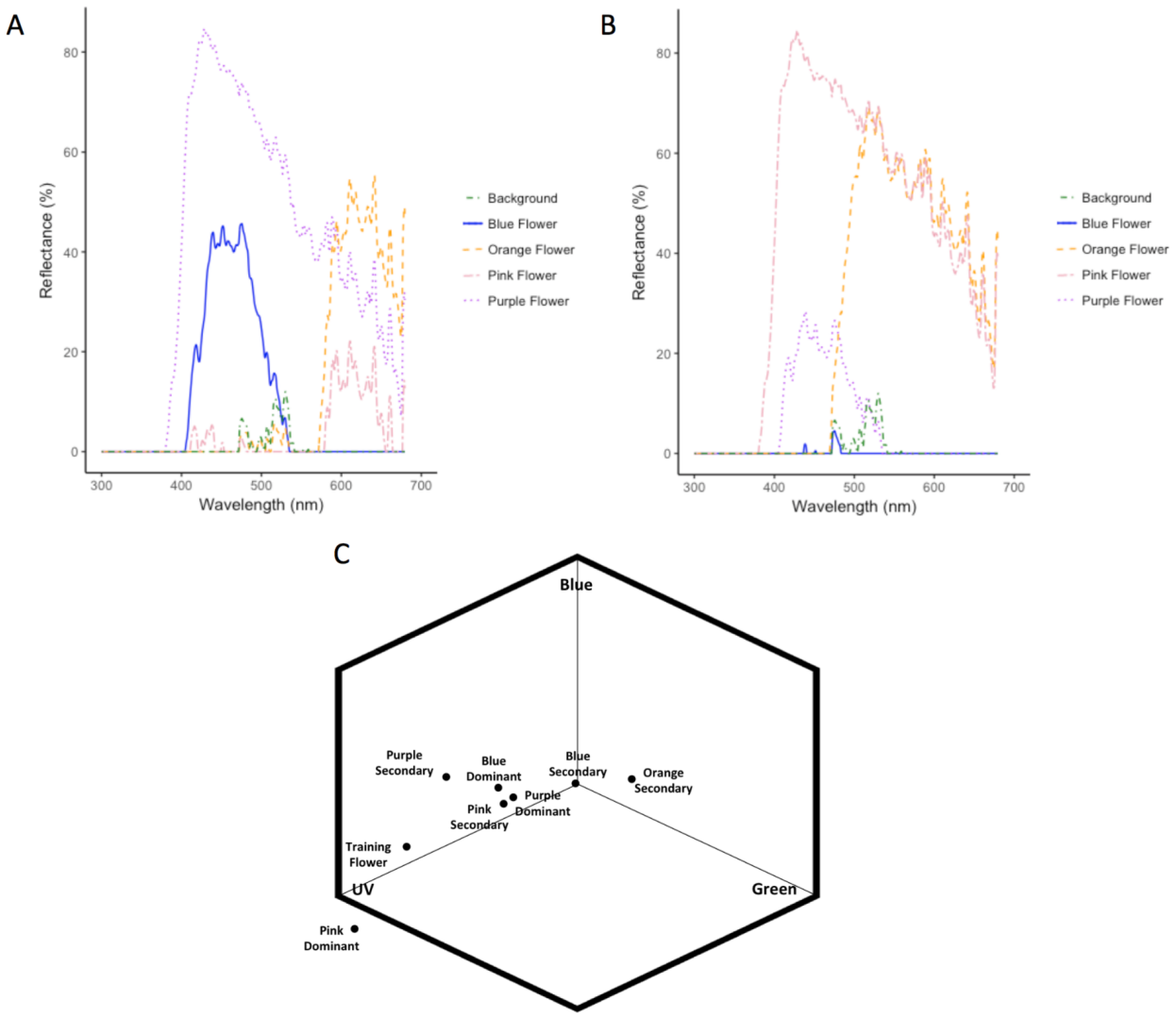

Figure S3. Colors of visually complex flowers used in experiment. (A) Spectral reflectance curves of the dominant color of visually complex flowers and background. (B) Spectral reflectance curves of the secondary color of visually complex flowers and background. Measurements taken with an Ocean Optics fiber optic spectrometer. (C) Colors of visually complex flowers (both dominant and secondary colors) and training flower depicted in hexagonal color space for *Bombus impatiens* (Chittka, 1992; Skorupski & Chittka, 2010). The hexagon is rooted in the green background against

which the flowers were presented to bees. The dominant color of orange visually complex flowers was plotted outside of this area.

