Reproductive isolation and gene flow vary among contact zones between incipient species

Catherine Lynn Debban
Snellville, GA

B.S., University of Georgia, 2012

A dissertation presented to the Graduate Faculty
of the University of Virginia in Candidacy for the Degree of
Doctor of Philosophy

Department of Biology

University of Virginia
July, 2019

ABSTRACT

The speciation process separates a single species into multiple lineages of independently evolving taxa, resulting in the diversification of the lineages into new species. However, it is not a unidirectional process that always results in speciation after it begins; secondary contact between partially-isolated lineages can either catalyze an increased rate of diversification via reinforcement or allow the lineages to merge back together into a single species. Studies comparing multiple natural contact zones between incipient species are necessary to understand what factors influence the outcome of secondary contact. In this study, I use the plant *Campanula americana* to test how consistent the outcome of secondary contact is across the range of incipient species. *Campanula americana* is divided into an Appalachian and a Western lineage that are separated by reproductive isolation, and are in contact in North Carolina, Pennsylvania, and Virginia. I found that in the North Carolina contact zone, there is low pre- and postzygotic reproductive isolation relative to allopatry, and a high level of gene flow between the lineages. Together, these findings indicate that these incipient lineages are merging together in North Carolina. By contrast, in the Pennsylvania and Virginia contact zones, gene flow between lineages is low. In Pennsylvania, postzygotic reproductive isolation is almost as high as it is in allopatry, and prezygotic isolation is higher than in allopatry, suggesting that reinforcement may be driving increased divergence in Pennsylvania. Together, this work demonstrates that even between the same two lineages, the outcome of secondary contact can vary among contact zones. Divergent outcomes are hypothesized to be driven by differences in initial levels of reproductive isolation between the lineages at time of contact and/or by the geographic structure of the contact zones. Secondary contact is a dynamic process whose outcome can be changed by factors that vary within a species.

ACKNOWLEDGEMENTS

I would like to thank the many people who have helped me as I worked on this dissertation. First, I would like to thank my advisor, Laura Galloway. She has provided incredible mentorship, guidance, and support that were crucial during my time as a graduate student. Her insight and enthusiasm always pushed me to be a better scientist, and I could not have asked for a better mentor.

I would also like to thank the members of my committee, who have been instrumental as I developed and carried out my research projects. Ben Blackman and Doug Taylor were excellent first readers who gave invaluable feedback on my work. Early in my dissertation, Ben helped me develop experiments, and more recently Doug has given key feedback as I analyzed my results. Alan Bergland's help with bioinformatics was crucial as I analyzed my sequencing data. Dave Carr and Bob Cox's feedback on my research has always been insightful.

The Galloway lab, past and present, has been a great set of colleagues who have been a great source of support, friendship, and feedback on my research. Thank you to Karen Barnard-Kubow and Brittany Sutherland, who helped me get oriented in the lab and were always willing to chat about my research. Thank you also to Matt Koski, Hanqin Wu, Hanna Makowski, and Holly Prendeville, and to all of the graduate students who have joined our lab group meeting, for your valuable feedback.

This dissertation physically could not have gotten done without logistical support from many people. Thank you to AnhThu Nguyen for assistance with planning and preparation of my sequencing libraries. Thank you also to everyone who has helped me with field work on this project. Judy Debban, Paul Debban, Rebecca Risser, and Jeffrey Risser all generously donated their time to help me find populations and plant experiments. Jeanette Engel provided me a base

of operations for field work in Pennsylvania. Thank you to Mountain Lake Biological Station and its staff, especially Jaime Jones and Eric Nagy. I would especially like to thank the undergraduates who helped me throughout my project: Dustin Wessels, Richard Hankins, Linnea Fraser, Claire Waterhouse, and Samuel Adu.

My family and friends have been invaluable as I completed this dissertation. I'd like to thank Judy, Paul, Elizabeth, and Theo Debban for their love and support. In addition, thank you to my friends, fellow graduate students, and housemates, whose friendship made this journey brighter.

Financial support for this dissertation was provided by an NSF Graduate Research Fellowship, a Rosemary Grant Award from the Society for the Study of Evolution, a Margaret Walton Scholarship for Mountain Lake, and the Mountain Lake Biological Station REU-Site grant.

TABLE OF CONTENTS

INTRODUCTION

The speciation process separates a single species into multiple lineages of independently evolving species, allowing for the diversification of the new species (de Queiroz 2007). An increase in reproductive isolation and a decrease in gene flow among these lineages is usually integral to the formation of new species. However, factors that increase reproductive isolation and decrease gene flow among lineages are not always irreversible (Mayr 1963). Geographic barriers separating lineages, such as glaciers or mountain ranges, can disappear or ranges can expand around them. Incompatibilities with a genetic basis can be purged when partially-isolated lineages are in contact (Felsenstein 1981), leading to potential for merging of differentiated lineages. Understanding how and when lineages merge during secondary contact is important to understand the factors that are most important to determining the outcome of this dynamic phase of the speciation process.

Reproductive isolation encompasses barriers to gene flow among lineages, ranging from geographic isolation due to ecological niche differences to low hybrid performance and reproduction (Coyne and Orr 2004). These barriers can happen at any life stage, either preventing the two lineages from breeding (prezygotic) or reducing the performance of hybrids (postzygotic); both lead to reduced gene flow between the lineages. Reproductive isolation can work through mechanisms intrinsic to the species or extrinsic (dependent on their environment). The most commonly discussed model of species formation is the allopatric model, in which a single species is separated into groups by a geographic barrier, and those groups begin to evolve as independent lineages (Mayr 1963). Reproductive isolation forms via drift or selection as new alleles fix in each new lineage. At first, any new alleles that fix must be compatible with the original genotypes, but later the new alleles must only be compatible with the genotypes of the

current population, and not with the other isolated lineage. In this way, new alleles that fix in each lineage may be incompatible with the other lineage when they come back into contact (Bateson 1909; Dobzhansky 1937; Muller 1942). While allopatric accumulation of reproductive isolation is a common mechanism, reproductive isolation can also arise in sympatry via reinforcement (Dobzhansky 1937), genome duplication, or niche diversification (Schluter et al. 1992), among other methods.

Reproductive isolation often does not completely separate the two allopatric lineages before they come back into contact, a process called secondary contact. This makes secondary contact an important and dynamic step in the speciation process, because it can allow lineages to merge together (Felsenstein 1981; Barton and Bengtsson 1986; Taylor et al. 2005; Kearns et al. 2018) or to diverge more quickly via reinforcement (Dobzhansky 1937; Noor 1999). Even a small amount of gene flow may be enough to allow lineages to overcome any barriers formed in allopatry (Felsenstein 1981). As hybrids are formed, backcross, and interbreed, the ones that are the most fit will be missing the alleles that cause incompatibilities, and have only the alleles from one lineage or the other at the loci involved in incompatibility. This decreases reproductive isolation and in turn allows more gene flow and more loss of reproductive isolation at the contact zone. However, under a wide variety of conditions selection favors individuals that do not interbreed with the other lineage, resulting in increases in prezygotic isolation at contact zones (Noor 1999; Servedio and Noor 2003; Bank et al. 2011).

Whether merging of lineages or reinforcement of speciation is the outcome of secondary contact depends on many conditions. Initial degree of reproductive isolation at contact can be an important factor in contact zone outcome; low amounts of reproductive isolation are more likely to be purged, while higher levels are likely to cause reinforcement (Coyne and Orr 2004). Effect

size of each reproductive isolating barrier can also influence outcome. When prezygotic barriers that arise in contact zones have small effects, they are less likely to complete speciation than if they have large effects (Bank et al. 2011). Genetic architectures of the incompatibilities are also important to the outcome of secondary contact. Tightly-linked incompatibilities are less likely to be broken apart by recombination, and therefore more likely to lead to speciation (Rieseberg 2001; Bank et al. 2012; Lindtke and Buerkle 2015). However, while cytonuclear incompatibilities (those involving both nuclear and either mitochondrial or chloroplast genomes) are unlinked, they can persist at higher rates of migration than nuclear-nuclear incompatibilities because cytoplasmic genomes do not recombine and are usually inherited solely from mothers (Höllinger and Hermisson 2017). Finally, the geographic structure of contact zones can be important to the outcome of secondary contact. When contact zones are a mosaic, reinforcement is more likely to result than if the contact zone is more clinal (Cain et al. 1999).

Reproductive isolation often varies across the range of a species, both between lineages of a species (McDermott and Noor 2011; Cutter 2012; Corbett-Detig et al. 2013) and between different populations of two separate species (Sweigart et al. 2007; Kozlowska et al. 2011; Mandeville et al. 2015). This variation is often common when extrinsic factors are important to reproductive isolation, because ecology varies across species' ranges (Hatfield and Schluter 1999). However, intrinsic reproductive isolation typically scales with genetic distance, as is the case in ring species (Irwin et al. 2001; Alcaide et al. 2014). Since the degree and type of reproductive isolation influences the outcome of secondary contact, variation in reproductive isolation across a range can yield different outcomes of secondary contact at different contact zones.

In this study I used *Campanula americana* to explore the variation in gene flow and reproductive isolation among contact zones between partially isolated lineages. *Campanula americana* is an herbaceous plant species that is divided into reproductively isolated lineages that are in contact in three locations across its range. Rapidly-evolving chloroplast markers divide *C. americana* into two lineages: an Appalachian lineage common in the Appalachian mountains from Tennessee to Pennsylvania, and a Western lineage stretching from the Appalachian mountains west to Kansas and east of the Appalachian mountains in Virginia (Barnard-Kubow et al. 2015). Hybrids between these two lineages have reduced germination and survival compared to their pure-lineage counterparts (Barnard-Kubow et al. 2016; Barnard-Kubow and Galloway 2017). Reductions in hybrid survival are largely driven by chlorosis of plants with Western cytoplasm on hybrid nuclear backgrounds; hybrids with Appalachian cytoplasm rarely express chlorosis (Barnard-Kubow et al. 2016). During the last glaciation, the Western lineage was confined to a southern refugium on the Gulf Coast and an eastern refugium on the Atlantic coast, and the Appalachian lineage to microrefugia within the Appalachian Mountains (Barnard-Kubow et al. 2015). Over the last 20,000 years, the Western lineage has expanded from the Gulf Coast north as far as Pennsylvania, and from the coast of Virginia west to the Appalachian Mountains. This has created contact zones between the Appalachian and Western lineages in North Carolina, Pennsylvania, and Virginia.

In Chapter 1, I tested whether postzygotic reproductive isolation is lower in two contact zones between the partially isolated Western and Appalachian lineages of *Campanula americana* than it is in allopatry. If allopatry represents an approximation of the initial reproductive isolation between the lineages at time of secondary contact, contact zones where the lineages are merging may have lower levels of reproductive isolation than in allopatry if they have purged their

reproductive incompatibilities. If the lineages are experiencing reinforcement, no difference in post-zygotic reproductive isolation is expected between allopatry and contact zones. I found that in the North Carolina contact zone, the lineages experienced no reproductive isolation due to germination, whereas germination contributed strongly to reproductive isolation in allopatry and in the Pennsylvania contact zone. An asymmetrical cytonuclear incompatibility was found in allopatry and persisted across both contact zones, reducing survival of hybrids with Western lineage dams. When hybrids from the North Carolina contact zone had Appalachian dams, they survived well and produced more flowers than their parents. The low postzygotic reproductive isolation in North Carolina and high postzygotic reproductive isolation in Pennsylvania and allopatry suggests that the Appalachian and Western lineages are merging in North Carolina but not in Pennsylvania.

In Chapter 2, I tested whether gene flow has occurred between the Appalachian and Western lineages of *C. americana* at three contact zones, and whether it differed between those zones. I found that gene flow was high in the North Carolina contact zone and low in the Pennsylvania and Virginia contact zones. This confirmed the findings of Chapter 1 that the lineages are merging in North Carolina but not in Pennsylvania or Virginia. By combining information from this study about phylogenetic relationships between populations of *C. americana* with previous studies of within-lineage reproductive isolation (Barnard-Kubow et al. 2016; Barnard-Kubow and Galloway 2017), I hypothesized that incompatibility due to germination arose in the Western lineage in populations that are now in allopatry and Pennsylvania, but not those in North Carolina or Virginia. This raised the possibility that the difference in contact outcome between North Carolina and Pennsylvania may have been due in part to differences in reproductive isolation at time of initial contact. However, this did not

account for the differences in contact outcome between North Carolina and Virginia, so other factors such as local habitat suitability and geographic structure of contact zones were likely also important.

In Chapter 3, I tested whether prezygotic reproductive isolation was different in two contact zones of *C. americana* than in allopatry. I expected that if lineages were merging, prezygotic reproductive isolation would be lower in contact zones than in allopatry, and if they have diverged via reinforcement, reproductive isolation would be higher in contact zones. Concordant with Chapters 1 and 2, I found that in North Carolina prezygotic reproductive isolation was lower than in allopatry. Whereas, in Pennsylvania, all prezygotic barriers present in allopatry were also present and the Western lineage begins flowering two weeks before the Appalachian lineage, conferring additional prezygotic isolation there. These data suggest that the Appalachian and Western lineages are merging in North Carolina, and that reinforcement may be contributing to the divergence of lineages in Pennsylvania.

Taken together, the findings presented here show that the outcome of secondary contact can vary across a range. While the lineages of *C. americana* are merging together in a North Carolina contact zone, they appear to be diverging in Virginia and Pennsylvania contact zones. This is likely driven at least partially by differences in initial reproductive isolation at time of contact, and is likely also driven by some other factor such as frequency of hybridization at the contact zones. This demonstrates the dynamic role of secondary contact between partially isolated lineages. The speciation process is not a unidirectional progression; factors that vary within lineages can have important effects on the trajectory of speciation.

References

Alcaide, M., E. S. C. Scordato, T. D. Price, and D. E. Irwin. 2014. Genomic divergence in a ring species complex. Nature 511:83–85.

Bank, C., J. Hermisson, and M. Kirkpatrick. 2011. Can reinforcement complete speciation? Evolution 66:229–239.

Bank, C., R. Bürger, and J. Hermisson. 2012. The limits to parapatric speciation: Dobzhansky–Muller incompatibilities in a continent–island model. Genetics 191:845–863.

Barnard-Kubow, K. B., and L. F. Galloway. 2017. Variation in reproductive isolation across a species range. Ecol Evol 7:9347–9357.

Barnard-Kubow, K. B., C. L. Debban, and L. F. Galloway. 2015. Multiple glacial refugia lead to genetic structuring and the potential for reproductive isolation in a herbaceous plant. American Journal of Botany 102:1842–1853.

Barnard-Kubow, K. B., N. So, and L. F. Galloway. 2016. Cytonuclear incompatibility contributes to the early stages of speciation. Evolution 70:2752–2766.

Barton, N., and B. O. Bengtsson. 1986. The barrier to genetic exchange between hybridising populations. Heredity 56:357–376.

Bateson, W. 1909. Heredity and variation in modern lights. Pp. 85–101 *in* A. C. Seward, ed. Darwin and Modern Science. Cambridge, MA.

Cain, M. L., V. Andreasen, and D. J. Howard. 1999. Reinforcing selection is effective under a relatively broad set of conditions in a mosaic hybrid zone. Evolution 53:1343–1353.

Corbett-Detig, R. B., J. Zhou, A. G. Clark, D. L. Hartl, and J. F. Ayroles. 2013. Genetic incompatibilities are widespread within species. Nature 504:135–137.

Coyne, J. A., and H. A. Orr. 2004. Speciation. Sinauer Associates, Inc., Sunderland, MA, USA.

Cutter, A. D. 2012. The polymorphic prelude to Bateson–Dobzhansky–Muller incompatibilities. TREE 27:210–219.

de Queiroz, K. 2007. Species concepts and species delimitation. Systematic Biology 56:879–886.

Dobzhansky, T. 1937. Genetics and the origin of species. Columbia University Press, New York.

Felsenstein, J. 1981. Skepticism towards Santa Rosalia, or Why are there so few kinds of animals? Evolution 35:124–138.

Hatfield, T., and D. Schluter. 1999. Ecological speciation in sticklebacks: environment-dependent hybrid fitness. Evolution 53:866–873.

Höllinger, I., and J. Hermisson. 2017. Bounds to parapatric speciation: A Dobzhansky-Muller incompatibility model involving autosomes, X chromosomes, and mitochondria. Evolution 71:1366–1380.

Irwin, D. E., J. H. Irwin, and T. D. Price. 2001. Ring species as bridges between microevolution and speciation. Genetica 112-113:223–243.

Kearns, A. M., M. Restani, I. Szabo, A. Schrøder-Nielsen, J. A. Kim, H. M. Richardson, J. M. Marzluff, R. C. Fleischer, A. Johnsen, and K. E. Omland. 2018. Genomic evidence of speciation reversal in ravens. Nature Communications 9:906.

Kozlowska, J. L., A. R. Ahmad, E. Jahesh, and A. D. Cutter. 2011. Genetic variation for postzygotic reproductive isolation between *Caenorhabditis briggsae* and *Caenorhabditis sp. 9*. Evolution 66:1180–1195.

Lindtke, D., and C. A. Buerkle. 2015. The genetic architecture of hybrid incompatibilities and their effect on barriers to introgression in secondary contact. Evolution 69:1987–2004.

Mandeville, E. G., T. L. Parchman, D. B. McDonald, and C. A. Buerkle. 2015. Highly variable reproductive isolation among pairs of *Catostomus* species. Molecular Ecology 24:1856–1872.

Mayr, E. 1963. Animal species and evolution. Bellknap Press of Harvard University, Cambridge, MA.

McDermott, S. R., and M. A. F. Noor. 2011. Genetics of hybrid male sterility among strains and species in the *Drosophila pseudoobscura* species group. Evolution 65:1969–1978.

Muller, H. J. 1942. Isolating mechanisms, evolution, and temperature. Biological Symposia 6:71-125.

Noor, M. A. F. 1999. Reinforcement and other consequences of sympatry. Heredity 83:503–508.

Rieseberg, L. H. 2001. Chromosomal rearrangements and speciation. TREE 16:351–358.

Schluter, D., and J. D. McPhail. 1992. Ecological character displacement and speciation in sticklebacks. American Naturalist 140: 85-108.

Servedio, M. R., and M. A. F. Noor. 2003. The role of reinforcement in speciation: theory and data. Annu. Rev. Ecol. Evol. Syst. 34:339–364.

Sweigart, A. L., A. R. Mason, and J. H. Willis. 2007. Natural variation for a hybrid incompatibility between two species of *Mimulus*. Evolution 61:141–151.

Taylor, E. B., J. W. Boughman, M. Groenenboom, M. Sniatynski, D. Schluter, and J. L. Gow. 2005. Speciation in reverse: morphological and genetic evidence of the collapse of a three-spined stickleback (*Gasterosteus aculeatus*) species pair. Molecular Ecology 15:343–355.

CHAPTER 1:

Postzygotic reproductive isolation between incipient species varies among contact zones

Abstract

Secondary contact between partially isolated lineages of a species is an important step in the process of speciation because it can lead to either the acceleration or reversal of the speciation process. However, the factors that are important in determining which outcome results from a particular contact zone are rarely tested in a natural setting. Here I test whether partially isolated lineages of the herb *Campanula americana* are merging together with secondary contact in two independent contact zones. I made F1 hybrids for multiple populations between genetically distinct Appalachian and Western lineages for both contact zones and populations in allopatry, then grew the hybrids and parental seed in a natural environment. A strong reduction in survival of hybrids with Western-lineage mothers was consistent across both contact zones and allopatry, indicating that this cytonuclear component of reproductive isolation persists at contact zones. Conversely, while hybrids between allopatric populations and in one contact zone had reduced germination, germination did not contribute to reproductive isolation in the other contact zone. In that zone, hybrids also produced more fruit than their parents. This reduction in postzygotic reproductive isolation relative to allopatry suggests that secondary contact is causing the lineages to merge together in one contact zone, but not the other. Variation in reproductive isolation between contact zones demonstrates that lineages may not respond consistently to secondary contact across their ranges, and therefore the outcome of secondary contact between two lineages depends on factors that vary within species.

Introduction

The process of speciation often begins when one species is separated into isolated groups by a barrier, such as a glacier or mountain range. In isolation, the groups evolve as independent lineages, accumulating genetic differences that can cause incompatibilities between the lineages (Bateson 1909; Dobzhansky 1937; Muller 1942). While long lasting, the geographic barriers keeping lineages apart are often impermanent because of processes like climate change and range expansion. When these barriers disappear, or species ranges expand around them, previously isolated lineages can come into secondary contact. In response to secondary contact, one or both of the lineages may develop additional prezygotic isolating barriers through a process called reinforcement in order to avoid mating with the other incompatible lineage (Dobzhansky 1937; Noor 1999; Matute 2010; Hopkins 2013). However, in many cases even a small amount of gene exchange may be enough for the lineages to purge incompatible alleles, reducing postzygotic isolation and allowing them to merge (Felsenstein 1981; Barton and Bengtsson 1986; Taylor et al. 2005; Seehausen et al. 2008; Lindtke and Buerkle 2015; Kearns et al. 2018). This dichotomy of outcomes means that secondary contact is an important stage of the speciation process because it can facilitate either the acceleration or the reversal of speciation.