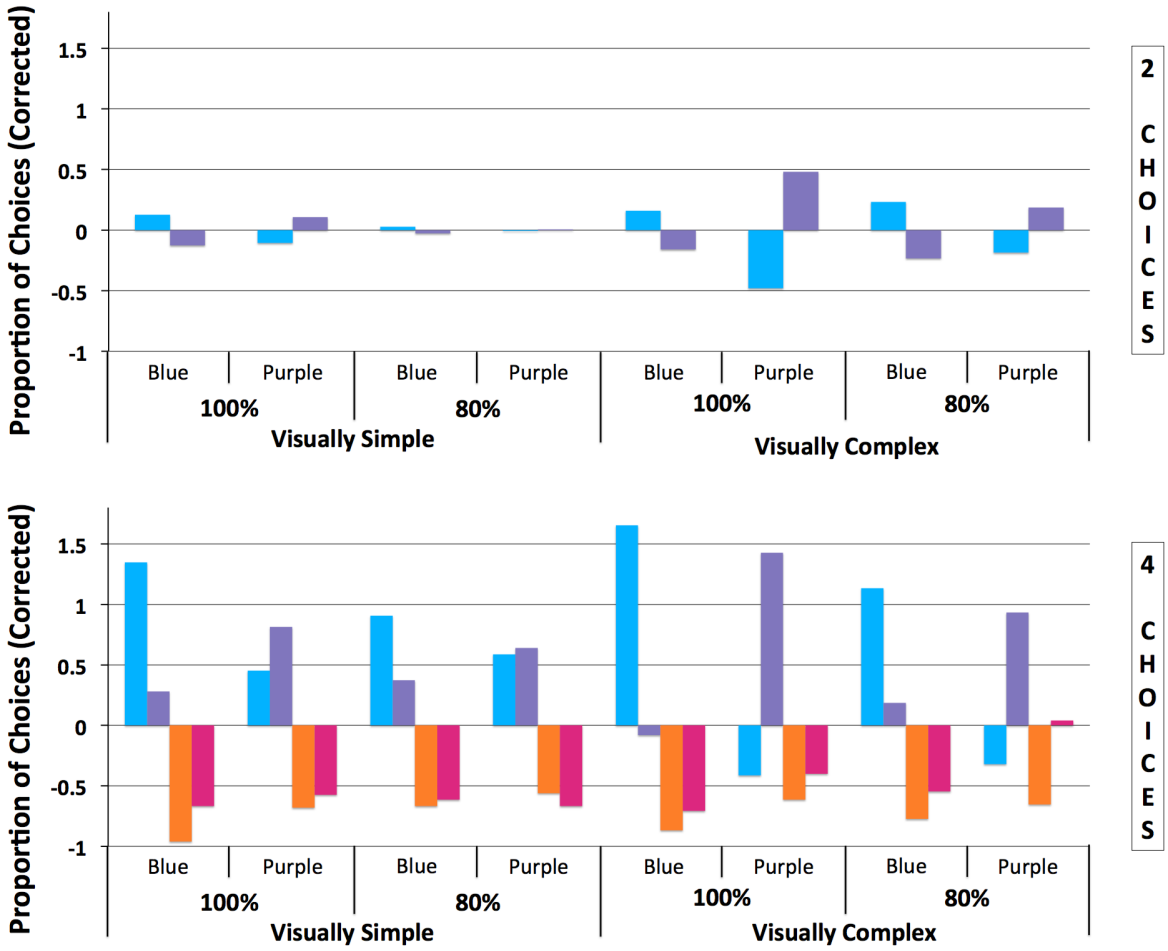

Figure S4. Average proportion of choices to each flower type according to each unique floral array. To facilitate comparison between two and four choice arrays, proportions have been corrected based on null expectations for random foraging; i.e. corrected proportion of choices = (observed - expected)/expected. Bar colors correspond to flower colors. 'Blue' and 'Purple' in the upper x-axis indicate color of focal flower. n = 3 per array.

Foraging Accuracy

To address the potential that bees could not reliably discriminate between the simple blue and purple flowers, we have included the following analysis of foraging accuracy. Foraging accuracy is a measure of how accurately bees foraged within their environment. For a bee making perfect choices, foraging accuracy = r in 100% reliable environments; r is the proportion of focal flowers selected. We use foraging accuracy to consider separately when analyzing discrimination as it provides a direct measure of how often the rewarding flower was chosen.