Many factors can change the dynamics and outcomes of secondary contact between partially isolated lineages. Frequency of contact, species relative abundance, and the geographic layout of contact zones can influence outcomes (Lepais et al. 2009). In particular, patchy and mosaic contact zones are more likely to result in reinforcement, while more continuous contact will more likely result in lineage collapse (Cain et al. 1999). Small differences in time spent in isolation can also influence the outcome of contact between the lineages. Once an initial incompatibility forms between the isolated groups, additional incompatibilities can follow more

easily because they do not experience the initial genotype of the population, and therefore do not need to be compatible with it (Matute et al. 2010; Moyle and Nakazato 2010). This means that once an incompatibility arises, more time spent in isolation can allow groups to rapidly accumulate additional isolating mechanisms, making it more difficult for them to merge back together. A well-known example of this phenomenon is ring species, where populations in one part of the range interbreed freely, but are isolated from each other at the other end of the range (Irwin et al. 2001; Alcaide et al. 2014). In non-ring species, gene flow between lineages can vary among contact zones due to ecological context (Lepais et al. 2009; Mandeville et al. 2015; Kingston et al. 2017). However, it is unknown how frequently the same outcome is reached when several contact zones between lineages are present and isolating mechanisms are primarily intrinsic.

The genetic architecture of isolating mechanisms is important to their persistence at contact zones. Tightly linked incompatibilities may be difficult to purge because recombination is not able to break them apart (Noor et al. 2001; Navarro and Barton 2003; Twyford and Friedman 2015), and theory and simulations suggest that reproductive isolation caused by incompatibilities between unlinked loci should break down easily by recombination (Barton 1979; Rieseberg et al. 1999; Lindtke and Buerkle 2015). Since loci in chloroplast genomes are not linked with nuclear genomes, cytonuclear incompatibilities might be expected to break down easily. However, unlinked incompatibilities may maintain reproductive isolation when those incompatibilities are strong (Bank et al. 2012). Cytonuclear incompatibilities especially can likely persist more easily than nuclear-nuclear incompatibilities because cytoplasmic genomes do not recombine and are usually inherited solely from mothers (Höllinger and Hermisson 2017). Since multiple barriers are often involved in reproductive isolation between lineages and each

barrier may arise from a different genetic architecture, secondary contact may lead to the degradation of some barriers but not others.

Cytonuclear incompatibilities may result in asymmetrical gene flow in which nuclear genes move into the group with a compatible cytoplasmic genome (Tiffin et al. 2001; Sardell and Uy 2016; Ley and Hardy 2017). This may lead to high prevalence of the compatible chloroplast at contact zones, and the spread of those chloroplasts within hybrids. If chloroplast genomes are sometimes inherited from fathers, the compatible chloroplast may "rescue" hybrids that would have otherwise been inviable (Barnard-Kubow et al. 2017). Biparental chloroplast inheritance may also allow the compatible chloroplast to more easily spread into populations of the usually-incompatible lineage via pollen movement, and subsequently persist by backcrossing within those populations (Rieseberg and Soltis 1991; Wolfe and Elisens 1995). However, biparental inheritance is rare, and most effects of cytonuclear incompatibilities are likely to manifest as asymmetries in gene flow.

In this study, I compared postzygotic reproductive isolation at two contact zones and allopatry of lineages of the species *Campanula americana* in order to understand the outcome of secondary contact. The last glaciation divided *C. americana*'s range into southern refugia along the Gulf Coast, eastern refugia along the Atlantic coast, and refugia within the Appalachian mountains (Barnard-Kubow et al. 2015). The groups of *C. americana* that resided in the Gulf and Eastern refugia now belong to the Western lineage, and the groups that resided in the Appalachian refugia now belong to the Appalachian lineage. There is postzygotic reproductive isolation between the lineages, manifesting as reductions in hybrid germination and survival (Barnard-Kubow et al. 2016; Barnard-Kubow and Galloway 2017). Reductions in survival are partially caused by a cytonuclear incompatibility between the Western chloroplast and

Appalachian nuclear DNA that causes F1 hybrids to have inactive chloroplasts (Barnard-Kubow et al. 2016). In contrast, germination reduction in hybrids appears to be caused by incompatibilities between nuclear genes (Barnard-Kubow et al. 2016; Barnard-Kubow and Galloway 2017). Since the last glaciation, the lineages have each expanded out of their refugia and come into secondary contact, with contact zones between the Appalachian and Western lineages in Pennsylvania and North Carolina.

Here I created hybrid and parental seed types from two contact zones and among allopatric populations, and evaluated them in natural habitats for germination, survival, and reproduction to estimate postzygotic reproductive isolation. In doing so, I addressed three questions about how secondary contact influences postzygotic reproductive isolation in *C. americana*: (1) Is postzygotic reproductive isolation lower in contact zones than in regions without secondary contact? (2) Do contact zones with different evolutionary histories also differ in reproductive isolation? (3) Do reproductive incompatibilities with distinct underlying genetic architecture differ in their strength among contact zones?

Methods

Study system

The American Bellflower (*Campanula americana* L. = *Campanulastrum americanum* Small, Campanulaceae) is an outcrossing autotetraploid annual or biennial herb (Gadella 1964; Galloway et al. 2003; Galloway and Etterson 2005). It grows in the eastern United States, with a distribution that spans south to Florida, north to Pennsylvania, east to coastal Virginia, and west to Kansas. It grows in partial-shade environments in forest edges, light gaps and roadsides. Seeds germinate in spring and fall, and plants bloom midsummer (June-August) after winter

vernalization (Baskin and Baskin 1984). Flowers are protandrous, opening as male and transitioning to female phase following pollen removal (Evanhoe and Galloway 2002).

Campanula americana is comprised of two major genetic lineages. A Western lineage extends from the Appalachians west to Nebraska and from Florida north to Michigan, and is also found east of the Blue Ridge Mountains (Fig. 1C). An Appalachian lineage is found in the Appalachian Mountains extending from Georgia to Pennsylvania (Fig. 1C) (Barnard-Kubow et al. 2015). Postzygotic reproductive isolation is strong between these lineages and is largely caused by cytonuclear incompatibility between the chloroplast of the Western lineage and the nuclear DNA of the Appalachian lineage (Barnard-Kubow et al. 2016). This cytonuclear interaction often gives hybrids with Western chloroplasts a white leafed (bleached) phenotype, reducing their survival. Although chloroplasts are typically maternally inherited, in *C. americana* individuals may also inherit chloroplasts from both parents, which can lead to partially chlorotic phenotypes with light green or variegated white-and-green leaves in hybrids (Barnard-Kubow et al. 2017). In addition to the cytonuclear incompatibility affecting survival, germination is reduced in Appalachian-Western hybrids regardless of maternal parent (Barnard-Kubow et al. 2016).

Populations & crossing

Focal populations of the Appalachian and Western lineages were chosen from two contact zones as well as from allopatry. For convenience I refer to each contact zone and allopatry as individual regions. I chose 26 focal populations from the three regions: six populations from a contact zone in Pennsylvania, twelve from a contact zone in North Carolina, and eight from allopatry (Fig. 1A, Table S1). Half of the populations within each region were from the Appalachian lineage, and half from the Western lineage. In each contact zone,

populations were chosen to represent a range of geographic distances from populations of the other lineage. The lineage of each population was confirmed as Appalachian or Western by Sanger sequencing chloroplast markers previously validated for their utility in distinguishing *C. americana* lineages (Barnard-Kubow et al. 2015).

I collected seeds by maternal family from each population and grew them in a controlled environment. I planted at least 20 seeds for each of the 26 populations (705 total), representing as many families as possible (12-30 families per population, mean 21). Additional seeds were planted if the germination rate of a population was low. Seeds were planted singly in a 3:1 mixture of Metromix and Turface, and germinated in a growth chamber set at 21°C day, 14°C night with 12-hour days for 44 days. I then moved the plants to a cold room to vernalize for 82 days at 5°C with 12-hour days. After vernalization, I transplanted rosettes into conetainers, arranged them in random order, and moved them into the greenhouse where lights extended day length to 16 hours. Plants were fertilized every other week until bolting, then weekly, and watered as needed.

Two types of crosses were made on each plant. I created "parental" type seeds by crossing individuals from the same population for each focal population (26 seed types). I also created F1 hybrids between Appalachian and Western populations within each region. Each population was crossed to two populations from the other lineage within the same region to make F1 between-population hybrids (Fig. 1B). I performed crosses in both directions so that Appalachian and Western populations each served as maternal plants. This resulted in 52 seed types (8 allopatric pairs + 12 North Carolina contact zone pairs + 6 Pennsylvania contact zone pairs = 26 pairs * 2 crossing directions each; Table S2). Flowers were emasculated prior to

crossing. On average, six pollinations were conducted for each of the 78 seed types. Fruits were collected after maturation and stored at room temperature until planting.

Growth-chamber germination & Field performance

To measure hybrid performance, I germinated seeds in growth chambers and transplanted rosettes into the field after vernalization. For each seed type, I planted 40 seeds across 20 cells (two seeds each) in randomized locations in germination flats (2080 seeds total). I recorded germination and leaf color of each seedling (bleached, light, variegated, or green; Fig. S1) every two weeks for eight weeks. Plants were then vernalized at 5ºC for two months. If two seedlings germinated and survived in the same cell, I removed a random seedling prior to vernalization. I planted seedlings in a site near a natural Appalachian allopatric population of *C. americana* (near Mountain Lake Biological Station, Fig. 1), that is similar in climate and intermediate in latitude to the Pennsylvania and North Carolina contact zones. Plots were situated on a hillside in light gaps of a forest, and native vegetation was left intact to provide a competitive environment. Rosettes were planted 0.25m apart into five fenced plots ("blocks") on March 25, 2017. A total of 899 plants were transplanted, fewer than the planned 1040 due to reduced germination and survival. On April 4, June 16, July 18, August 13, and September 26, I recorded survival and for the latter two dates the number of fruits on each plant.

Field germination and seedling survival

Germination and early survival were also determined under natural conditions. I planted hybrid and parental type seeds into the same site in which adult performance was measured. Seeds were planted in 24 blocks. Each block was divided into 48 cells by 32mm square wax-coated paper sleeves, filled with a mixture of field soil and potting mix. 18 replicates of each seed type, divided evenly among maternal plants, were distributed across blocks (936 replicates

total). I randomized seed types within blocks, planting 10 seeds per cell. For 16 of 78 seed types, fewer than 10 seeds were planted per replicate (mean 5.4 seeds) due to inadequate seed numbers. I planted seeds in the field on September 1, 2016, at the beginning of the natural seed dispersal period. On September 23, October 9, October 19, and November 6, I recorded the number of seedlings in each cell. By November 6, germination had nearly ceased due to cool temperatures. Proportion germination and seedling survival was calculated as the maximum number of seedlings in each cell divided by the number of seeds planted. Since the cells were checked every 2 weeks, any germinant that died before being recorded would count the same as those that did not germinate. Therefore, this measurement contains elements of both germination and survival.

Statistical analysis

I used an Aster Model to analyze cumulative fitness across the life cycle. Aster modeling is a statistical tool that allows users to analyze cumulative fitness while accounting for different probability distributions at each life stage (Geyer et al. 2007). The model included the life stages of growth-chamber germination, survival and fruit production. Field germination could not be included in this model since it was assessed using different individuals than the later life stages whereas growth chamber germinants were planted in field plots and followed through reproduction. The model included fixed effects of cross type (hybrid and parent), region (allopatry, Pennsylvania, and North Carolina), maternal lineage (Appalachian and Western), and all interactions of these effects. Random effects included maternal population (nested within region) and paternal population (nested within region). A significant cross-type effect with hybrid plants having lower fitness than parental plants indicates reproductive isolation. A cross-type*region effect indicates that reproductive isolation varies among regions. Finally, a cross-

type*maternal-lineage interaction indicates that the direction of a cross (Appalachian mother or Western mother) influences the amount of reproductive isolation.

I then examined reproductive isolation at each life stage by comparing hybrid performance to the performance of parental-type plants. Field germination/seedling survival, growth-chamber germination, combined growth chamber (seedling) and field (adult) survival, and field fruit production of surviving plants were analyzed separately using a generalized linear mixed model. Model effects were the same as for cumulative fitness, but the random effect of block was included for field germination/seedling survival and fruit production to account for differences between field plots. Models for growth-chamber germination and survival used binary distributions (logit link), the model for field fruit production used a lognormal distribution (identity link), and the model for field germination/seedling survival used a binomial distribution (logit link). To explore the source of significant cross-type*region interactions, I ran separate models using the same structure but excluding one region from each; only cross-type*region is reported for these models. I performed Tukey tests among the means for the cross-type*maternal-lineage interaction for survival and field germination to identify which treatment combinations differed from each other.

Finally, to explore the relative effects of separate life stages to overall reproductive isolation, I calculated each life stage's absolute contribution to reproductive isolation. For each life stage, I calculated reproductive isolation (RI) using the equation described by Sobel and Chen (Sobel and Chen 2014):

$$RI = 1 - 2 \times (Hybrid / (Parent + Hybrid))$$

Calculated in this way, positive values indicate reproductive isolation (bounded by 1), and negative values indicate heterosis (bounded by -1). I then used the method described by Ramsey

and colleagues to calculate the absolute contribution (AC) of each life stage to cumulative reproductive isolation (Ramsey et al. 2003):

$$AC_{Germination} = RI_{Germination}$$

$$AC_{Survival} = RI_{Survival} \times (1 - AC_{Germination})$$

$$AC_{Fruit\ Production} = RI_{Fruit\ Production} \times (1 - (AC_{Germination} + AC_{Survival}))$$

Reproductive isolation at later life stages can only influence the plants that made it through earlier life stages, so the contributions of later life stages to overall reproductive isolation is modified by earlier life stages.

I then evaluated the possible contribution of biparental chloroplast inheritance to these results by exploring leaf color of the hybrids. First, I determined whether maternal lineage contributed to variation in seedling leaf color in the hybrids. I combined all regions for this analysis because region did not affect hybrid leaf color. I performed a chi-square test of independence to determine whether maternal lineage (Appalachian or Western) affected how many hybrids of each leaf color phenotype (green, variegated, light, and bleached) were produced. Next, to test the effect of hybrid seedling leaf color on survival and fruit production, I performed generalized linear models. The maternal lineage, seedling leaf color, and the interaction between the two were included as fixed effects. The maternal and paternal populations were included as random effects. All bleached-phenotype seedlings died (see Results) and therefore were excluded from the analysis.

Results

Reproductive isolation for cumulative fitness varied between lineages and among regions (cross type*maternal lineage and cross type*region, Table 1A). Hybrids with Western mothers

performed ~80-97% worse than parental populations, while hybrids with Appalachian mothers performed 78% worse to 26% better than their parents (Fig. 2). Reproductive isolation was generally smaller in contact zones than in allopatry, and was much smaller in the North Carolina contact zone than in the Pennsylvania contact zone (Fig. 2). Hybrids from allopatry and Pennsylvania always performed at least 50% worse than their parents no matter the lineage of their mother. However, Appalachian-mother hybrids from North Carolina demonstrated no reproductive isolation and performed 26% better than their parents (Fig. 2).

Few hybrids from Western mothers survived (cross type*maternal lineage, Table 1A, Fig. 3). Parental plants of both lineages survived equally well and hybrids from Appalachian mothers survived at levels 85% as well as their parents (Fig. 3A). Fewer hybrids with Western-mothers survived than with Appalachian-mothers (~15% as many). Similarly, fewer hybrids with Western mothers germinated and survived as seedlings in the field than those with Appalachian mothers (Table 1A, Fig. 3B).

For growth-chamber germination, fruit production, and field germination/seedling survival, hybrids from allopatric and Pennsylvania populations performed worse than parents, but hybrids from populations from the North Carolina contact zone did not (cross type*region, Table 1, Fig. 4). Hybrids from North Carolina had similar germination rates as their parents in a controlled environment whereas those from Pennsylvania and, to a greater extent, allopatry performed worse than their parents (Table 1A, B, Fig. 4A). In nature, North Carolina hybrids again germinated in similar proportions as their parents (Table 1B, Fig. 4B). However, as in controlled environments, germination rate of allopatric and Pennsylvania hybrids in nature was worse than their parents, with Pennsylvania hybrids germinating at an intermediate rate, higher than allopatric hybrids but lower than parents (Table 1B, Fig. 4B). For fruit production, the

pattern of reproductive isolation was the same for Pennsylvania and allopatry (Table 1B); hybrids from those regions produced fewer fruits than parent-type plants (Fig. 4C). In contrast, hybrids from North Carolina produced more fruit than their parents (Table 1B; Fig. 4C).

The sum of contributions to reproductive isolation across life stages shows patterns of overall reproductive isolation consistent with ASTER models of cumulative fitness (Fig. 2, 5). Survival contributed the most to reproductive isolation in hybrids with Western mothers. Germination was important to reproductive isolation in allopatry and Pennsylvania for both maternal lineages. In North Carolina hybrids with Appalachian mothers, high fruit production resulted in heterosis. When both crossing directions are combined, reproductive isolation was greatest in allopatry, intermediate in Pennsylvania, and lowest in North Carolina.

The distribution of F1 leaf color differed between hybrids that have Appalachian and Western mothers ($\chi^2 = 635.18$, df 3, $p<0.0001$; Fig. 6A, Fig. S1). Hybrids with an Appalachian mother were typically green (>80%) while hybrids with a Western mother were often bleached (43%). The remaining offspring of Western mothers were light (23%) or variegated (28%) but rarely green (5%) while those of from Appalachian mothers are only occasionally light or variegated. No bleached individuals survived beyond the cotyledon stage. Survival of light and variegated hybrids was intermediate to bleached and green seedlings, with a reduction of at least 50% relative to green seedlings (Table 2, Fig. 6B). Similarly, light and variegated hybrids produced fewer fruits than green hybrids (Table 2, Fig. 6C).

Discussion

In this study, I compared performance of hybrids and parent-type plants in a natural setting across two contact zones and allopatry in *Campanula americana* to explore the dynamics of postzygotic reproductive isolation at secondary contact zones. Postzygotic reproductive isolation was generally smaller in contact zones than in allopatry (Table 1A, Fig. 2, 4, & 5). The North Carolina contact zone had much lower reproductive isolation than allopatry or the Pennsylvania contact zone (Table 1B, Fig. 2, 4, & 5). The lower reproductive isolation in North Carolina relative to allopatry was driven by a lack of reproductive isolation due to germination (Fig. 4A & 4B) and heterosis due to increased fruit production in Appalachian-mother hybrids relative to parents (Fig. 4C). These reproductive incompatibilities have nuclear-nuclear genetic architectures (Barnard-Kubow et al. 2016). In contrast, reproductive isolation due to survival persists across the range of *C. americana* (Table 1A), and is underlain by a cytonuclear incompatibility (Barnard-Kubow et al. 2016). In total, reproductive isolation at contact zones between lineages of *C. americana* varies between contact zones and among incompatibilities with different underlying genetic architecture.

Cytonuclear incompatibility between lineages persists across the range

Hybrids with western mothers had much lower survival than those with Appalachian mothers, creating strong asymmetric reproductive isolation that is constant across all regions (Table 1A, Fig. 3). This asymmetrical reproductive isolation is consistent with a cytonuclear genetic architecture underlying incompatibility in *C. americana* (Barnard-Kubow et al. 2016). The chloroplast genome of the Western lineage is incompatible with Appalachian nuclear DNA, resulting in bleached leaves and reduced hybrid survival (Barnard-Kubow and Galloway 2017). Since chloroplasts are mainly inherited from the mother, this causes hybrids with Western

mothers to die much more often than hybrids with Appalachian mothers. Asymmetrical barriers caused by cytonuclear incompatibilities are common during plant speciation (Greiner et al. 2011). They can arise rapidly in allopatry (Martin et al. 2017), often via changes to regulatory networks as cytoplasmic and nuclear genomes co-adapt (Johnson 2010; Greiner and Bock 2013; Case et al. 2016). The Campanulaceae has a chloroplast genome that evolves rapidly relative to other plant families (Knox 2014; Barnard-Kubow et al. 2014), which may make this family especially predisposed to evolve cytonuclear incompatibilities.

I found that genetic architecture of incompatibilities is important to the persistence of each reproductive barrier at contact zones. The cytonuclear incompatibility creating reproductive isolation for survival was strong in all regions, while the nuclear-nuclear incompatibilities involved in reduced germination and fruit production were absent in at least one contact zone. This is concordant with theoretical modeling, which suggests that cytonuclear barriers to gene flow are often stronger than nuclear-nuclear barriers, especially when gene flow is male-biased as is often the case in plants (Höllinger and Hermisson 2017). These models show that cytonuclear barriers can be maintained at higher levels of gene flow than nuclear-nuclear barriers, suggesting that gene flow in North Carolina may be high enough to remove nuclear-nuclear barriers, but not so high as to remove cytonuclear incompatibilities. My study corroborates theory that nuclear-nuclear barriers are lost more easily than cytonuclear barriers.

While plants usually inherit chloroplasts from their mothers, in many species including *C. americana,* plants can inherit chloroplasts from either or both parents (Corriveau and Coleman 1988; Zhang et al. 2003; Snijder et al. 2007; Li et al. 2013; McCauley 2013; Barnard-Kubow et al. 2017). In *C. americana*, a paternally-inherited Appalachian chloroplast can allow a hybrid to survive despite having a Western mother, which would typically create hybrids with low survival

(Barnard-Kubow et al. 2017). Hybrids containing both Western and Appalachian chloroplasts, gained through biparental chloroplast inheritance, displayed intermediate leaf color phenotypes, with either pale green leaves or variegated green and white leaves (Fig. 6A & S1). Plants with intermediate phenotypes also had intermediate levels of fitness relative to green and bleached plants (Fig 6B & 6C). In a previous study, 58% of variegated seedlings had only Appalachian chloroplasts by the time they flowered via a process called vegetative sorting, as the more functional Appalachian chloroplast became more common as the plants grew (Barnard-Kubow et al. 2017). Since the fitness of intermediate phenotype plants in my study was similar to those that underwent vegetative sorting (Fig. 6), it is likely that the intermediate phenotype plants in this study contained only the Appalachian chloroplast when they flowered. Therefore, if Western-mother F1 hybrids survive in the field, they will likely pass on the Appalachian chloroplast to their offspring.