For foraging accuracy, we performed a repeated measures ANOVA of blocks of 25 choices for each bee and factors of the experimental design. To begin answering the question of whether bees can discriminate between the simple blue and purple flowers, we can take the strongest situation for learning from this analysis, where reliability of reward is 100%, and flowers are visually simple. If bees are learning the pairing of flower and reward, they should increase in accuracy over time. Using contrasts within the ANOVA analysis, we tested the difference between the accuracy of the first block and the fourth block for both choice number treatments. With four choices, bees show evidence of learning the flower-reward pairing (p=0.018), however there is not a significant increase in accuracy when bees only have two choices (p=0.258). A figure depicting this analysis is given below.

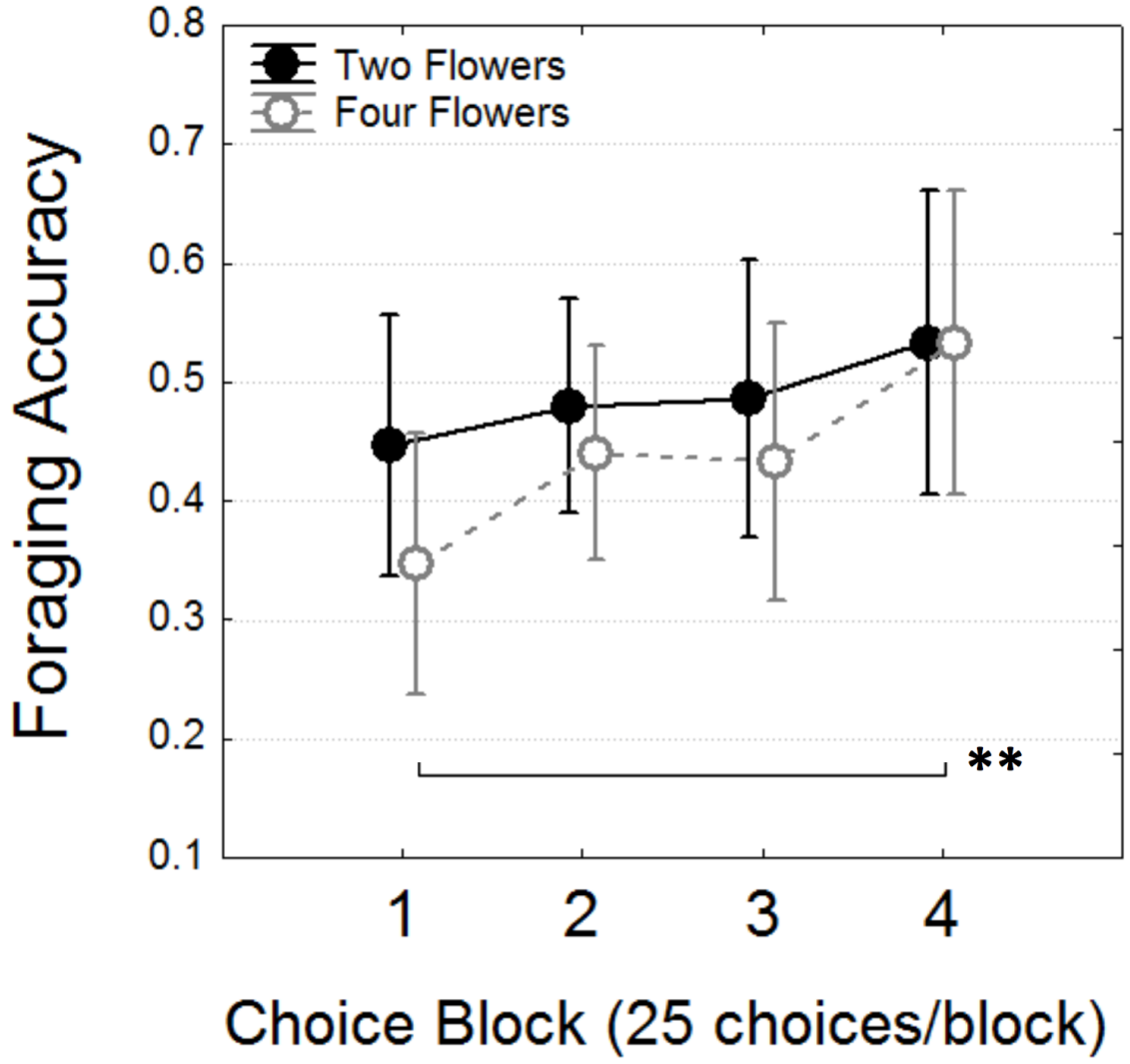

Figure S5. Interaction between our main effect of choice number with individual experience (i.e. choice block) for foraging accuracy, when rewards are 100% reliable and flowers are visually simple. A statistically significant difference (p<0.05) between the first and fourth decision blocks for four choices is denoted by **. Significant differences were determined using contrasts. Error bars are 95% CIs. n = 6 per treatment.