Asymmetric reproductive isolation due to survival was uniform across both contact zones and allopatry. Therefore, it appears that the alleles contributing to cytonuclear incompatibility are retained in contact zones despite gene flow (Chapter 2). Hybrids usually survived when Appalachian plants are their mothers, but they survived only rarely when Western plants served as mothers (Fig. 3A). A common outcome of cytonuclear incompatibilities between lineages is asymmetrical gene flow (Tiffin et al. 2001; Greiner and Bock 2013; Ley and Hardy 2017); cytonuclear incompatibility in *C. americana* would be expected to result in gene flow that proceeds from Western populations into those with Appalachian chloroplasts. In *C. americana*, seeds passively disperse by gravity and therefore seed dispersal is limited (Galloway 2005), so most gene movement is expected via pollen. Since viable hybrids are produced on Appalachian mothers, gene flow will likely proceed asymmetrically via pollen from Western populations into

Appalachian populations. Alternatively, in the rare event that a hybrid from a Western-mother survives due to inheritance of its father's chloroplast, it may pass on that Appalachian chloroplast to its offspring within a Western population. If such a plant can establish and backcross in the Western population, producing high-fitness hybrids, its offspring may retain the Appalachian chloroplast, but with hybrid or Western-like nuclear DNA after subsequent backcrosses. In other words, rare biparental chloroplast inheritance events may facilitate the spread of the Appalachian lineage's chloroplast within Western lineage populations (Rieseberg and Soltis 1991). Thus, the cytonuclear incompatibility and biparental chloroplast inheritance may allow Western nuclear genes to flow into Appalachian populations, and, although rarely, Appalachian chloroplasts to flow into Western populations.

Reproductive isolation varies among contact zones

Reproductive isolation differed between contact zones and allopatry; isolation was highest in allopatry, intermediate in Pennsylvania, and lowest in North Carolina (Fig. 2, 4, & 5). While germination contributed strongly to isolation in allopatry, its effects in Pennsylvania were less dramatic, and hybrids from North Carolina germinated as often as parent-type seeds (Fig. 4A & 4B). Similarly, while fruit production contributed to isolation in allopatry and Pennsylvania, North Carolina hybrids produced more fruit than parents (Fig. 4C). Since hybrids with Appalachian mothers did not experience survival breakdown (Fig. 3), Appalachian-mother hybrids from North Carolina had higher overall fitness than their parents even though the cytonuclear incompatibility persists across the range (Fig 1 & 5). The absence of reproductive isolation due to germination and fruit production in North Carolina suggests that the lineages are merging together in that contact zone.

It is surprising that hybrids with Appalachian mothers from the North Carolina contact zone have higher fitness than their parents because it means that in one crossing direction and contact zone hybridization is favorable (Fig. 2 & 5). Parental populations may have differentiated in part by drift, and therefore gaining genetic load as mildly deleterious alleles increased in frequency in isolation (Fenster and Galloway 2000). These alleles are masked when heterozygous, resulting in heterosis when plants from different populations cross. Past work in *C. americana* has found high rates of inbreeding depression in this species, supporting the presence of genetic load in these populations (Galloway and Etterson 2007). In the absence of reproductive isolating barriers, hybrids can perform better than parents because hybridization masks the presence of genetic load, which was likely the cause of hybrid vigor in North Carolina hybrids with Appalachian mothers.

High hybrid fitness in one crossing direction in North Carolina suggests that gene flow may proceed quickly and asymmetrically there. The small reduction in reproductive isolation in Pennsylvania relative to allopatry may have been due to some gene exchange there, although the magnitude is lower than in North Carolina. Higher hybrid fitness in contact zones relative to allopatry is evidence of gene flow; as lineages interbreed, incompatibilities are purged (Felsenstein 1981; Barton and Bengtsson 1986; Lindtke and Buerkle 2015). Increased gene flow and decreased reproductive isolation can feed back on each other to allow the lineages to interbreed more and more easily over time (Todesco et al. 2016). As the alleles responsible for reproductive isolation are removed from the populations in contact, gene flow proceeds more easily and in turn provides the opportunity for selection to act on those alleles and remove them from the populations (Taylor et al. 2005; Seehausen et al. 2008; Garrick et al. 2014). Since postzygotic reproductive isolation was low in North Carolina relative to allopatry, it is likely that

gene flow in that contact zone has allowed selection to remove at least some of the incompatibilities there.

The reproductive incompatibilities separating the lineages of *C. americana* are intrinsic (Barnard-Kubow and Galloway 2017) and therefore might be expected to be uniform across ecological contexts, resulting in uniform responses to secondary contact (Coyne and Orr 2004; Kingston et al. 2017). However, intrinsic isolating barriers may vary across the range due to genetic variation in the alleles that underlie the barriers (Cutter 2012; Corbett-Detig et al. 2013). The contact zones each showed different patterns of reproductive isolation relative to allopatry (Fig 2 & 5). One explanation for this is that reproductive isolation may have been stronger at initial contact in Pennsylvania than in North Carolina due to range-wide genetic structure. Alternatively, extrinsic factors that differ between Pennsylvania and North Carolina may have modified the ability of the lineages to merge (Mandeville et al. 2015; Cuter and Gray 2016).

Several differences between the contact zones could explain the difference in reproductive isolation between them. The simplest explanation is time in contact. The Western lineage of *C. americana* migrated northward from refugia near the Gulf Coast following the Last Glacial Maximum (Barnard-Kubow et al. 2015). Therefore, it would have come into contact with the Appalachian lineage in North Carolina much earlier than in Pennsylvania. Lower reproductive isolation in North Carolina may indicate that these populations have had more time to purge incompatible alleles. Also, as the Western lineage migrated north, it may have accumulated additional incompatibilities with the Appalachian lineage, increasing reproductive isolation at the time of initial contact and thereby making purging all incompatibilities less likely. Reproductive isolating barriers accumulate rapidly in allopatry after initial barriers form (Moyle and Nakazato 2010; Matute et al. 2010), making this explanation plausible. Finally, the

North Carolina contact zone is in the center of highly suitable habitat (Barnard-Kubow et al. 2015), where there may be many populations and lots of opportunity for gene flow. In contrast, the Pennsylvania contact zone is in less suitable habitat on the edge of the range, where populations may be sparser and opportunities for hybridization rare. These differences may lead to a lower rate of migration between lineages in Pennsylvania than in North Carolina. With a reduction in gene flow between lineages in Pennsylvania, rather than merging, reinforcement may occur (Chapter 3), maintaining postzygotic reproductive isolation and increasing prezygotic isolating mechanisms (Noor 1999; Hopkins and Rausher 2012). Additional work is necessary to determine the underlying causes of differences between the contact zones.

Conclusion

When partially isolated lineages come into secondary contact before speciation is complete, lineages may collapse, or they may instead facilitate speciation via reinforcement. This dichotomy of outcomes means that secondary contact is a key stage during the speciation process. In the North Carolina contact zone between partially-isolated lineages of *C. americana*, postzygotic reproductive isolation is lower than in allopatry, indicating that the lineages are merging after contact. However, postzygotic reproductive isolation in the Pennsylvania contact zone between the same lineages is almost as high as in allopatry. Variation in reproductive isolation between contact zones demonstrates that many factors influence the outcome of secondary contact. Genetic architecture of reproductive barriers is particularly important; cytonuclear incompatibilities persist even when nuclear-nuclear incompatibilities are purged. In addition, the contrast between high reproductive isolation in Pennsylvania and low reproductive isolation in North Carolina demonstrates that lineages may not respond consistently to secondary

contact across their ranges. The differences between the contact zones in *C. americana* show that variation within species can dramatically affect the outcome of secondary contact, and therefore the trajectory of speciation.

Acknowledgements

I would like to thank Rebecca and Jeffrey Risser for assistance planting the common garden, Karen Barnard-Kubow for collecting seed for *C. americana* populations, and Mountain Lake Biological Station for logistical support. I would also like to thank the Galloway lab for comments on and discussion of this manuscript. This work was supported by an NSF GRFP and by NSF DEB-1457686.

Table 1. ANOVA to evaluate reproductive isolation in fitness components and cumulative fitness between *C. americana* populations from different genetic lineages sampled from contact zones and allopatry. (A) The effects of crossing between or within lineages on fitness components assessed by ANOVA, and cumulative fitness assessed by ASTER modeling. (B) Interactions between cross type and region for pairs of regions calculated using models with the same structure as A. Crosses were conducted between populations from different lineages and within populations (cross type) in two contact zones and in allopatry (region). Both Appalachian and Western populations served as mothers in crosses (maternal lineage). F-values shown; maternal population and paternal population were included in the model as random effects for all analyses, and block as a random effect for field germination and fruit production (not shown). + p<0.1, * p<0.05, ** p<0.01, *** p<0.001.

A Effect	Num DF	Growth-Chamber Germination	Survival	Fruit Production	Field Germination /Seedling survival	Aster df	Cumulative Fitness
Cross Type	1	130.28***	172.35***	2.66	9.63**	7	323.20***
Region[a]	2	2.89+	1.91	12.93***	16.47***	8	101.00***
Maternal Lineage	1	1.40	70.26***	6.74**	2.40	8	309.24***
Cross Type*Region	2	31.50***	1.15	3.31*	4.55*	2	64.21***
Region*Maternal Lineage	2	2.05	0.29	0.44	1.59	2	2.44
Cross Type*Maternal Lineage	1	2.71	57.06***	0.05	12.36***	1	17.90***
Cross Type*Region*Maternal Lineage	2	0.06	0.05	0.11	3.12*	2	4.01
Den DF		2343	1131	525	800		

a. Den DF Region: 15 GC Germ, Survival and Fruit Production; 20 Field Germination

B Cross Type*Region: pairwise comparison	Num DF	Growth-Chamber Germination	Fruit Production	Field Germination /Seedling survival
Allopatry vs PA	1	9.65**	0.15	3.42+
Allopatry vs NC	1	60.80***	3.08+	6.28*
PA vs NC	1	26.90***	5.32*	0.40
Den DF		1524-1633	295-382	482-581

Table 2. The effects of seedling leaf color phenotype and maternal lineage on survival and fruit production, assessed by ANOVA. Crosses were conducted between populations from different lineages, and both Appalachian and Western lineage populations served as mothers in crosses (maternal lineage). Offspring of these crosses had bleached, light, variegated, or green leaves (leaf color). F-values shown; maternal population and paternal population were included in the model as random effects (not shown).
* $p<0.05$, *** $p<0.001$.

	Num DF	Survival	Fruit Production
Maternal lineage	1	26.07***	0.05
Leaf color	2	24.77***	4.33*
Maternal lineage*Leaf color	2	1.81	0.32
Den DF		510	209

Figure 1. Locations of *Campanula americana* populations (A), population crossing scheme used in this study (B), and range of each lineage of *C. americana* (C). Appalachian lineage populations are marked in green, and Western lineage populations in purple. North Carolina and Pennsylvania contact zones are designated with boxes, and all other populations are classified as allopatric. The common garden location is denoted by a yellow star.

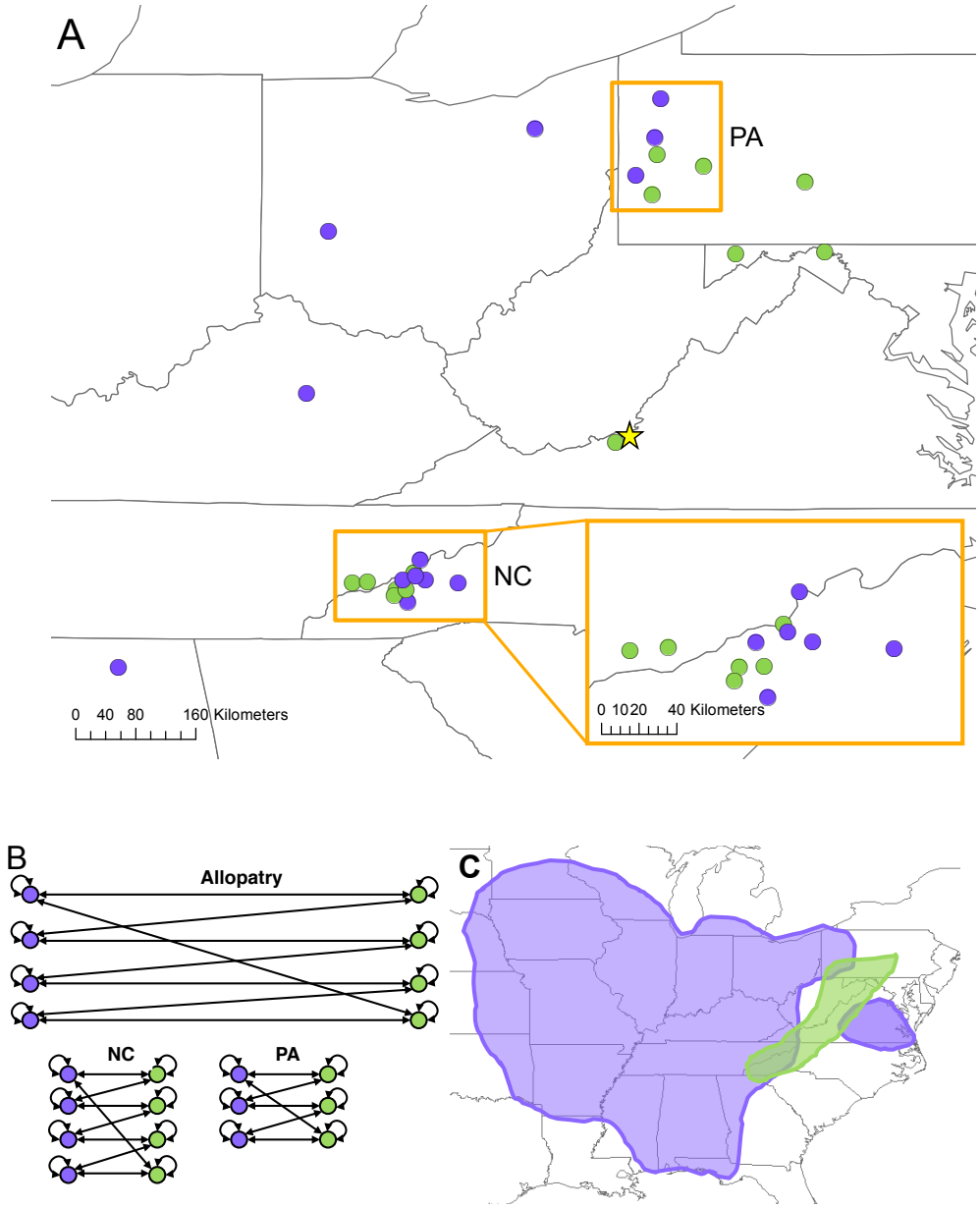


Figure 2. Least square means of cumulative fitness of between-lineage hybrids and parental populations evaluated for crosses with Appalachian-lineage mothers and Western-lineage mothers from contact zones in Pennsylvania (PA) and North Carolina (NC), as well as populations in allopatry. Green indicates plants with Appalachian-lineage mothers, and purple indicates plants with Western-lineage mothers. Error bars represent one standard error.

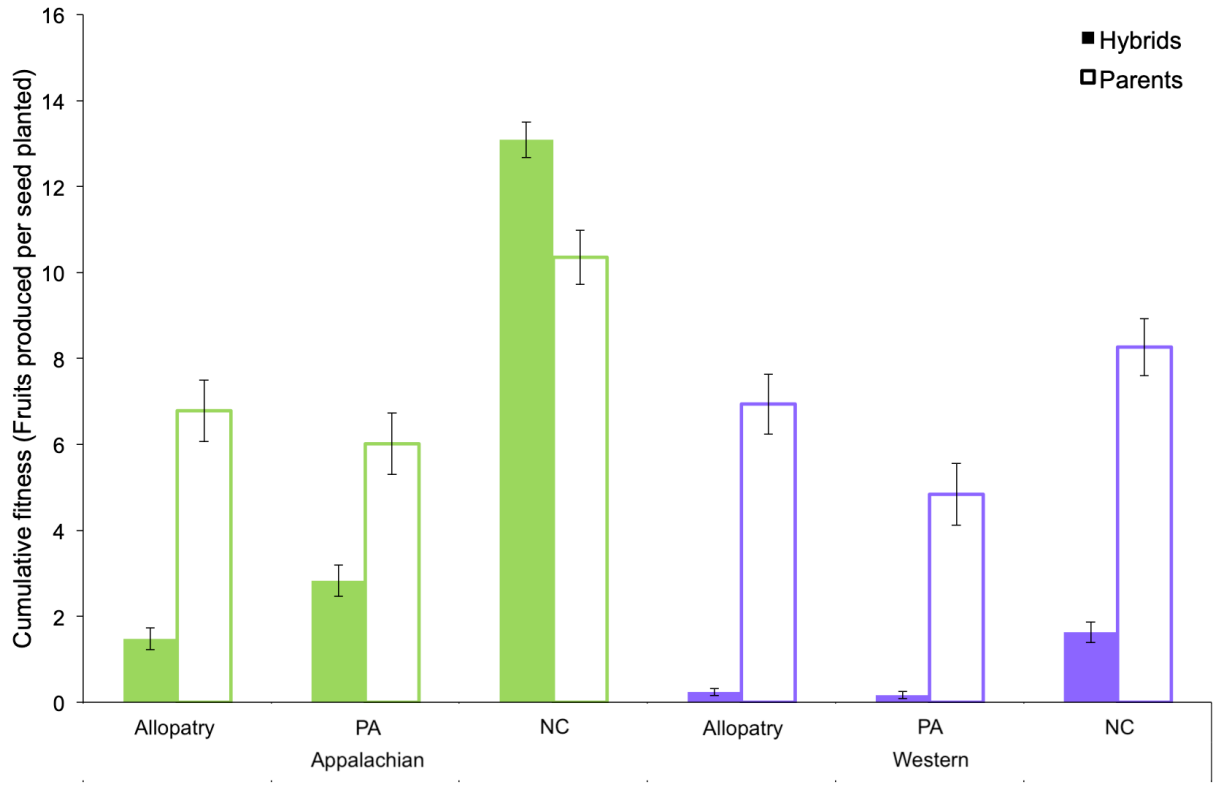


Figure 3. Least square means of survival (A) and germination and seedling survival in the field (B) of between-lineage hybrids and parental populations for crosses with Appalachian-lineage mothers and Western-lineage mothers. Green represents crosses with Appalachian mothers, and purple crosses with Western mothers. Error bars represent one standard error. Bars different at alpha = 0.05 are marked with different letters.

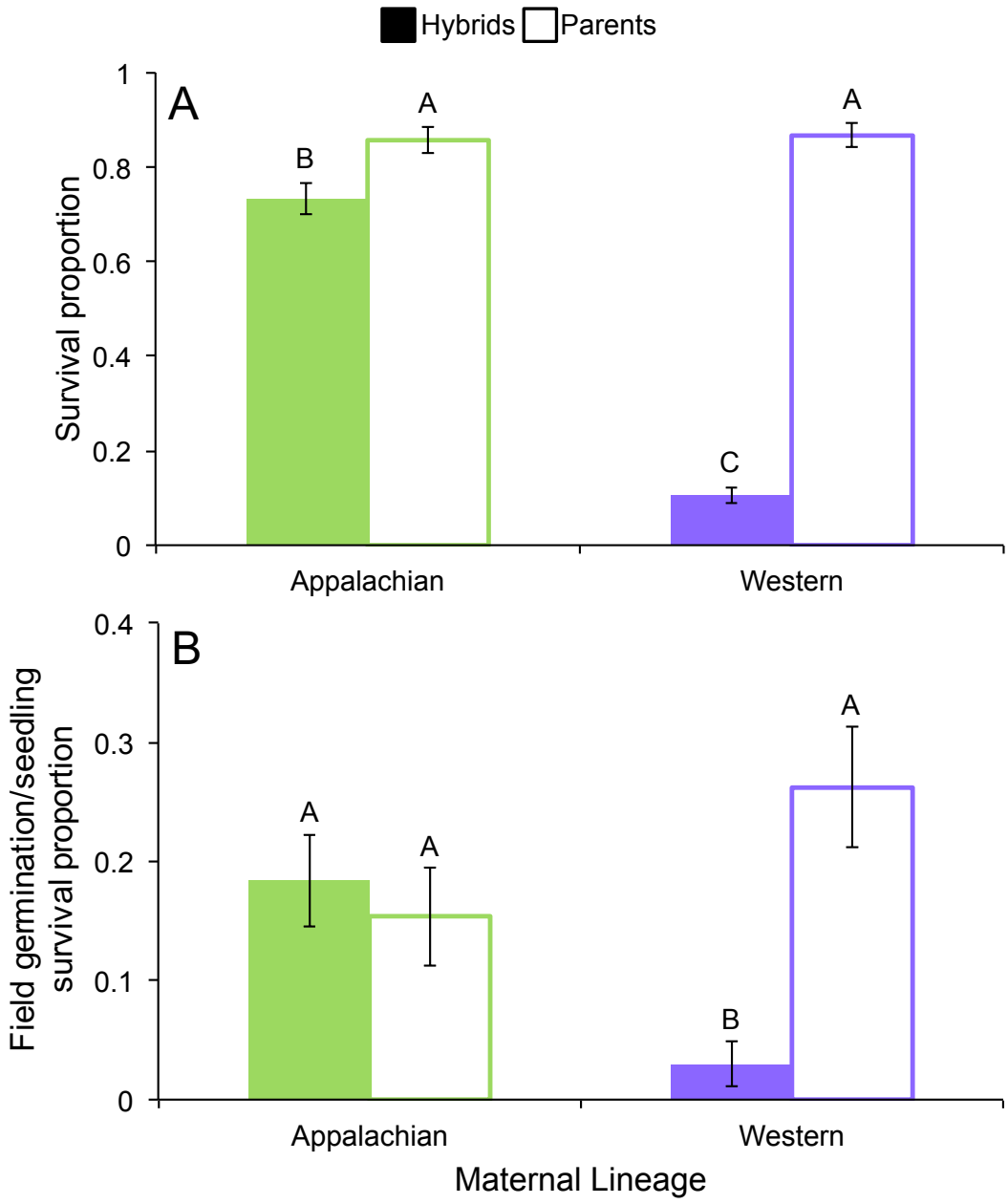


Figure 4. Least square means of performance traits of between-lineage hybrids and parental populations from contact zones in Pennsylvania (PA) and North Carolina (NC), as well as allopatry. Traits measured are germination in the growth chamber (A), germination and seedling survival in the field (B), and field fruit production (C). Error bars represent one standard error.

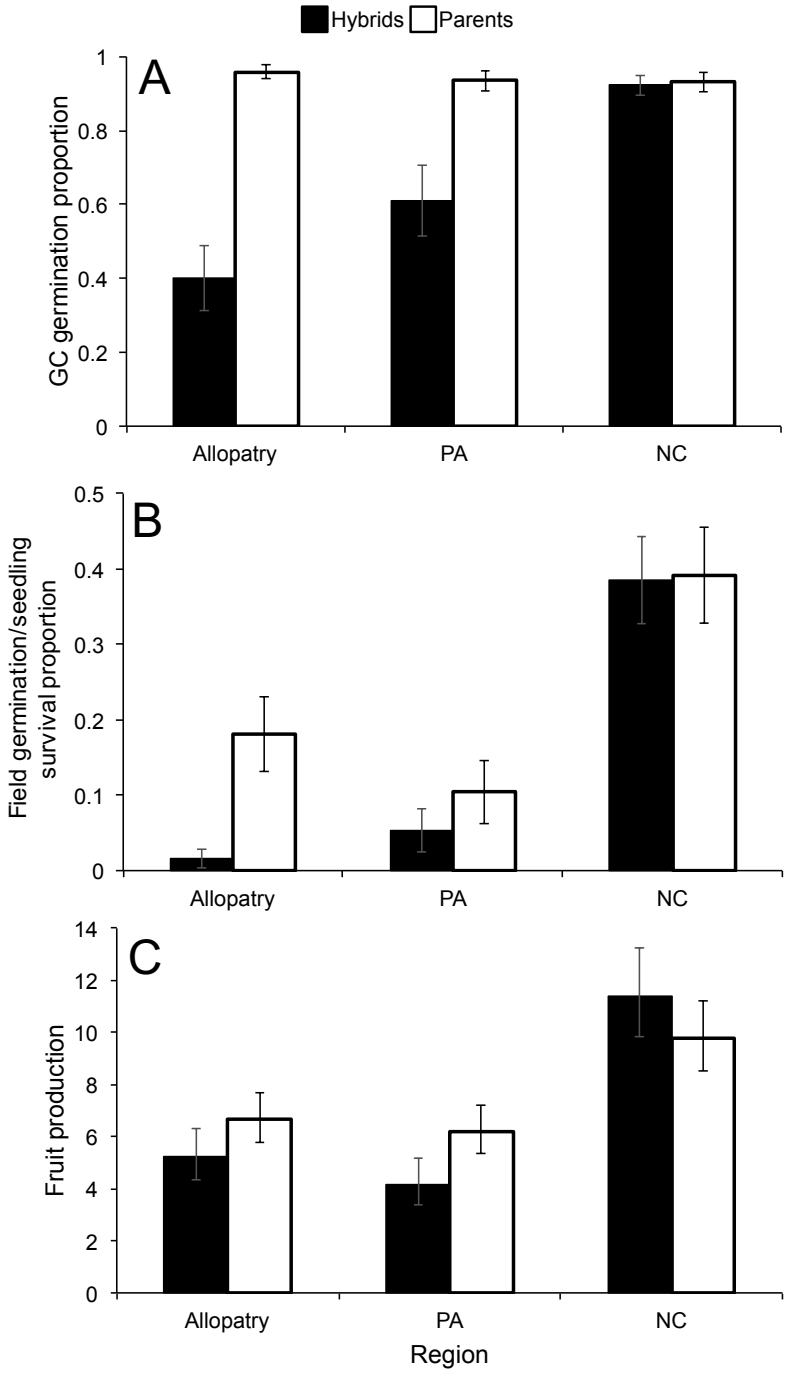


Figure 5. Absolute contributions to reproductive isolation between Appalachian and Western lineages of *Campanula americana* evaluated for populations from contact zones in Pennsylvania (PA) and North Carolina (NC) as well as allopatry. Positive values indicate variation in these traits contributes to reproductive isolation, and negative values indicate heterosis. Bars outlined in green represent crosses with Appalachian mothers, and bars outlined in purple represent crosses with Western mothers. Overall reproductive isolation (total bar height) is the sum of traits throughout the life cycle.

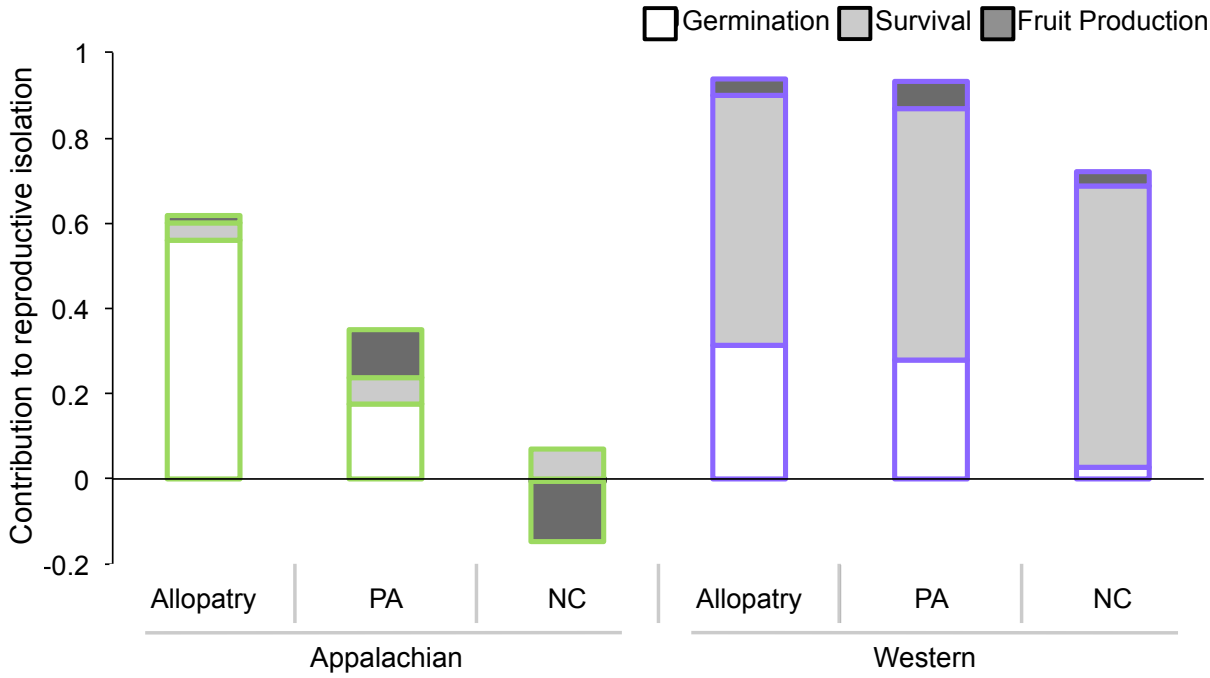


Figure 6. Number of plants (A) and least square means of survival (B) and fruit production (C) of hybrids with each leaf color phenotype. Hybrid phenotype was scored as bleached (white-leaved), light (pale green), variegated (green and white tissue in distinct patches), and green (typical dark green color). Bars outlined in green represent Appalachian-mother crosses, and bars outlined in purple represent Western-mother crosses.

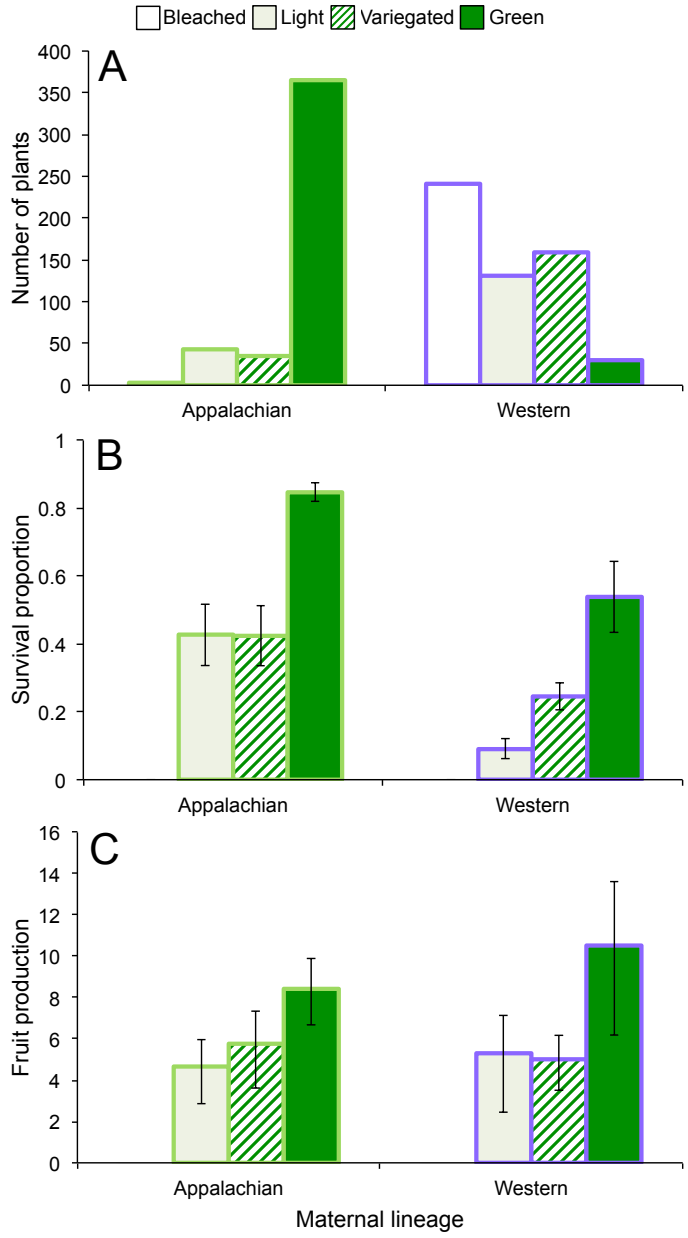


Table S1. Populations of *Campanula americana* were sampled from three regions, including allopatry and contact zones in North Carolina and Pennsylvania. Populations from both the Appalachian and Western lineages were included from each zone.

Population	Lineage	Region	Latitude	Longitude
MD5	Appalachian	Allopatry	39.614	-79.116
PA95	Appalachian	Allopatry	40.475	-78.281
VA73	Appalachian	Allopatry	37.353	-80.552
WV98	Appalachian	Allopatry	39.632	-78.043
AL_BG	Western	Allopatry	34.656	-86.517
KY51	Western	Allopatry	37.934	-84.259
OH119	Western	Allopatry	39.885	-83.997
OH64	Western	Allopatry	41.115	-81.518
NC109E	Appalachian	North Carolina	35.787	-82.973
NC110	Appalachian	North Carolina	35.582	-83.186
NC130	Appalachian	North Carolina	35.516	-83.210
NC91	Appalachian	North Carolina	35.586	-83.066
TN113	Appalachian	North Carolina	35.660	-83.710
TN92	Appalachian	North Carolina	35.676	-83.526
NC105	Western	North Carolina	35.703	-82.833
NC106	Western	North Carolina	35.667	-82.443
NC107	Western	North Carolina	35.943	-82.895
NC108	Western	North Carolina	35.701	-83.106
NC109A	Western	North Carolina	35.748	-82.954
NC114	Western	North Carolina	35.436	-83.048
PA101	Appalachian	Pennsylvania	40.664	-79.501
PA102	Appalachian	Pennsylvania	40.322	-80.111
PA104	Appalachian	Pennsylvania	40.802	-80.055
PA103	Western	Pennsylvania	40.550	-80.311
PA27	Western	Pennsylvania	41.008	-80.083
PA94	Western	Pennsylvania	41.467	-80.011

Table S2. Pairs of populations crossed to produce reciprocal F1 hybrids from each of three regions.

Region	Appalachian Population	Western Population
Allopatry	MD5	KY51
	MD5	OH64
	PA95	AL_BG
	PA95	OH119
	VA73	AL_BG
	VA73	KY51
	WV98	OH64
	WV98	OH119
North Carolina	NC109E	NC109A
	NC109E	NC107
	NC110	NC108
	NC110	NC114
	NC130	NC105
	NC130	NC107
	NC91	NC106
	NC91	NC109A
	TN113	NC106
	TN113	NC108
	TN92	NC105
	TN92	NC114
Pennsylvania	PA101	PA27
	PA101	PA94
	PA102	PA103
	PA102	PA94
	PA104	PA103
	PA104	PA27

Figure S1. Seedling leaf colors including (A) bleached, (B) light, (C) variegated, and (D) green (typical) phenotypes.

Citations

Alcaide, M., E. S. C. Scordato, T. D. Price, and D. E. Irwin. 2014. Genomic divergence in a ring species complex. Nature 511:83–85.

Bank, C., R. Bürger, and J. Hermisson. 2012. The limits to parapatric speciation: Dobzhansky–Muller incompatibilities in a continent–island model. Genetics 191:845–863.

Barnard-Kubow, K. B., and L. F. Galloway. 2017. Variation in reproductive isolation across a species range. Ecol Evol 7:9347–9357.

Barnard-Kubow, K. B., C. L. Debban, and L. F. Galloway. 2015. Multiple glacial refugia lead to genetic structuring and the potential for reproductive isolation in a herbaceous plant. American Journal of Botany 102:1842–1853.

Barnard-Kubow, K. B., D. B. Sloan, and L. F. Galloway. 2014. Correlation between sequence divergence and polymorphism reveals similar evolutionary mechanisms acting across multiple timescales in a rapidly evolving plastid genome. BMC Evol Biol 14.

Barnard-Kubow, K. B., M. A. McCoy, and L. F. Galloway. 2017. Biparental chloroplast inheritance leads to rescue from cytonuclear incompatibility. New Phytol 213:1466–1476.

Barnard-Kubow, K. B., N. So, and L. F. Galloway. 2016. Cytonuclear incompatibility contributes to the early stages of speciation. Evolution 70:2752–2766.

Barton, N. H. 1979. The dynamics of hybrid zones. Heredity 43:341–359.

Barton, N., and B. O. Bengtsson. 1986. The barrier to genetic exchange between hybridising populations. Heredity 56:357–376.

Baskin, J. M., and C. C. Baskin. 1984. The ecological life cycle of *Campanula americana* in northcentral Kentucky. Bulletin of the Torrey Botanical Club 111:329–337.

Bateson, W. 1909. Heredity and variation in modern lights. Pp. 85–101 *in* A. C. Seward, ed. Darwin and Modern Science. Cambridge, MA.

Cain, M. L., V. Andreasen, and D. J. Howard. 1999. Reinforcing selection is effective under a relatively broad set of conditions in a mosaic hybrid zone. Evolution 53:1343–1353.

Case, A. L., F. R. Finseth, C. M. Barr, and L. Fishman. 2016. Selfish evolution of cytonuclear hybrid incompatibility in *Mimulus*. Proceedings of the Royal Society B: Biological Sciences 283:20161493–9.

Corbett-Detig, R. B., J. Zhou, A. G. Clark, D. L. Hartl, and J. F. Ayroles. 2013. Genetic incompatibilities are widespread within species. Nature 504:135–137.

Corriveau, J. L., and A. W. Coleman. 1988. Rapid screening method to detect potential biparental inheritance of plastid DNA and results for over 200 angiosperm species. American

Journal of Botany 75:1443–1458.

Coyne, J. A., and H. A. Orr. 2004. Speciation. Sinauer Associates, Inc., Sunderland, MA, USA.

Cutter, A. D. 2012. The polymorphic prelude to Bateson–Dobzhansky–Muller incompatibilities. TREE 27:210–219.

Cutter, A. D., and J. C. Gray. 2016. Ephemeral ecological speciation and the latitudinal biodiversity gradient. Evolution 70:2171–2185.

Dobzhansky, T. 1937. Genetics and the origin of species. Columbia University Press, New York.

Evanhoe, L., and L. F. Galloway. 2002. Floral longevity in *Campanula americana* (Campanulaceae): A comparison of morphological and functional gender phases. American Journal of Botany 89:587–591.

Felsenstein, J. 1981. Skepticism towards Santa Rosalia, or Why are there so few kinds of animals? Evolution 35:124–138.

Fenster, C. B., and L. F. Galloway. 2000. Population differentiation in an annual legume: genetic architecture. Evolution 54:1157–1172.

Gadella, T. W. J. 1964. Cytotaxonomic studies in the genus *Campanula*. Wentia 11:1–104.

Galloway, L. F. 2005. Maternal effects provide phenotypic adaptation to local environmental conditions. New Phytol 166:93–100.

Galloway, L. F., and J. R. Etterson. 2007. Inbreeding depression in an autotetraploid herb: a three cohort field study. New Phytol 173:383–392.

Galloway, L. F., and J. R. Etterson. 2005. Population differentiation and hybrid success in *Campanula americana*: geography and genome size. J. Evol. Biol. 18:81–89.

Galloway, L. F., J. R. Etterson, and J. L. Hamrick. 2003. Outcrossing rate and inbreeding depression in the herbaceous autotetraploid, *Campanula americana*. Heredity 90:308–315.

Garrick, R. C., E. Benavides, M. A. Russello, C. Hyseni, D. L. Edwards, J. P. Gibbs, W. Tapia, C. Ciofi, and A. Caccone. 2014. Lineage fusion in Galápagos giant tortoises. Molecular Ecology 23:5276–5290.

Geyer, C. T., S. Wagenius, and R. G. Shaw. 2007. Aster models for life history analysis. Biometrika 94:415–426.

Greiner, S., and R. Bock. 2013. Tuning a ménage à trois: Co-evolution and co-adaptation of nuclear and organellar genomes in plants. Bioessays 35:354–365.

Greiner, S., U. Rauwolf, J. Meurer, and R. G. Herrmann. 2011. The role of plastids in plant speciation. Molecular Ecology 20:671–691.

Hopkins, R. 2013. Reinforcement in plants. New Phytol 197:1095–1103.

Hopkins, R., and M. D. Rausher. 2012. Pollinator-mediated selection on flower color allele drives reinforcement. Science 335:1090–1092.

Höllinger, I., and J. Hermisson. 2017. Bounds to parapatric speciation: A Dobzhansky-Muller incompatibility model involving autosomes, X chromosomes, and mitochondria. Evolution 71:1366–1380.

Irwin, D. E., S. Bensch, and T. D. Price. 2001. Speciation in a ring. Nature 409:333–337.

Johnson, N. A. 2010. Hybrid incompatibility genes: remnants of a genomic battlefield? Trends in Genetics 26:317–325.

Kearns, A. M., M. Restani, I. Szabo, A. Schrøder-Nielsen, J. A. Kim, H. M. Richardson, J. M. Marzluff, R. C. Fleischer, A. Johnsen, and K. E. Omland. 2018. Genomic evidence of speciation reversal in ravens. Nature Communications 9:906.

Kingston, S. E., T. L. Parchman, Z. Gompert, C. A. Buerkle, and M. J. Braun. 2017. Heterogeneity and concordance in locus-specific differentiation and introgression between species of towhees. J. Evol. Biol. 30:474–485.

Knox, E. B. 2014. The dynamic history of plastid genomes in the Campanulaceae sensu lato is unique among angiosperms. PNAS 201403363.

Lepais, O., R. J. Petit, E. Guichoux, J. E. Lavabre, F. Alberto, A. Kremer, and S. Gerber. 2009. Species relative abundance and direction of introgression in oaks. Molecular Ecology 18:2228–2242.

Ley, A. C., and O. J. Hardy. 2017. Hybridization and asymmetric introgression after secondary contact in two tropical african climber species, *Haumania danckelmaniana* and *Haumania liebrechtsiana* (Marantaceae). Int. J. Plant Sci. 178:421–430.

Li, D., X. Qi, X. Li, L. Li, C. Zhong, and H. Huang. 2013. Maternal inheritance of mitochondrial genomes and complex inheritance of chloroplast genomes in *Actinidia* Lind.: evidences from interspecific crosses. Mol Genet Genomics 288:101–110.

Lindtke, D., and C. A. Buerkle. 2015. The genetic architecture of hybrid incompatibilities and their effect on barriers to introgression in secondary contact. Evolution 69:1987–2004.

Mandeville, E. G., T. L. Parchman, D. B. McDonald, and C. A. Buerkle. 2015. Highly variable reproductive isolation among pairs of *Catostomus* species. Molecular Ecology 24:1856–1872.

Martin, H., P. Touzet, M. Dufay, C. Godé, E. Schmitt, E. Lahiani, L. F. Delph, and F. Van Rossum. 2017. Lineages of *Silene nutans* developed rapid, strong, asymmetric postzygotic reproductive isolation in allopatry. Evolution 71:1519–1531.