ANOVA Tables

Table S1. Bateman's Index ANOVA table.

Effect	SS	DF	MS	F	p
Intercept	2.210232	1	2.210232	18.48029	0.000107
Signal Complexity	1.098907	1	1.098907	9.18823	0.004259
Reward Reliability	0.902317	1	0.902317	7.54449	0.008980
Choice Number	8.489642	1	8.489642	70.98397	0.000000
Signal Complexity x Reward Reliability	0.632041	1	0.632041	5.28465	0.026817
Signal Complexity x Choice Number	0.628035	1	0.628035	5.25116	0.027275
Reward Reliability x Choice Number	0.143344	1	0.143344	1.19853	0.280164
Signal Complexity x Reward Reliability x Choice Number	0.523646	1	0.523646	4.37833	0.042789
Error	4.783977	40	0.119599		
Choice Block	0.470736	3	0.156912	4.39759	0.005670
Choice Block x Signal Complexity	0.026201	3	0.008734	0.24477	0.864918
Choice Block x Reward Reliability	0.062114	3	0.020705	0.58026	0.629081
Choice Block x Choice Number	0.134806	3	0.044935	1.25935	0.291513
Choice Block x Signal Complexity x Reward Reliability	0.151197	3	0.050399	1.41247	0.242515
Choice Block x Signal Complexity x Choice Number	0.178574	3	0.059525	1.66822	0.177492
Choice Block x Reward Reliability x Choice Number	0.056381	3	0.018794	0.52671	0.664770
Choice Block x Signal Complexity x Reward Reliability x Choice Number	0.241902	3	0.080634	2.25983	0.084967
Error	4.281765	120	0.035681		

Significant terms ($p<0.05$) are indicated in red.

Table S2. Jacobs' Index ANOVA table.

Effect	SS	DF	MS	F	p
Intercept	20.92748	1	20.92748	134.4229	0.000000
Signal Complexity	1.14662	1	1.14662	7.3651	0.009764
Reward Reliability	0.79329	1	0.79329	5.0955	0.029518
Choice Number	5.18383	1	5.18383	33.2971	0.000001
Signal Complexity x Reward Reliability	0.00713	1	0.00713	0.0458	0.831688
Signal Complexity x Choice Number	0.09198	1	0.09198	0.5908	0.446608
Reward Reliability x Choice Number	0.02664	1	0.02664	0.1711	0.681355
Signal Complexity x Reward Reliability x Choice Number	0.00248	1	0.00248	0.0159	0.900256
Error	6.22736	40	0.15568		
Choice Block	2.20669	3	0.73556	20.9119	0.000000
Choice Block x Signal Complexity	0.38928	3	0.12976	3.6891	0.013905
Choice Block x Reward Reliability	0.06273	3	0.02091	0.5945	0.619796
Choice Block x Choice Number	0.30489	3	0.10163	2.8893	0.038364
Choice Block x Signal Complexity x Reward Reliability	0.01528	3	0.00509	0.1448	0.932818
Choice Block x Signal Complexity x Choice Number	0.07443	3	0.02481	0.7053	0.550640
Choice Block x Reward Reliability x Choice Number	0.04928	3	0.01643	0.4670	0.705849
Choice Block x Signal Complexity x Reward Reliability x Choice Number	0.05594	3	0.01865	0.5301	0.662474
Error	4.22091	120	0.03517		

Significant terms (p<0.05) are indicated in red.

Table S3. Foraging efficiency ANOVA table.

Effect	SS	DF	MS	F	p
Intercept	62995.71	1	62995.71	267.1179	0.000000
Signal Complexity	62.42	1	62.42	0.2647	0.609761
Reward Reliability	128.68	1	128.68	0.5457	0.464412
Choice Number	404.48	1	404.48	1.7151	0.197795
Signal Complexity x Reward Reliability	18.40	1	18.40	0.0780	0.781424
Signal Complexity x Choice Number	228.60	1	228.60	0.9693	0.330763
Reward Reliability x Choice Number	94.58	1	94.58	0.4010	0.530163
Signal Complexity x Reward Reliability x Choice Number	13.26	1	13.26	0.0562	0.813790
Error	9433.39	40	235.83		
Choice Block	4564.37	3	1521.46	55.5592	0.000000
Choice Block x Signal Complexity	196.49	3	65.50	2.3918	0.071967
Choice Block x Reward Reliability	96.79	3	32.26	1.1782	0.321037
Choice Block x Choice Number	63.29	3	21.10	0.7704	0.512765
Choice Block x Signal Complexity x Reward Reliability	62.07	3	20.69	0.7555	0.521245
Choice Block x Signal Complexity x Choice Number	60.41	3	20.14	0.7353	0.532924
Choice Block x Reward Reliability x Choice Number	69.07	3	23.02	0.8408	0.474105
Choice Block x Signal Complexity x Reward Reliability x Choice Number	111.20	3	37.07	1.3536	0.260366
Error	3286.14	120	27.38		