Matute, D. R. 2010. Reinforcement can overcome gene flow during speciation in *Drosophila*.

Current Biology 20:2229–2233.

Matute, D. R., I. A. Butler, D. A. Turissini, and J. A. Coyne. 2010. A test of the snowball theory for the rate of evolution of hybrid incompatibilities. Science 329:1518–1521.

McCauley, D. E. 2013. Paternal leakage, heteroplasmy, and the evolution of plant mitochondrial genomes. New Phytol 200:966–977.

Moyle, L. C., and T. Nakazato. 2010. Hybrid incompatibility "snowballs" between Solanum species. Science 329:1521–1523.

Muller, H. J. 1942. Isolating mechanisms, evolution, and temperature. Biological Symposia 6:71-125.

Navarro, A., and N. H. Barton. 2003. Chromosomal speciation and molecular divergence—accelerated evolution in rearranged chromosomes. Science 300:321–324.

Noor, M. A. F. 1999. Reinforcement and other consequences of sympatry. Heredity 83:503–508.

Noor, M. A. F., K. L. Grams, L. A. Bertucci, and J. Reiland. 2001. Chromosomal inversions and the reproductive isolation of species. PNAS 98:12084–12088.

Ramsey, J., H. D. Bradshaw Jr., and D. W. Schemske. 2003. Components of reproductive isolation between the monkeyflowers *Mimulus lewisii* and *M. cardinalis* (Phrymaceae). Evolution 57:1520–1534.

Rieseberg, L. H., and D. E. Soltis. 1991. Phylogenetic consequences of cytoplasmic gene flow in plants. Evolutionary Trends in Plants 5:65–84.

Rieseberg, L. H., J. Whitton, and K. Gardner. 1999. Hybrid zones and the genetic architecture of a barrier to gene flow between two sunflower species. Genetics 152:713–727.

Sardell, J. M., and J. A. C. Uy. 2016. Hybridization following recent secondary contact results in asymmetric genotypic and phenotypic introgression between island species of *Myzomela* honeyeaters. Evolution 70:257–269.

Seehausen, O., G. Takimoto, D. Roy, and J. Jokela. 2008. Speciation reversal and biodiversity dynamics with hybridization in changing environments. Molecular Ecology 17:30–44.

Snijder, R. C., F. S. Brown, and J. M. van Tuyl. 2007. The role of plastome-genome incompatibility and biparental plastid inheritance in interspecific hybridization in the genus *Zantedeschia* (Araceae). Floriculture and Ornamental Biotechnology 1:150–157.

Sobel, J. M., and G. F. Chen. 2014. Unification of methods for estimating the strength of reproductive isolation. Evolution 68:1511–1522.

Taylor, E. B., J. W. Boughman, M. Groenenboom, M. Sniatynski, D. Schluter, and J. L. Gow. 2005. Speciation in reverse: morphological and genetic evidence of the collapse of a three-spined

stickleback (*Gasterosteus aculeatus*) species pair. Molecular Ecology 15:343–355.

Tiffin, P., S. Olson, and L. C. Moyle. 2001. Asymmetrical crossing barriers in angiosperms. Proceedings of the Royal Society B: Biological Sciences 268:861–867.

Todesco, M., M. A. Pascual, G. L. Owens, K. L. Ostevik, B. T. Moyers, S. Hübner, S. M. Heredia, M. A. Hahn, C. Caseys, D. G. Bock, and L. H. Rieseberg. 2016. Hybridization and extinction. Evolutionary Applications 9:892–908.

Twyford, A. D., and J. Friedman. 2015. Adaptive divergence in the monkey flower *Mimulus guttatus* maintained by a chromosomal inversion. Evolution 69:1476–1486.

Wolfe, A. D., and W. J. Elisens. 1995. Evidence of chloroplast capture and pollen-mediated gene flow in *Penstemon* sect. *Peltanthera* (Scrophulariaceae). Systematic Botany 20:395–412.

Zhang, Q., Y. Liu, Sodmergen. 2003. Examination of the Cytoplasmic DNA in Male Reproductive Cells to Determine the Potential for Cytoplasmic Inheritance in 295 Angiosperm Species. Plant Cell Physiology 44:941–951.

CHAPTER 2:

Variation in gene flow among three contact zones between partially isolated lineages

Abstract

Secondary contact between partially isolated lineages can either facilitate speciation via reinforcement or allow the lineages to collapse back into a single species. If the lineages experience reinforcement, gene flow at the contact zones will be low, and if they collapse together, gene flow will be high. Factors that determine which of these outcomes happen in a particular contact zone include ecological factors and level and type of intrinsic reproductive isolation. Both of these often vary across a species range, and therefore the outcome of secondary contact can vary among contact zones. However, few studies have explored the degree to which outcomes of secondary contact vary within a species. In this study, I test whether gene flow varies among three contact zones between the partially isolated Appalachian and Western lineages of the herb *Campanula americana*. I performed double-digest restriction-associated DNA sequencing (ddRADseq) on multiple populations of each lineage from each contact zone and from allopatry. Gene flow between lineages was high in one contact zone, but low in the other two zones. Despite a cytonuclear incompatibility present between the lineages across the range, I find little evidence of asymmetrical gene flow between the lineages. Since the contact zones differ in both initial levels of intrinsic reproductive isolation and in ecological habitat suitability, both intrinsic genetic variation and ecological variation likely contribute to variation in contact zone outcome.

Introduction

The process of speciation divides a single species into multiple new independently evolving lineages, which themselves become new species (de Queiroz 2007). In order to evolve independently, gene flow between the lineages must be very low. Low levels of gene flow are usually achieved through geographic isolation or through genetically conferred reproductive isolating barriers, or both (Coyne and Orr 2004). If a geographic barrier arises, blocking within-species gene flow, the two isolated lineages begin accumulating genetic differences which can lead to reproductive isolation (Bateson 1909; Dobzhansky 1937; Muller 1942). If the lineages come into contact before they are completely isolated by genetic mechanisms, the gene flow at those contact zones can facilitate increased divergence of the lineages via reinforcement, reducing gene flow (Dobzhansky 1937; Noor 1999; Matute 2010). Alternatively, contact can facilitate the loss of reproductive isolation and lead to high levels of gene flow, allowing the lineages to merge (Felsenstein 1981; Barton and Bengtsson 1986; Seehausen et al. 2008). These divergent outcomes make secondary contact between partially isolated lineages a key phase in the speciation process.

Many factors, including initial magnitude of reproductive isolation (Bank et al. 2011), geographic structure of contact zones (Cain et al. 1999), and genetic architecture of reproductive incompatibilities (Bank et al. 2012a; Lindtke and Buerkle 2015), can determine the outcome of secondary contact. These factors may vary across species' ranges, especially when they are ecologically driven because ecology often varies across species' ranges (Coyne and Orr 2004). However, ecology is not the only factor influencing the outcome of secondary contact that varies across a species range; degree of reproductive isolation between a pair of lineages often scales with genetic distance because the genetic basis of isolating barriers can vary across a species

range (Irwin 2005; Corbett-Detig 2013). Factors that increase the frequency of gene flow between lineages at contact zones, such as low initial reproductive isolation and high density of populations at contact zones, make it more likely that the lineages will merge. Even if there are multiple barriers to reproductive isolation, if they are not tightly linked it may be possible for them to be lost at contact zones, allowing for more gene flow between lineages (Bank et al. 2012b). Since reproductive isolation varies within species (Cutter 2012; Corbett-Detig et al. 2013; Mandeville et al. 2015), and the magnitude of reproductive isolation contributes to the outcome of secondary contact, it is likely that different contact zones between the same pair of lineages may have different outcomes. However, few studies have explored the degree to which outcomes of secondary contact vary within a species.

When partially isolated lineages meet, their response to contact is often asymmetrical. Reinforcement is often stronger in one species than in the other (Hopkins and Rausher 2012; Yukilevich 2012; Bewick and Dyer 2014). Similarly, when secondary contact facilitates gene flow between two lineages, that gene flow is often asymmetrical (Burgess et al. 2005; Field et al. 2010; Natalis and Wesselingh 2012; Hülber et al. 2015; Ley and Hardy 2017). In many cases, these asymmetrical responses to secondary contact are driven by preexisting asymmetries in reproductive isolation (Tiffin et al. 2001). Cytonuclear incompatibilities, genome copy differences, genomic imprinting, and sex-specific differences can all contribute to asymmetric barriers that affect one lineage more than the other. A hallmark of cytonuclear incompatibility is that the nuclear genome has higher fitness on one of the two cytoplasmic genomes. In this case, I would predict that gene flow would progress more easily into the lineage with the more compatible cytoplasmic genome. When at least one genetic incompatibility is asymmetrical, I expect the response to secondary contact to be asymmetrical as well.

Campanula americana is an herbaceous plant species that is divided into partially reproductively isolated lineages that are in contact in three locations across its range. Rapidly-evolving chloroplast markers divide *C. americana* into two lineages: an Appalachian lineage common in the Appalachian mountains from Tennessee to Pennsylvania, and a Western lineage stretching from the Appalachian mountains west to Kansas and east of the Appalachian mountains in Virginia (Fig. 1) (Barnard-Kubow et al. 2015). Hybrids of these lineages have reduced germination and survival compared to their pure-lineage counterparts (Barnard-Kubow et al. 2016; Barnard-Kubow and Galloway 2017). Reductions in hybrid survival are largely driven by chlorosis of plants with Western chloroplasts on hybrid nuclear backgrounds; hybrids with Appalachian chloroplasts rarely express chlorosis (Barnard-Kubow et al. 2016, Chapter 1). During the last glaciation, the Western lineage was confined to a southern refugium on the Gulf Coast and an eastern refugium on the Atlantic coast, and the Appalachian lineage to microrefugia within the Appalachian Mountains (Barnard-Kubow et al. 2015). Over the last 20,000 years, the Western lineage has expanded from the Gulf Coast north as far as Pennsylvania, and from the coast of Virginia west to the Appalachian Mountains. This has created contact zones between the Appalachian and Western lineages in North Carolina, Pennsylvania, and Virginia (Fig. 1). A crossing study of *C. americana* at contact zones found that while reproductive isolation between lineages is high in a Pennsylvania contact zone, all isolating barriers except cytonuclear incompatibility are absent from the North Carolina contact zone (Chapter 1). Here, I use multiple populations for each contact zone as well as allopatry to test the predictions that (1) there is gene flow between lineages at contact zones; (2) within contact zones, gene flow is greater in North Carolina than in Pennsylvania as predicted by levels of reproductive isolation (Chapter 1); (3) where gene flow occurs it is asymmetrical, with plants with Appalachian (broadly compatible)

chloroplasts containing more mixed ancestry than those with Western (incompatible) chloroplasts.

Methods

The American Bellflower (*Campanula americana* L. = *Campanulastrum americanum* Small, Campanulaceae) is an monocarpic autotetraploid annual or biennial herb found in the eastern United States (Gadella 1964; Galloway et al. 2003; Galloway and Etterson 2005). Plants are largely outcrossing, but fully self compatible (Galloway et al. 2003). It grows in frequently disturbed partial-shade environments in forest edges, light gaps and roadsides. Seeds germinate in spring and fall, and plants bloom midsummer (June-August) after winter vernalization (Baskin and Baskin 1984). Flowers are protandrous, opening as male and transitioning to female phase following pollen removal (Evanhoe and Galloway 2002).

To test for gene flow between lineages of *Campanula americana*, I sampled populations of each lineage (Appalachian and Western) from each of four regions: three contact zones (North Carolina, Pennsylvania, and Virginia) and allopatry. Contact zones were defined as regions where populations were within 80 km of a population of the other lineage. If fresh leaf tissue was not available for collection, I germinated seeds in growth chambers and used the resulting seedlings for DNA extractions. In total, 326 plants from 29 populations were included in analyses: 17 Western populations (7 Allopatric, 5 NC, 3 PA, and 2 VA) and 12 Appalachian populations (3 Allopatric, 4 NC, 3 PA, and 2 VA) (Fig. 1, Table S1). In addition, I sampled 19 individuals from an outgroup population of *Triodanis perfoliata*, one of *Campanula americana*'s closest relatives (Crowl et al. 2016).

I used a double digest restriction site associated DNA sequencing (ddRADseq) protocol to obtain SNPs from across the genome (Peterson et al. 2012). I performed DNA extractions on either fresh or dried tissue in 96-well plates using a modified CTAB chloroform extraction protocol. Next gen sequencing libraries were built with a Biomek NXp liquid handling robot. Samples were randomized within plates, then digested with the enzymes ApoI HF and SphI HF simultaneously (NEB, Inc.). I then ligated one of 48 unique adapters to the DNA of each sample (Table S2), purified with Ampure beads, and size-selected for fragments 450-550bp long using a BluePippin machine (Sage Science, Inc.). After size selection, I made 9 pools of 48 samples, each with unique adaptors, and performed PCR on each pool using a different primer for each one, giving each sample a unique combination of indices (Fig. S6). PCR products were normalized based on qPCR, then pooled and sent to BGI for sequencing on two lanes of Illumina HiSeq 4000.

To analyze gene flow in parts of the genome with different inheritance patterns, I created four datasets from my sequencing reads: a dataset of all of the reads produced (full dataset), a dataset of reads aligned to the *C. americana* chloroplast genome (chloroplast dataset), a dataset of reads aligned to the *C. americana* mitochondrial genome (mitochondria dataset), and a dataset of reads that did not align to either the *C. americana* chloroplast or mitochondria genomes (nuclear dataset). I removed samples in the bottom 5th percentile of read number (less than 15,000 reads per individual). Reads were filtered and assembled using the program STACKS, version 1.48 (Catchen et al. 2011; 2013). Stacks parameters (-m 8 -M 4 -n 4) were determined by running Stacks on a subset of 13 samples, permuting each parameter (-m tested at 4 and 8, -M and –n tested at 0-9 and held equal to each other). I chose parameter values to maximize polymorphism without combining loci (Paris et al. 2017).

I made phylogenies to infer the genetic relationships between populations. Phylogenies were constructed using the program RAxML 8.2.4 (Stamatakis 2014). For each dataset, I ran 100 bootstraps using the GTR gamma method, with *Triodanis perfoliata* sequences as an outgroup. In addition, I used the program BEAST 2 to build a population-level phylogeny using a Bayesian method (Bouckaert et al. 2014). I used the nuclear, chloroplast, and mitochondrial datasets as partitions with independent clock and site models, and used bModelTest to determine site models (Bouckaert 2017). I ran a 10,000,000 generation MCMC chain with a coalescent constant population prior. I time-calibrated the phylogeny based on estimates of a split between *T. perfoliata* and *C. americana* 11.78 million years ago from a fossil-calibrated phylogeny of Campanulaceae (Mansion et al. 2012).

I explored the possibility of gene flow by evaluating population structure and admixture within populations using the program STRUCTURE (Pritchard et al. 2000). The program STRUCTURE clusters samples into populations using a MCMC approach to find linkage disequilibrium between loci. I used SNPs from loci present in at least 25 of the 29 populations. Using a burn-in of 10,000 and 30,000 reps I ran STRUCTURE 3 times for each k (assumed population number) of 1-6 separately for each region. The best k was selected using the Evanno method with the program STRUCTURE Harvester (Evanno et al. 2005; Earl and vonHoldt 2012). I clustered runs of STRUCTURE using the program CLUMPP (Jakobsson and Rosenberg 2007), and graphed using the program Distruct (Rosenberg 2004).

I then tested for gene flow across the phylogeny of *C. americana*. I did this using the Treemix program which creates bifurcating tree models with migration edges between branches (Pickrell and Pritchard 2012). I tested these models allowing for 0-10 migration edges between branches. Only loci present in all populations were included in this analysis. Running Treemix

on nuclear-only data accounts for shared the genes of Appalachian and Western North Carolina populations by placing them together on the phylogenetic tree, so migration edges are not added within that region. Therefore, I only present Treemix analyses using the full dataset.

Finally, I conducted two different tests to evaluate gene flow between pairs of populations in contact zones. For both tests, I used the nuclear dataset to best assess gene movement at contact zones. First, I performed Four Population tests which ask whether gene flow between two sympatric (i.e. contact) populations makes them appear more closely related than expected relative to two allopatric populations included in the test (Patterson et al. 2012; Peter 2016). I used software included on the Treemix package to perform Four Population tests on sets of populations with the format ($Western_{Allopatry}$, $Western_{Contact}$; $Appalachian_{Allopatry}$, $Appalachian_{Contact}$), defining lineage by cytoplasmic genotype. Finally, to confirm findings from Four Population tests and to distinguish gene flow from incomplete lineage sorting, I conducted ABBA/BABA tests (Durand et al. 2011). I used the package evobiR in R to calculate D statistics for my populations, using 1,000 bootstraps to calculate Z scores. For both Four Populations and ABBA/BABA tests, I calculated Z score significance cutoffs for multiple testing comparisons.

Results

Sequencing and dataset creation

Sequencing produced a median of 1.2 million reads per individual. Of these, an average of 3% of reads aligned to chloroplast sequence, and 10% aligned to mitochondrial sequence.

Evolutionary relatedness of populations

To assess the evolutionary relationships between populations, I constructed phylogenies separately using the nuclear, cytoplasmic, and full datasets. Phylogenies were concordant

between the mitochondrial and chloroplast datasets (Fig. S1C & S1D), showing deep divergence between the Appalachian and Western lineages, and more recent divergence within each lineage (Fig. 2B). The phylogeny built from the nuclear dataset provides more interpopulation resolution than the cytoplasmic phylogeny, but is discordant with it (Fig. 2C). The cytoplasmic phylogenies separate populations into Appalachian and Western lineages (see also Fig. S1). However, the nuclear phylogeny places populations from the North Carolina contact zone close to one another on the tree, regardless of lineage, while maintaining lineage separation between the Appalachian and Western populations in Pennsylvania and Virginia contact zones. Phylogenies built using Bayesian (Fig. 2, BEAST) and Maximum Likelihood methods (Fig. S1, RAxML) are largely concordant, but the Bayesian nuclear phylogeny places all North Carolina populations within the Western lineage (Fig. 2C), while the Maximum Likelihood nuclear phylogeny places them basally in the tree (Fig. S1B).

Population structure

For each region, the STRUCTURE models with the best support have two populations when using the full dataset (Fig. S2). At Pennsylvania and Virginia contact zones, STRUCTURE divides individuals into two populations that cleanly correspond to their cytoplasmic lineage (Fig. 3 & S3). However, the North Carolina region shows a large proportion of shared ancestry within populations (Fig. 3 & S3). This dichotomy is even more pronounced in models built from nuclear-only datasets. In those, STRUCTURE divides individuals into populations that cleanly correspond to their cytoplasmic lineage for Pennsylvania and Virginia zones (Fig. 3 & S4), but individuals from the North Carolina zone are indistinguishable as one lineage or the other (Fig. 3 & S4). When all populations were run in the same model with the nuclear dataset, allopatric,

Virginia, and Pennsylvania populations were divided by lineage, and North Carolina populations were identified as a single cluster that was neither Appalachian nor Western (Fig. S5).

Gene flow between lineages

In comparisons of gene flow across the phylogeny, gene flow was limited to North Carolina. Treemix models placed two migration edges between North Carolina populations of the Appalachian and Western branches of the phylogeny (Fig. 4). The first migration edge was from an Appalachian population to the ancestor of four Western populations, and the second was from a Western population to an Appalachian population.

Tests for gene flow found a large amount of gene flow in North Carolina, and only limited amounts in the other contact zones. Four Population tests demonstrated that gene flow is likely to be present in North Carolina, but little to none in Pennsylvania or Virginia (Fig. 5, Table S3). ABBA BABA tests similarly predicted high levels of gene flow among many of the populations in North Carolina, but little in any other region (Fig. 5, Table S4).

Discussion

Variation in gene flow among contact zones

In this study I found evidence of extensive gene flow at one contact zone between Appalachian and Western lineages of *Campanula americana*, but little to none at the other two contact zones. Discordance between cytoplasmic and nuclear phylogenies reveals a history of admixture in North Carolina (Fig. 2B & 2C). North Carolina populations retain distinct lineage identity in their chloroplast and mitochondria DNA (Fig. 2B), but gene flow has at least partially homogenized their nuclear DNA. Specifically, while phylogenies built from the full dataset divide populations by chloroplast lineage (Fig. 2A), phylogenies built from only nuclear data show North Carolina populations as closely related irrespective of lineage (Fig. 2C). Also, in

North Carolina the two lineages are indistinguishable in tests of population structure using nuclear markers (Fig. 3) and tests support gene exchange (Fig. 4 & 5). An alternative explanation for a nuclear genome that is homogenized across the contact zone in North Carolina could be chloroplast capture; this idea is explored below. In contrast, there is little evidence of gene flow in allopatry or in the Pennsylvania and Virginia contact zones. Nuclear and whole genome phylogenies separate populations by lineage and within that, contact zone (Fig. 2A & 2B). There is also clear population structure within each contact zone (Fig. 3), and limited evidence of gene exchange (Fig. 4 & 5). In summary, the Appalachian and Western lineages in the Pennsylvania and Virginia contact zones are genetically distinct while in North Carolina the lineages appear to be merging.

Variation in reproductive isolation in different regions may be driven by variation in either extrinsic or intrinsic barriers (Coyne and Orr 2004). Extrinsic barriers are expected to vary across species ranges, since they depend on ecological factors that often vary geographically (Coyne and Orr 2004; Lepais et al. 2009; Mandeville et al. 2015; Eaton et al. 2015). However, intrinsic barriers (those that are controlled genetically) can also differ within species (McDermott and Noor 2011), both among populations of a species pair (Matute et al. 2014) and among population pairs within a species (Corbett-Detig et al. 2013). Accumulation of reproductive isolation with genetic distance can create this variation in intrinsic reproductive isolation within a species (Irwin et al. 2001; Alcaide et al. 2014). The differences I find between contact zones of the lineages of *C. americana* are likely driven by ecological variation and variation in intrinsic reproductive isolation at time of initial contact across the range.

Combining information on gene exchange presented here with controlled environment estimates of reproductive isolation within and between lineages (Barnard-Kubow et al. 2016;

Barnard-Kubow and Galloway 2017) suggests that the preexisting levels of reproductive isolation due to germination varied among contact zones at initial contact. Hybrids between the lineages in Pennsylvania and allopatry show reduced germination relative to parents, but crosses between Appalachian and Western lineages in North Carolina have no reduction in germination (Chapter 1). Previous studies (Barnard-Kubow et al. 2016); (Barnard-Kubow and Galloway 2017) found reduced germination in hybrids between North Carolina and allopatric Western plants, but not between North Carolina and allopatric Appalachian plants or between Virginia Western and Appalachian plants. Since only populations in the Pennsylvania/Allopatric group of the Western lineage (Fig 2A) are involved in hybrid germination breakdown, alleles leading to that breakdown arose in the Pennsylvania/allopatric Western group but not in the Virginia/North Carolina Western group, rather than being purged after contact in North Carolina. This means that intrinsic barriers were likely stronger at initial contact in Pennsylvania because the Western lineage may have accumulated additional isolating barriers (germination) as it migrated from its refugia, to North Carolina, to its current extent. This difference in initial levels of reproductive isolation may at least partially explain the differences in gene flow between North Carolina and Pennsylvania.

Since the Virginia contact zone lacks reproductive isolation due to germination similar to the North Carolina zone, but has low gene flow like the Pennsylvania zone, preexisting levels of reproductive isolation do not fully explain the differences in outcome of secondary contact. Therefore, it is likely that some extrinsic barriers are at play here. Populations are sparser at the Pennsylvania and Virginia contact zones, likely because of their proximity to less suitable habitat at the range edge (Barnard-Kubow et al. 2015), so frequency of contact may be lower in Pennsylvania and Virginia than in North Carolina. Fragmented, mosaic contact zones are more

likely to lead to reinforcement than to lineages merging (Cain et al. 1999), so this is one likely reason for the differences between the North Carolina and Virginia contact zones. In total, variation in both intrinsic and extrinsic barriers across the range is likely to contribute to differences in outcomes of secondary contact in *C. americana*.

Asymmetry

Studies of postzygotic reproductive isolation show that a cytonuclear incompatibility is ubiquitous in *C. americana* (Chapter 1). Because the Appalachian chloroplast is compatible with both Appalachian and Western nuclear DNA, but the Western chloroplast is incompatible with Appalachian nuclear DNA, I expect to only find hybrid nuclear DNA on Appalachian cytoplasmic backgrounds. However, I found only weak evidence of asymmetrical gene flow that would likely result from these incompatibilities; both Appalachian- and Western-chloroplast plants from North Carolina show generally mixed ancestry. In greenhouse crosses, hybrids from North Carolina plants with Western lineage mothers have inactive chloroplasts and high mortality, but hybrids with Appalachian mothers have high fitness (Chapter 1). Because of this, I expected to find that wild individuals with mixed-lineage ancestry would have Appalachian chloroplast haplotypes. North Carolina populations are placed sister to the Western lineage on the nuclear phylogeny, which suggests that the nuclear DNA of those populations is more similar to the Western lineage than the Appalachian lineage. However, since mixed-ancestry individuals were found to have either Western or Appalachian chloroplast haplotypes, gene flow appears to be symmetrical.

One potential explanation for similar nuclear DNA but distinction in chloroplast DNA among populations in North Carolina is chloroplast capture (Rieseberg and Soltis 1991). In chloroplast capture, the chloroplast from one lineage becomes common in populations with the

nuclear DNA of another lineage through backcrossing. If this was the case in *C. americana*, I would expect the Appalachian chloroplast to be common on a Western nuclear background because of the pattern of compatibility. North Carolina populations of *C. americana* are sister to allopatric and Pennsylvania populations of the Western lineage (Fig. 2C), which may support this scenario. However, nuclear population structure shows that North Carolina is a separate cluster from the Appalachian and Western lineages (Fig. 3 & S4), suggesting that gene flow began long enough ago for new linkage disequilibrium patterns to emerge and that the nuclear DNA in that region is neither fully Western nor fully Appalachian. In addition, the nuclear component of the cytonuclear interaction is present in North Carolina (Chapter 1), which would be unlikely in the case of chloroplast capture. In total, chloroplast capture is unlikely to explain the pattern of gene flow in *C. americana* in the North Carolina contact zone.

It is surprising that the cytonuclear incompatibility persists even on an extensively mixed nuclear background; one might expect that the nuclear component of the incompatibility would be purged in this case (Felsenstein 1981), or that one chloroplast lineage would spread across the entire region (Rieseberg and Soltis 1991). *Campanula americana* sometimes inherits both its maternal and its paternal chloroplasts (Barnard-Kubow et al. 2017), which would allow Appalachian chloroplasts to spread even to previously Western populations with hybridization. However, since all populations contain only one lineage's chloroplast and the Western chloroplast persists in North Carolina, this doesn't appear to be driving the overall pattern of chloroplast distribution in North Carolina. It is possible that some selective advantage to the nuclear component of the incompatibility is allowing cytonuclear reproductive isolation to persist in the otherwise-merged North Carolina region.

Conclusion

In this study I found evidence of extensive gene flow at one secondary contact zone, but little to no gene flow at two other contact zones. This supports results from previous studies that found that postzygotic reproductive isolation is reduced in sympatry in North Carolina relative to allopatry, but is high in the Pennsylvania contact zone. High gene flow and low reproductive isolation suggests that the lineages are merging where they are in contact in North Carolina. However, there is little evidence of a similar process in Pennsylvania or Virginia. This is likely due to differences in initial reproductive isolation at secondary contact, structure of the contact zones, or other ecological factors. In North Carolina where gene flow is high, a cytonuclear incompatibility remains, which makes hybrids with Appalachian mothers more fit than hybrids with Western lineage mothers (Chapter 1). However, there is no concordant asymmetry in direction of gene flow. Instead, plants with each cytoplasmic genotype have mixed nuclear ancestry. This study finds variation among three contact zones in the amount of gene flow resulting from secondary contact between partially isolated lineages of *C. americana*. This suggests that variation in reproductive isolation and habitat suitability across the range of a species can change the outcome of secondary contact, thereby altering the outcome of speciation.

Acknowledgements

I thank Karen Barnard-Kubow, Rebecca Risser, Judy Debban, Paul Debban, Jeanette Engel, Susan Munch, and Marge Van Tassel for help finding populations of *Campanula americana* for this study. I also thank AnhThu Nguyen for assistance with sequencing library construction, Andrew Crowl for advice on outgroup species, Alan Bergland for advice on analyses, and the Galloway lab for discussion. This work was supported by an NSF GRFP and a Rosemary Grant Award from the Society for the Study of Evolution.

Figure 1. (A) Range of the Western (purple) and Appalachian (green) lineages of *Campanula americana*. (B) Locations of populations used in this study. Populations from the Appalachian lineage are marked in green, and those from the Western lineage are marked in purple. See Table S1 for exact population locations.

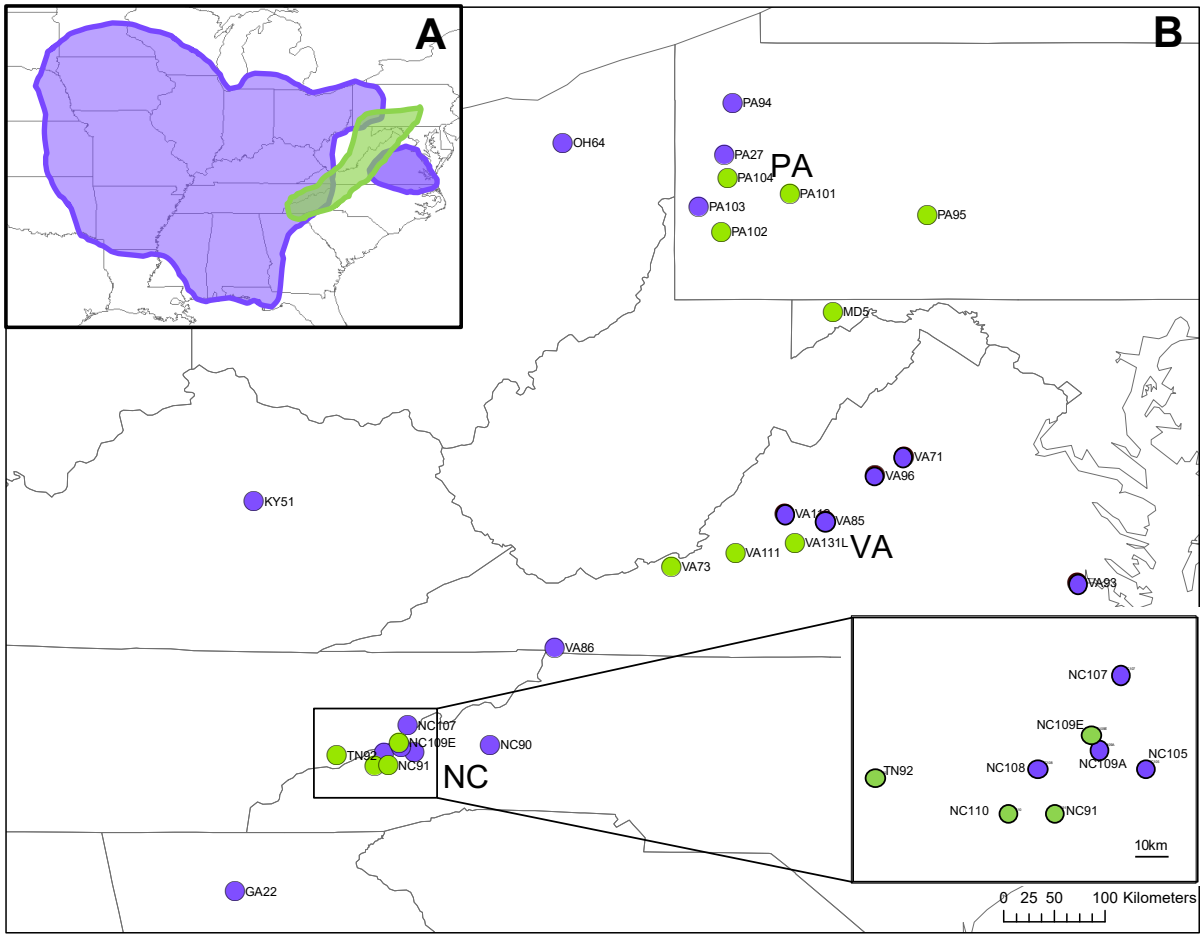


Figure 2. Population trees built with a Bayesian method from (A) the full dataset, (B) the chloroplast and mitochondrial datasets combined, and (C) the nuclear dataset. Nodes are marked with posterior probabilities. Western lineage populations are marked in purple, and Appalachian lineage populations are marked in green.

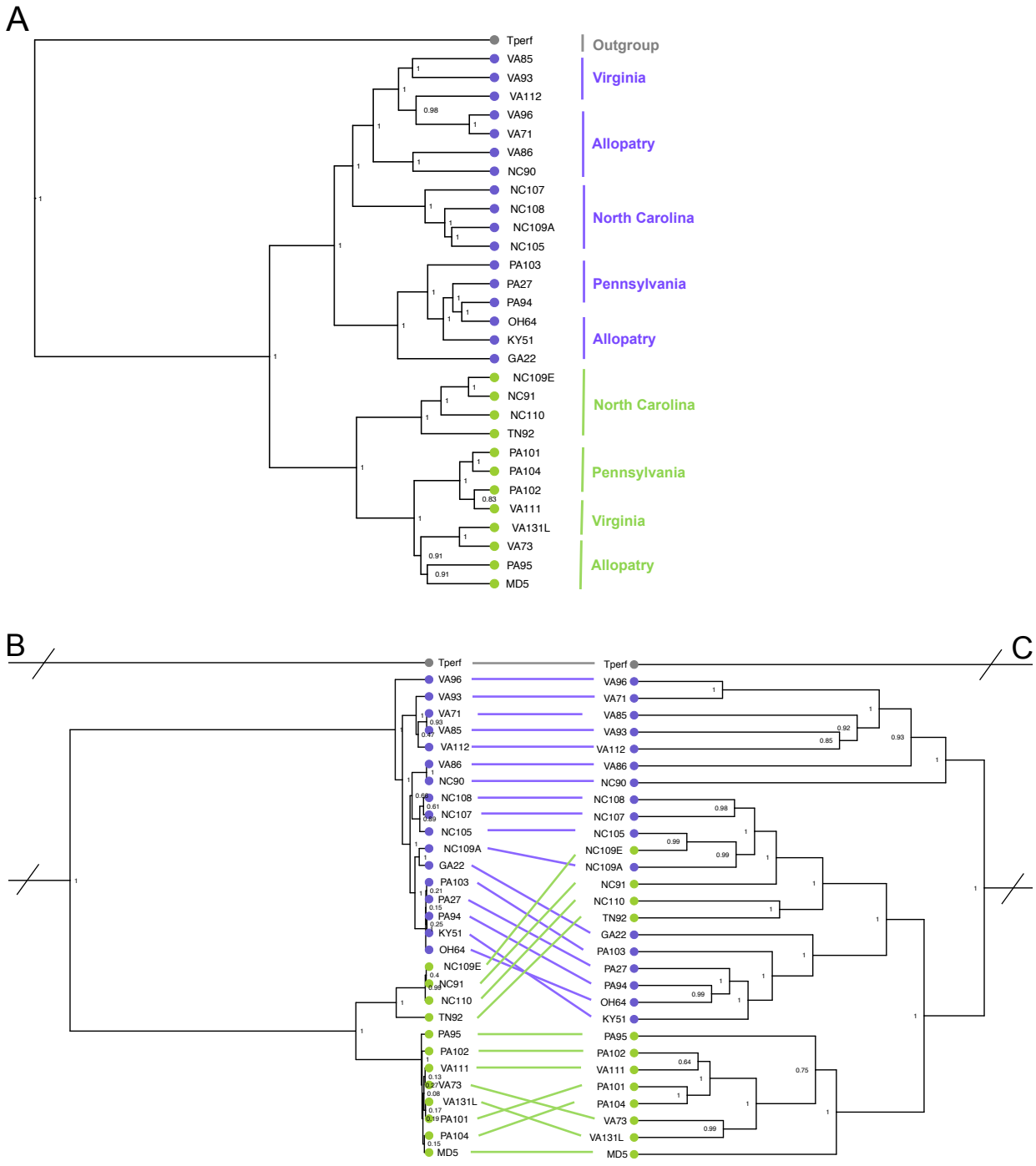


Figure 3. STRUCTURE plots with the best k values of Allopatry, Pennsylvania, Virginia, and North Carolina. Top bar represents the cytoplasmic lineage of individuals in each population below. Middle plots were built using the full dataset, and bottom plots were built using the nuclear dataset. Populations in each region are arranged by geography; populations closest to the other lineage geographically are closest to the other lineage in the STRUCTURE plots. Clusters are colored by the lineages of the populations in which they are most common: purple for Western and green for Appalachian. Clusters equally common in multiple lineages are colored grey.

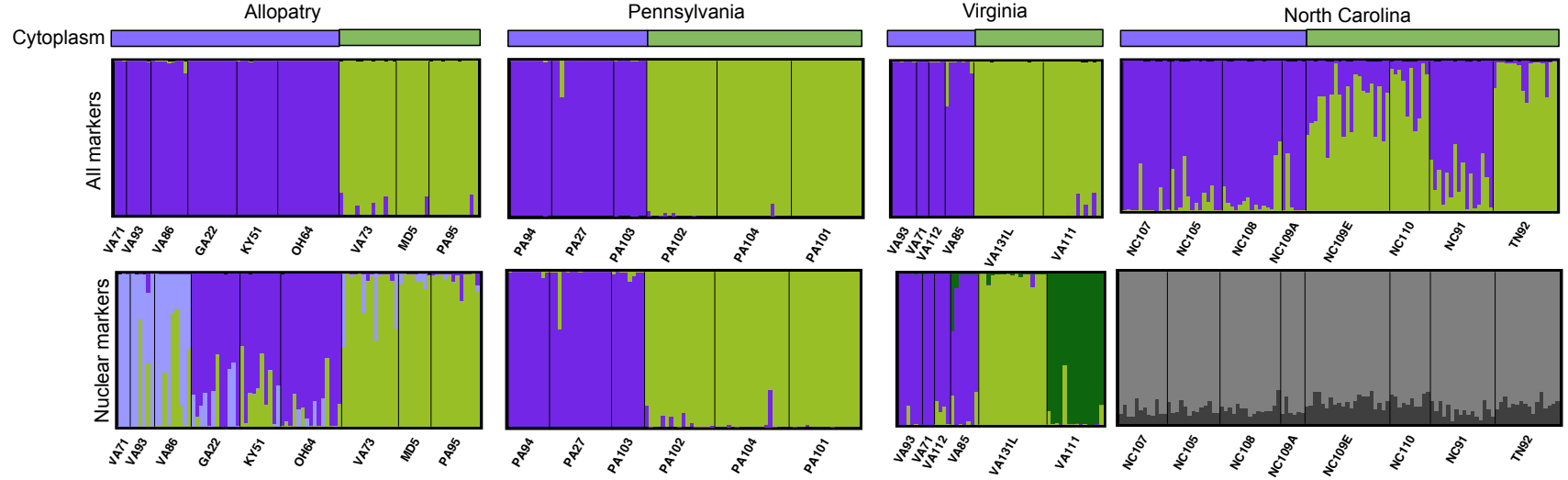


Figure 4. Phylogenetic tree including two hypothesized migration events (marked by orange arrows), created with the program Treemix. Population names are colored by their cytoplasmic lineage: Appalachian in green and Western in purple.

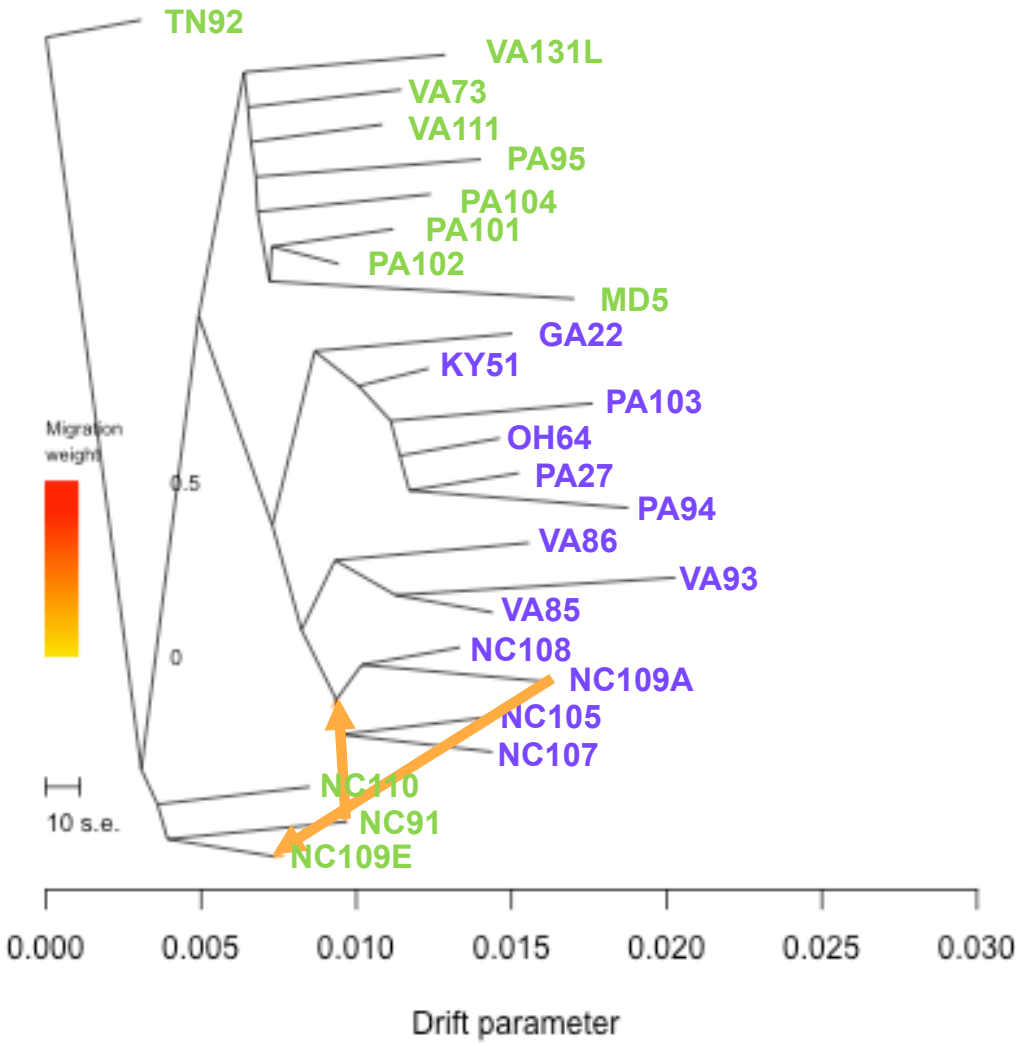


Figure 5. Maps showing gene flow between populations, as predicted by significant Four Population tests (left) and ABBA-BABA tests (right). Gene flow line weight determined by number of significant tests indicating gene flow (See Tables S3 and S4).