Significant terms (p<0.05) are indicated in red.

Table S4. Decision latency ANOVA table.

Effect	SS	DF	MS	F	p
Intercept	1.34324	1	1.34324	123.3771	0.000000
Signal Complexity	0.00509	1	0.00509	0.4676	0.498034
Reward Reliability	0.00362	1	0.00362	0.3327	0.567280
Choice Number	0.02290	1	0.02290	2.1030	0.154808
Signal Complexity x Reward Reliability	0.02005	1	0.02005	1.8415	0.182381
Signal Complexity x Choice Number	0.00316	1	0.00316	0.2906	0.592854
Reward Reliability x Choice Number	0.01039	1	0.01039	0.9547	0.334403
Signal Complexity x Reward Reliability x Choice Number	0.03584	1	0.03584	3.2922	0.077115
Error	0.43549	40	0.01089		
Choice Block	0.48729	3	0.16243	23.9244	0.000000
Choice Block x Signal Complexity	0.05728	3	0.01909	2.8121	0.042310
Choice Block x Reward Reliability	0.00193	3	0.00064	0.0947	0.962848
Choice Block x Choice Number	0.04536	3	0.01512	2.2271	0.088531
Choice Block x Signal Complexity x Reward Reliability	0.04424	3	0.01475	2.1719	0.094881
Choice Block x Signal Complexity x Choice Number	0.02584	3	0.00861	1.2688	0.288260
Choice Block x Reward Reliability x Choice Number	0.01107	3	0.00369	0.5437	0.653324
Choice Block x Signal Complexity x Reward Reliability x Choice Number	0.08080	3	0.02693	3.9670	0.009776
Error	0.81472	120	0.00679		

Significant terms (p<0.05) are indicated in red.

Table S5. Foraging accuracy ANOVA table.

Effect	SS	DF	MS	F	p
Intercept	40.99603	1	40.99603	1114.527	0.000000
Signal Complexity	0.12813	1	0.12813	3.483	0.069328
Reward Reliability	0.05880	1	0.05880	1.599	0.213427
Choice Number	0.00013	1	0.00013	0.004	0.952291
Signal Complexity x Reward Reliability	0.01470	1	0.01470	0.400	0.530874
Signal Complexity x Choice Number	0.04083	1	0.04083	1.110	0.298382
Reward Reliability x Choice Number	0.00163	1	0.00163	0.044	0.834173
Signal Complexity x Reward Reliability x Choice Number	0.01080	1	0.01080	0.294	0.590921
Error	1.47133	40	0.03678		
Choice Block	0.46837	3	0.15612	12.671	0.000000
Choice Block x Signal Complexity	0.21720	3	0.07240	5.876	0.000889
Choice Block x Reward Reliability	0.12973	3	0.04324	3.510	0.017458
Choice Block x Choice Number	0.07480	3	0.02493	2.024	0.114228
Choice Block x Signal Complexity x Reward Reliability	0.01903	3	0.00634	0.515	0.672776
Choice Block x Signal Complexity x Choice Number	0.00597	3	0.00199	0.161	0.922110
Choice Block x Reward Reliability x Choice Number	0.03343	3	0.01114	0.904	0.441180
Choice Block x Signal Complexity x Reward Reliability x Choice Number	0.01213	3	0.00404	0.328	0.804927
Error	1.47853	120	0.01232		

Significant terms (p<0.05) are indicated in red.

References

Chittka, L. (1992). The colour hexagon: a chromaticity diagram based on photoreceptor excitations as a generalized representation of colour opponency. *Journal of Comparative Physiology A*. 170, 533-543.

Skorupski, R. and Chittka L. (2010). Photoreceptor spectral sensitivity in the bumblebee, *Bombus impatiens* (Hymenoptera: Apidae). *PLoS ONE*. 5, e12049.