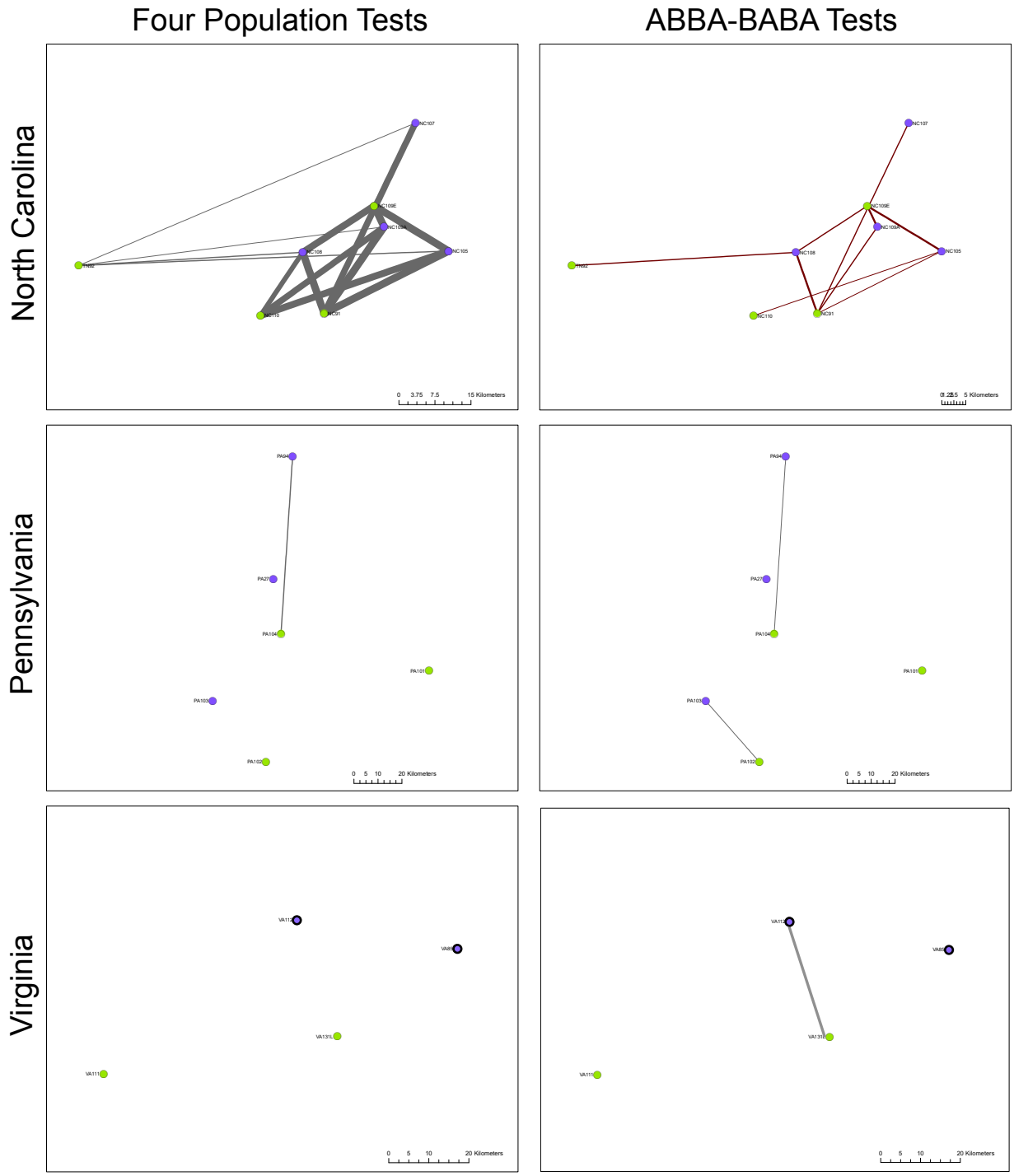


Table S1. Location, lineage, and number of individuals analyzed for each population in this study.

Population	Samples	Lineage	Region	Latitude	Longitude
MD5	8	Appalachian	Allopatry	39.61	-79.12
PA95	12	Appalachian	Allopatry	40.48	-78.28
VA73	14	Appalachian	Allopatry	37.35	-80.55
GA22	12	Western	Allopatry	34.47	-84.43
KY51	10	Western	Allopatry	37.93	-84.26
OH64	15	Western	Allopatry	41.12	-81.52
VA71	3	Western	Allopatry	38.33	-78.49
VA86	10	Western	Allopatry	36.63	-81.59
VA93	4	Western	Allopatry	37.21	-76.95
VA96	1	Western	Allopatry	38.16	-78.75
NC109E	19	Appalachian	NC Contact	35.79	-82.97
NC110	10	Appalachian	NC Contact	35.58	-83.19
NC91	17	Appalachian	NC Contact	35.59	-83.07
TN92	17	Appalachian	NC Contact	35.68	-83.53
NC105	13	Western	NC Contact	35.70	-82.83
NC107	12	Western	NC Contact	35.94	-82.90
NC108	15	Western	NC Contact	35.70	-83.11
NC109A	6	Western	NC Contact	35.75	-82.95
NC90	1	Western	NC Contact	35.77	-82.16
PA101	17	Appalachian	PA Contact	40.66	-79.50
PA102	17	Appalachian	PA Contact	40.32	-80.11
PA104	12	Appalachian	PA Contact	40.80	-80.06
PA103	8	Western	PA Contact	40.55	-80.31
PA27	15	Western	PA Contact	41.01	-80.08
PA94	10	Western	PA Contact	41.47	-80.01
VA111	13	Appalachian	VA Contact	37.48	-79.99
VA131L	17	Appalachian	VA Contact	37.55	-79.46
VA112	4	Western	VA Contact	37.82	-79.55
VA85	7	Western	VA Contact	37.76	-79.19
Triodanis perfoliata	19	Outgroup	Outgroup	37.42	-80.38

Table S2. Sequences of P1 and P2 adapter oligonucleotides and PCR indices used in library preparation.

NAME	BARCODE	SEQ (5' --> 3')
Index1_EcoRI_P1a	ATCACG	ACACTCTTTCCCTACACGACGCTCTTCCGATCTATCACG
Index1_EcoRI_P1b		/5Phos/AATTCGTGATAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
Index2_EcoRI_P1a	CGATGT	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCGATGT
Index2_EcoRI_P1b		/5Phos/AATTACATCGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
Index3_EcoRI_P1a	TTAGGC	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTTAGGC
Index3_EcoRI_P1b		/5Phos/AATTGCCTAAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
Index4_EcoRI_P1a	TGACCA	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTGACCA
Index4_EcoRI_P1b		/5Phos/AATTTGGTCAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
Index5_EcoRI_P1a	ACAGTG	ACACTCTTTCCCTACACGACGCTCTTCCGATCTACAGTG
Index5_EcoRI_P1b		/5Phos/AATTCACTGTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
Index6_EcoRI_P1a	GCCAAT	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGCCAAT
Index6_EcoRI_P1b		/5Phos/AATTATTGGCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
Index7_EcoRI_P1a	CAGATC	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCAGATC
Index7_EcoRI_P1b		/5Phos/AATTGATCTGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
Index8_EcoRI_P1a	ACTTGA	ACACTCTTTCCCTACACGACGCTCTTCCGATCTACTTGA
Index8_EcoRI_P1b		/5Phos/AATTTCAAGTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
Index9_EcoRI_P1a	GATCAG	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGATCAG
Index9_EcoRI_P1b		/5Phos/AATTCTGATCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
Index10_EcoRI_P1a	TAGCTT	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTAGCTT
Index10_EcoRI_P1b		/5Phos/AATTAAGCTAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
Index11_EcoRI_P1a	GGCTAC	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGGCTAC
Index11_EcoRI_P1b		/5Phos/AATTGTAGCCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
Index12_EcoRI_P1a	CTTGTA	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCTTGTA
Index12_EcoRI_P1b		/5Phos/AATTTACAAGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
Index13_EcoRI_P1a	AGTCAA	ACACTCTTTCCCTACACGACGCTCTTCCGATCTAGTCAA
Index13_EcoRI_P1b		/5Phos/AATTTTGACTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT

Index14_EcoRI_P1a	AGTTCC	ACACTCTTTCCCTACACGACGCTCTTCCGATCTAGTTCC
Index14_EcoRI_P1b		/5Phos/AATTGGAACTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
Index15_EcoRI_P1a	ATGTCA	ACACTCTTTCCCTACACGACGCTCTTCCGATCTATGTCA
Index15_EcoRI_P1b		/5Phos/AATTTGACATAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
Index16_EcoRI_P1a	CCGTCC	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCCGTCC
Index16_EcoRI_P1b		/5Phos/AATTGGACGGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
Index17_EcoRI_P1a	GTAGAG	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGTAGAG
Index17_EcoRI_P1b		/5Phos/AATTCTCTACAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
Index18_EcoRI_P1a	GTCCGC	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGTCCGC
Index18_EcoRI_P1b		/5Phos/AATTGCGGACAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
Index19_EcoRI_P1a	GTGAAA	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGTGAAA
Index19_EcoRI_P1b		/5Phos/AATTTTTCACAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
Index20_EcoRI_P1a	GTGGCC	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGTGGCC
Index20_EcoRI_P1b		/5Phos/AATTGGCCACAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
Index21_EcoRI_P1a	GTTTCG	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGTTTCG
Index21_EcoRI_P1b		/5Phos/AATTCGAAACAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
Index22_EcoRI_P1a	CGTACG	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCGTACG
Index22_EcoRI_P1b		/5Phos/AATTCGTACGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
Index23_EcoRI_P1a	GAGTGG	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGAGTGG
Index23_EcoRI_P1b		/5Phos/AATTCCACTCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
Index24_EcoRI_P1a	GGTAGC	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGGTAGC
Index24_EcoRI_P1b		/5Phos/AATTGCTACCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
Index25_EcoRI_P1a	ACTGAT	ACACTCTTTCCCTACACGACGCTCTTCCGATCTACTGAT
Index25_EcoRI_P1b		/5Phos/AATTATCAGTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
Index26_EcoRI_P1a	ATGAGC	ACACTCTTTCCCTACACGACGCTCTTCCGATCTATGAGC
Index26_EcoRI_P1b		/5Phos/AATTGCTCATAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
Index27_EcoRI_P1a	ATTCCT	ACACTCTTTCCCTACACGACGCTCTTCCGATCTATTCCT
Index27_EcoRI_P1b		/5Phos/AATTAGGAATAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT

Index28_EcoRI_P1a	CAAAAG	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCAAAAG
Index28_EcoRI_P1b		/5Phos/AATTCTTTTGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
Index29_EcoRI_P1b	CAACTA	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCAACTA
Index29_EcoRI_P1b		/5Phos/AATTTAGTTGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
Index30_EcoRI_P1a	CACCGG	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCACCGG
Index30_EcoRI_P1b		/5Phos/AATTCCGGTGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
Index31_EcoRI_P1a	CACGAT	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCACGAT
Index31_EcoRI_P1b		/5Phos/AATTATCGTGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
Index32_EcoRI_P1a	CACTCA	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCACTCA
Index32_EcoRI_P1b		/5Phos/AATTTGAGTGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
Index33_EcoRI_P1a	CAGGCG	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCAGGCG
Index33_EcoRI_P1b		/5Phos/AATTCGCCTGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
Index34_EcoRI_P1a	CATGGC	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCATGGC
Index34_EcoRI_P1b		/5Phos/AATTGCCATGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
Index35_EcoRI_P1a	CATTTT	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCATTTT
Index35_EcoRI_P1b		/5Phos/AATTAAAATGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
Index36_EcoRI_P1a	CCAACA	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCCAACA
Index36_EcoRI_P1b		/5Phos/AATTTGTTGGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
Index37_EcoRI_P1a	CGGAAT	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCGGAAT
Index37_EcoRI_P1b		/5Phos/AATTATTCCGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
Index38_EcoRI_P1a	CTAGCT	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCTAGCT
Index38_EcoRI_P1b		/5Phos/AATTAGCTAGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
Index39_EcoRI_P1a	CTATAC	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCTATAC
Index39_EcoRI_P1b		/5Phos/AATTGTATAGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
Index40_EcoRI_P1a	CTCAGA	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCTCAGA
Index40_EcoRI_P1b		/5Phos/AATTTCTGAGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
Index41_EcoRI_P1a	GACGAC	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGACGAC
Index41_EcoRI_P1b		/5Phos/AATTGTCGTCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT

Index42_EcoRI_P1a	TAATCG	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTAATCG
Index42_EcoRI_P1b		/5Phos/AATTCGATTAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
Index43_EcoRI_P1a	TACAGC	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTACAGC
Index43_EcoRI_P1b		/5Phos/AATTGCTGTAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
Index44_EcoRI_P1a	TATAAT	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTATAAT
Index44_EcoRI_P1b		/5Phos/AATTATTATAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
Index45_EcoRI_P1a	TCATTC	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTCATTC
Index45_EcoRI_P1b		/5Phos/AATTGAATGAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
Index46_EcoRI_P1a	TCCCGA	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTCCCGA
Index46_EcoRI_P1b		/5Phos/AATTTCGGGAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
Index47_EcoRI_P1a	TCGAAG	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTCGAAG
Index47_EcoRI_P1b		/5Phos/AATTCTTCGAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
Index48_EcoRI_P1a	TCGGCA	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTCGGCA
Index48_EcoRI_P1b		/5Phos/AATTTGCCGAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
SphI_P2a		/5Phos/AGATCGGAAGAGCGAGAACAA
SphI_P2b		GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCATG

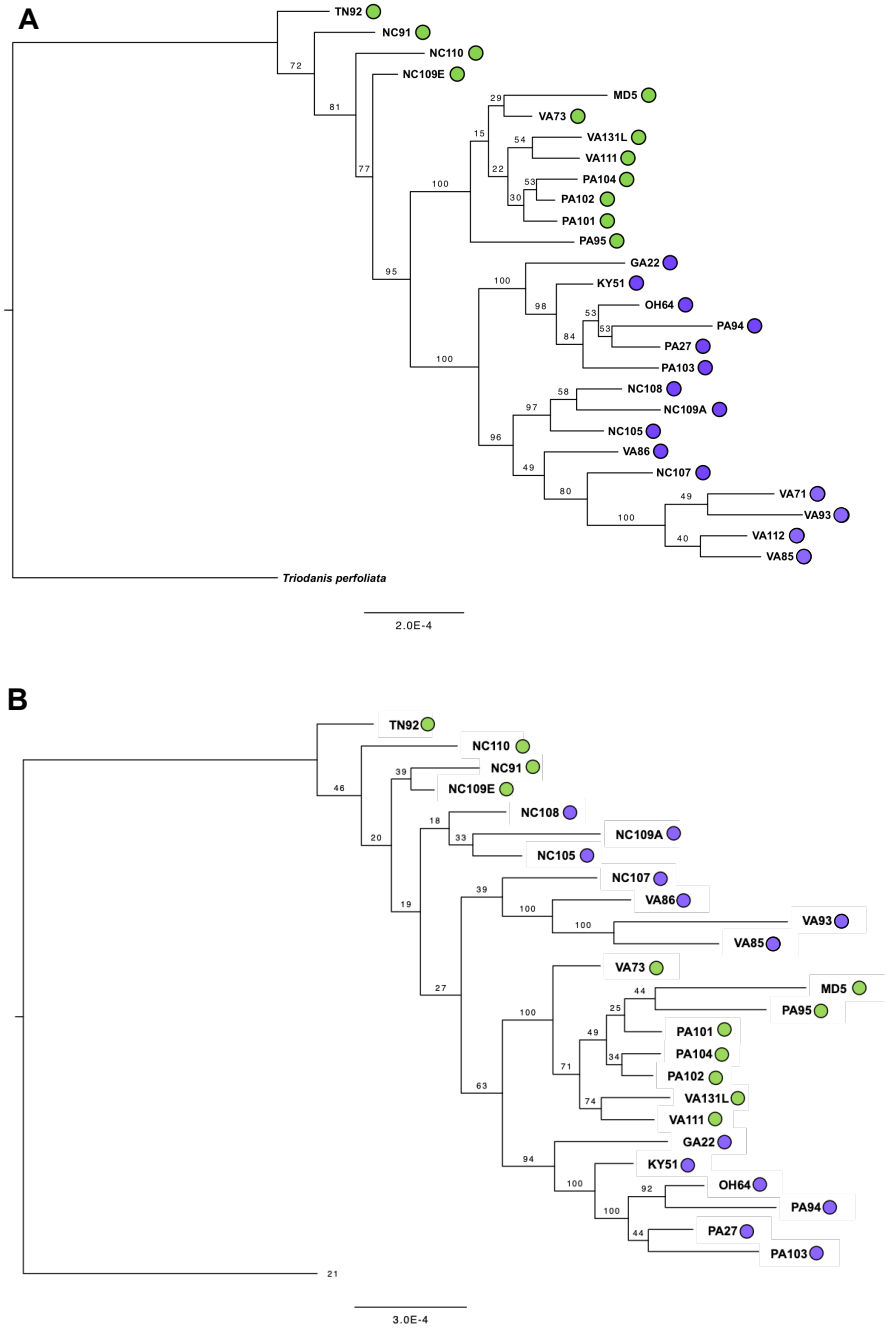
Table S3. Four Population test means by sympatric Appalachian and Western population pair. Z scores marked with a * denote significant gene flow after Bonferroni correction for multiple testing. Number of significant tests after multiple testing correction are out of 12 except for the Virginia region, which are out of 3 because of fewer available allopatric populations.

Region	Appalachian population	Western Population	Mean Z	Significant tests
NC	NC109E	NC105	6.0784*	12
NC	NC109E	NC107	2.3715	5
NC	NC109E	NC108	4.5702*	11
NC	NC109E	NC109A	8.7868*	12
NC	NC110	NC105	4.7645*	11
NC	NC110	NC107	0.3517	0
NC	NC110	NC108	3.5306	8
NC	NC110	NC109A	4.4847*	10
NC	NC91	NC105	6.1346	12
NC	NC91	NC107	2.9063	10
NC	NC91	NC108	6.6912*	12
NC	NC91	NC109A	10.1243*	12
NC	TN92	NC105	2.3098	2
NC	TN92	NC107	1.3967	1
NC	TN92	NC108	2.7043	2
NC	TN92	NC109A	2.5623	1
PA	PA101	PA103	-1.4296	0
PA	PA101	PA27	0.6097	0
PA	PA101	PA94	0.7279	0
PA	PA102	PA103	0.3185	0
PA	PA102	PA27	0.2032	0
PA	PA102	PA94	-0.1275	0
PA	PA104	PA103	-1.3472	0
PA	PA104	PA27	1.3986	0
PA	PA104	PA94	2.7770	4
VA	VA111	VA85	-0.5735	0
VA	VA131L	VA85	-0.4675	0

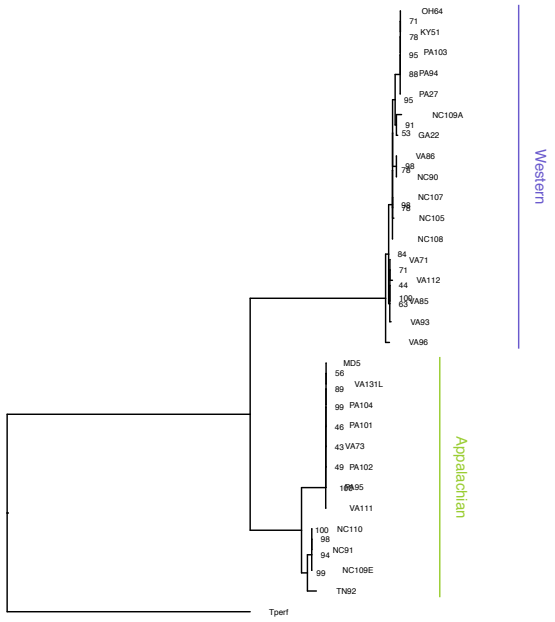
Table S4. ABBA-BABA test means by sympatric Appalachian and Western population pair. Z scores marked with a * denote significant gene flow after Bonferroni correction for multiple testing. Number of significant tests after multiple testing correction are out of 7.

Region	Appalachian population	Western population	Mean D	Mean Z	Significant Tests
NC	NC109E	NC105	0.40	3.40	3
NC	NC109E	NC107	0.11	1.23	0
NC	NC109E	NC108	0.33	2.90	1
NC	NC109E	NC109A	0.41	3.33	3
NC	NC110	NC105	0.28	2.21	1
NC	NC110	NC107	-0.20	0.11	0
NC	NC110	NC108	0.04	1.38	0
NC	NC110	NC109A	0.13	1.67	0
NC	NC91	NC105	0.22	2.40	1
NC	NC91	NC107	0.30	2.60	2
NC	NC91	NC108	0.40	3.44	3
NC	NC91	NC109A	0.30	3.07	2
NC	TN92	NC105	0.03	1.31	0
NC	TN92	NC107	0.02	1.08	0
NC	TN92	NC108	0.26	2.13	2
NC	TN92	NC109A	0.07	1.24	0
PA	PA101	PA103	-0.10	1.05	0
PA	PA101	PA27	-0.03	1.03	0
PA	PA101	PA94	-0.35	-2.67	0
PA	PA102	PA103	0.18	1.77	1
PA	PA102	PA27	0.13	1.34	0
PA	PA102	PA94	-0.15	-1.83	0
PA	PA104	PA103	0.12	1.21	0
PA	PA104	PA27	-0.02	0.82	0
PA	PA104	PA94	0.03	1.18	1
VA	VA111	VA112	-0.05	0.87	0
VA	VA111	VA85	-0.12	0.35	0
VA	VA131L	VA112	0.40	2.05	2
VA	VA131L	VA85	0.00	0.95	0

Figure S1. Population trees built with a maximum likelihood method from (A) the full dataset, (B) the nuclear dataset, (C) the chloroplast dataset, and (D) the mitochondrial dataset. Colors represent cytoplasmic lineages: Appalachian (green) and Western (purple).



C


OH64
KY51
PA103
PA94
PA27
NC109A
GA22
VA86
NC90
NC107
NC105
NC108
VA71
VA112
VA85
VA93
VA96
Western
MD5
VA131L
PA104
PA101
VA73
PA102
VA111
NC110
NC91
NC109E
TN92
Appalachian
Tperf

D

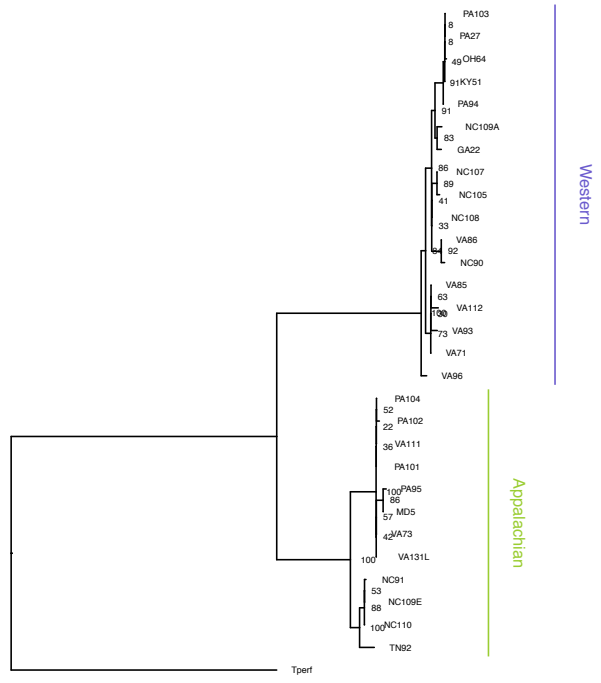

PA103
PA27
OH64
KY51
PA94
NC109A
GA22
NC107
NC105
NC108
VA86
NC90
VA85
VA112
VA93
VA71
VA96
Western
PA104
PA102
VA111
PA101
PA95
MD5
VA73
VA131L
NC91
NC109E
NC110
TN92
Appalachian
Tperf

Figure S2. Evanno plots showing best k values for analysis with STRUCTURE.

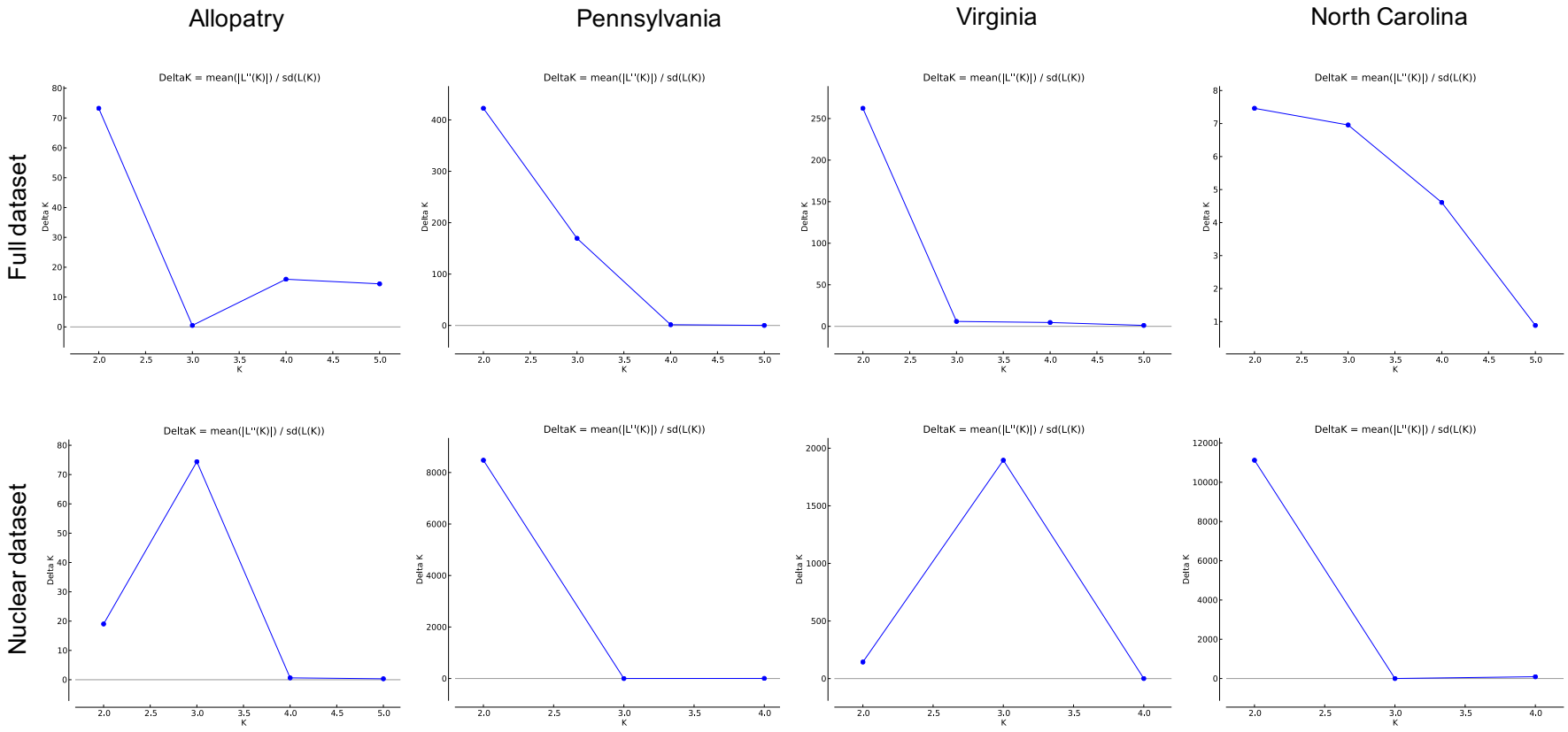


Figure S3. STRUCTURE plots for all k values 2-5 for the full dataset of (from left to right) Allopatry, Pennsylvania, Virginia, and North Carolina. Best value of k for each region is marked with a *. Populations in each region are arranged by geography; populations closest to the other lineage geographically are closest to the other lineage on the STRUCTURE plots. Clusters are colored by the lineages of the populations in which they are most common: purple for Western and green for Appalachian. Clusters equally common in multiple lineages are colored grey.

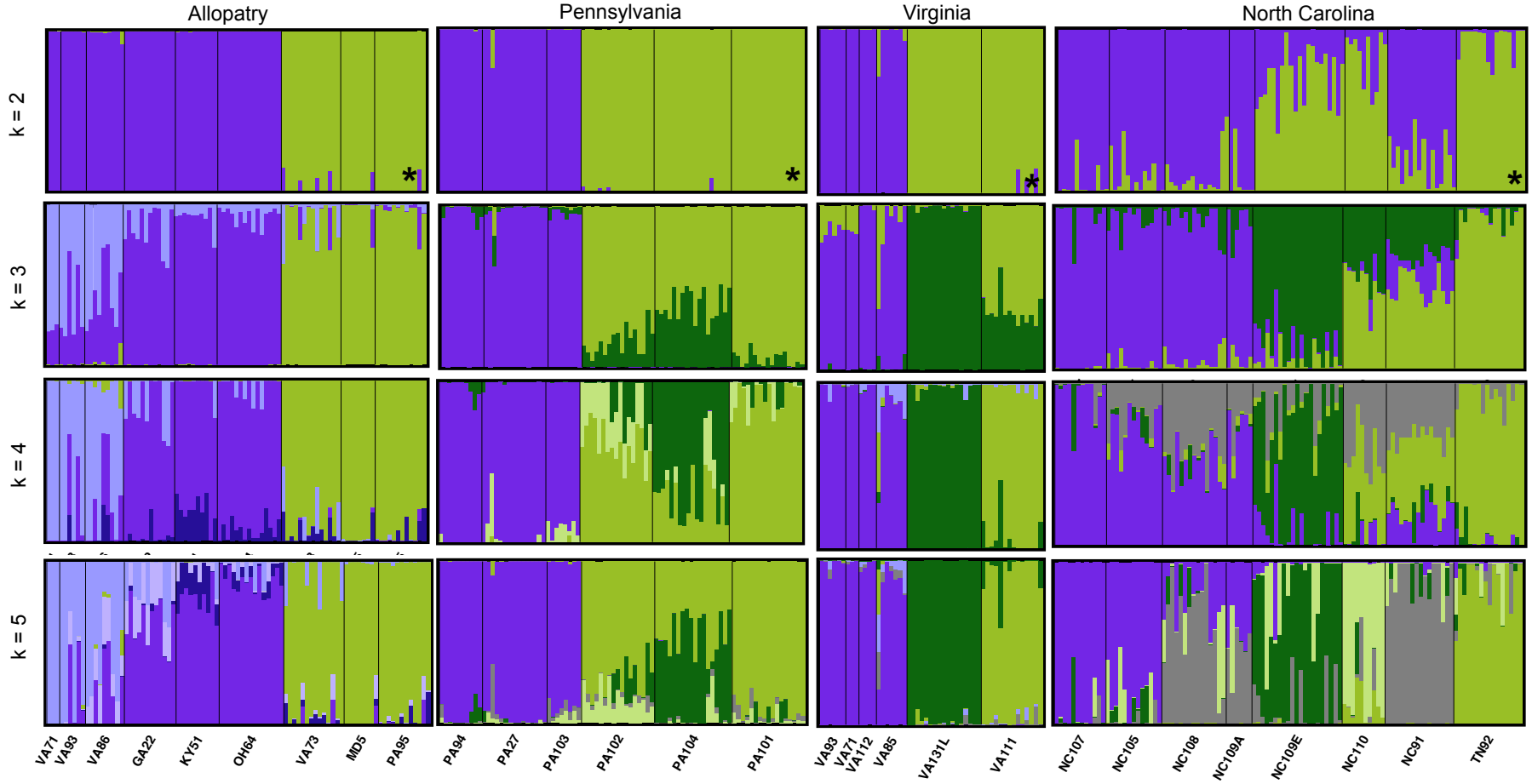


Figure S4. STRUCTURE plots for all k values 2-5 for the nuclear dataset of (from left to right) Allopatry, Pennsylvania, Virginia, and North Carolina. Best value of k for each region is marked with a *. Populations in each region are arranged by geography; populations closest to the other lineage geographically are closest to the other lineage on the STRUCTURE plots. Clusters are colored by the lineages of the populations in which they are most common: purple for Western and green for Appalachian. Clusters equally common in multiple lineages are colored grey.

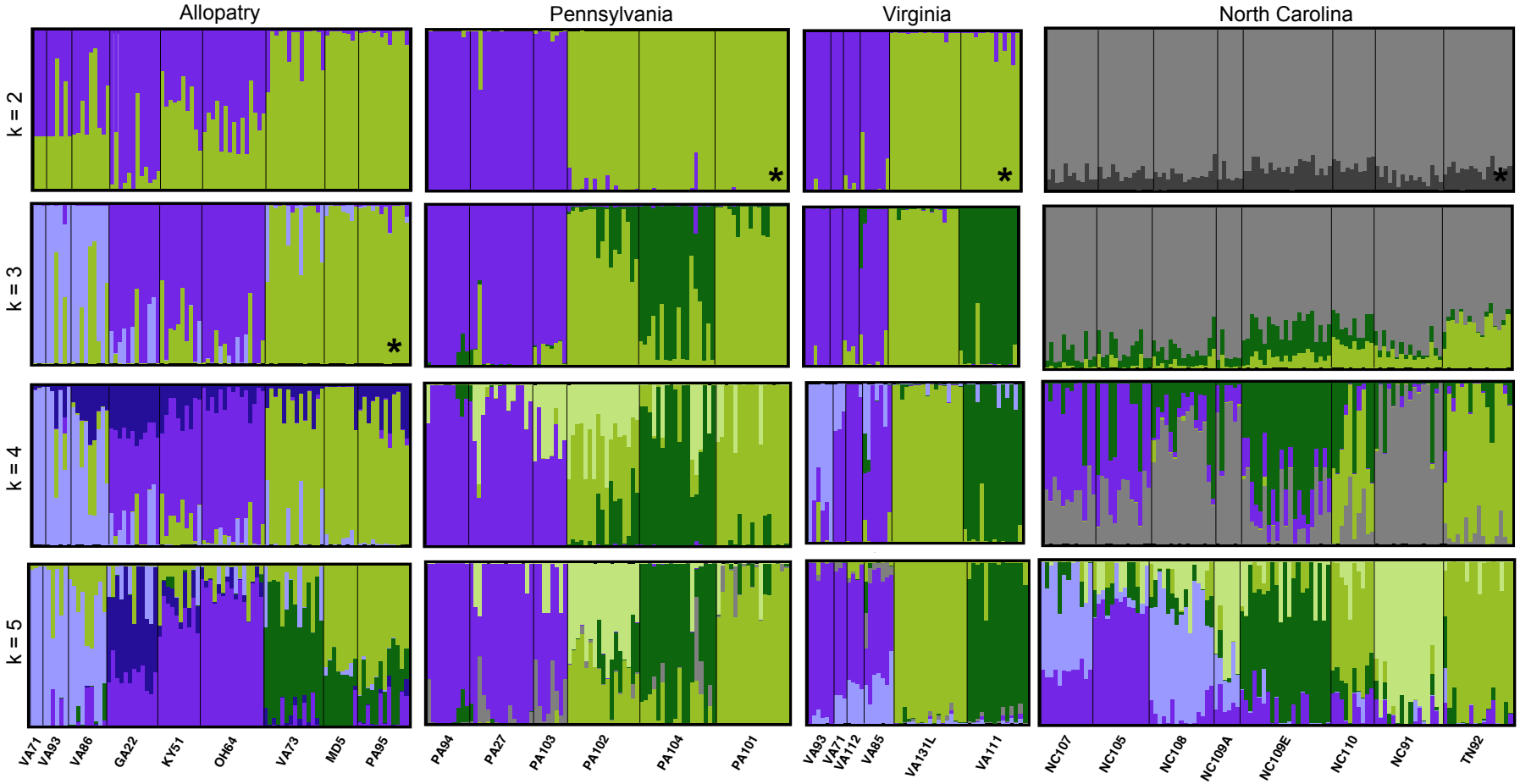


Figure S5. STRUCTURE plots for the best k value (5) for the nuclear dataset of (from left to right) Allopatry, Virginia, Pennsylvania, and North Carolina, created with a single model for all populations. Populations in each region are arranged by geography; populations closest to the other lineage geographically are closest to the other lineage on the STRUCTURE plots. Clusters are colored by the lineages of the populations in which they are most common: purple for Western and green for Appalachian. Clusters equally common in multiple lineages are colored grey.

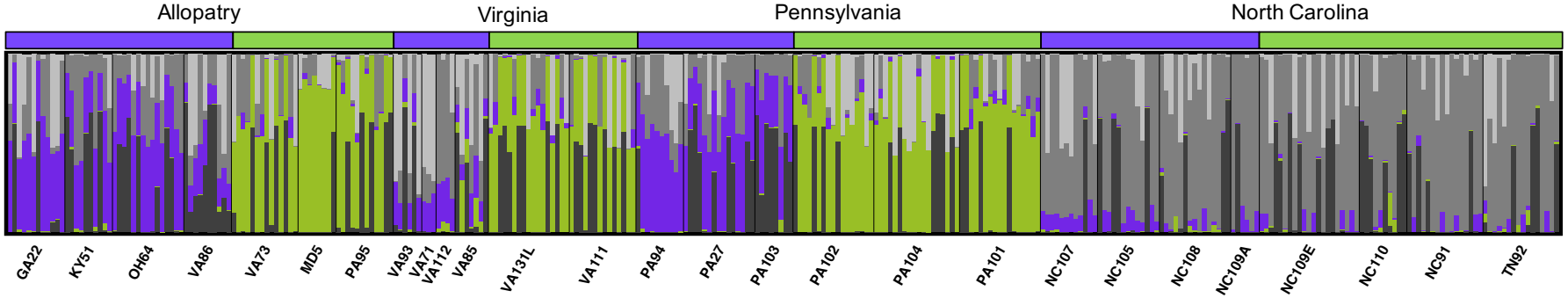


Figure S6. Schematic of adapter sequences and PCR indices used in library construction.

References

Alcaide, M., E. S. C. Scordato, T. D. Price, and D. E. Irwin. 2014. Genomic divergence in a ring species complex. Nature 511:83–85.

Bank, C., J. Hermisson, and M. Kirkpatrick. 2011. Can reinforcement complete speciation? Evolution 66:229–239.

Bank, C., R. Bürger, and J. Hermisson. 2012. The limits to parapatric speciation: Dobzhansky–Muller incompatibilities in a continent–island model. Genetics 191:845–863.

Barnard-Kubow, K. B., and L. F. Galloway. 2017. Variation in reproductive isolation across a species range. Ecol Evol 7:9347–9357.

Barnard-Kubow, K. B., C. L. Debban, and L. F. Galloway. 2015. Multiple glacial refugia lead to genetic structuring and the potential for reproductive isolation in a herbaceous plant. American Journal of Botany 102:1842–1853.

Barnard-Kubow, K. B., M. A. McCoy, and L. F. Galloway. 2017. Biparental chloroplast inheritance leads to rescue from cytonuclear incompatibility. New Phytol 213:1466–1476.

Barnard-Kubow, K. B., N. So, and L. F. Galloway. 2016. Cytonuclear incompatibility contributes to the early stages of speciation. Evolution 70:2752–2766.

Barton, N., and B. O. Bengtsson. 1986. The barrier to genetic exchange between hybridising populations. Heredity 56:357–376.

Baskin, J. M., and C. C. Baskin. 1984. The ecological life cycle of *Campanula americana* in northcentral Kentucky. Bulletin of the Torrey Botanical Club 111:329–337.

Bateson, W. 1909. Heredity and variation in modern lights. Pp. 85–101 *in* A. C. Seward, ed. Darwin and Modern Science. Cambridge, MA.

Bewick, E. R., and K. A. Dyer. 2014. Reinforcement shapes clines in female mate discrimination in *Drosophila subquinaria*. Evolution 68:3082–3094.

Bouckaert, R. 2017. bModelTest: Bayesian phylogenetic site model averaging and model comparison. BMC Evolutionary Biology 17:42.

Bouckaert, R., J. Heled, D. Kühnert, T. Vaughan, C.-H. Wu, D. Xie, M. A. Suchard, A. Rambaut, and A. J. Drummond. 2014. BEAST 2: A software platform for Bayesian evolutionary analysis. PLoS Comput Biol 10:e1003537–6.

Burgess, K. S., M. Morgan, L. Deverno, and B. C. Husband. 2005. Asymmetrical introgression between two Morus species (*M. alba*, *M. rubra*) that differ in abundance. Molecular Ecology 14:3471–3483.

Cain, M. L., V. Andreasen, and D. J. Howard. 1999. Reinforcing selection is effective under a relatively broad set of conditions in a mosaic hybrid zone. Evolution 53:1343–1353.

Catchen, J. M., A. Amores, P. A. Hohenlohe, W. A. Cresko, and J. H. Postlethwait. 2011. Stacks: Building and genotyping loci de novo from short-read sequences. G3 1:171–182.

Catchen, J. M., P. A. Hohenlohe, S. Bassham, A. Amores, and W. A. Cresko. 2013. Stacks: an analysis tool set for population genomics. Molecular Ecology 22:3124–3140.

Corbett-Detig, R. B., J. Zhou, A. G. Clark, D. L. Hartl, and J. F. Ayroles. 2013. Genetic incompatibilities are widespread within species. Nature 504:135–137.

Coyne, J. A., and H. A. Orr. 2004. Speciation. Sinauer Associates, Inc., Sunderland, MA, USA.

Crowl, A. A., N. W. Miles, C. J. Visger, K. Hansen, T. Ayers, R. Haberle, and N. Cellinese. 2016. A global perspective on Campanulaceae: Biogeographic, genomic, and floral evolution. American Journal of Botany 103:233–245.

Cutter, A. D. 2012. The polymorphic prelude to Bateson–Dobzhansky–Muller incompatibilities. TREE 27:210–219.

de Queiroz, K. 2007. Species concepts and species delimitation. Systematic Biology 56:879–886.

Dobzhansky, T. 1937. Genetics and the origin of species. Columbia University Press, New York.

Durand, E. Y., N. Patterson, D. Reich, and M. Slatkin. 2011. Testing for ancient admixture between closely related populations. Molecular Biology and Evolution 28:2239–2252.

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

Eaton, D. A. R., A. L. Hipp, A. González-Rodríguez, and J. Cavender-Bares. 2015. Historical introgression among the American live oaks and the comparative nature of tests for introgression. Evolution 69:2587–2601.

Evanhoe, L., and L. F. Galloway. 2002. Floral longevity in *Campanula americana* (Campanulaceae): A comparison of morphological and functional gender phases. American Journal of Botany 89:587–591.

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

Felsenstein, J. 1981. Skepticism towards Santa Rosalia, or Why are there so few kinds of animals? Evolution 35:124–138.

Field, D. L., D. J. Ayre, R. J. Whelan, and A. G. Young. 2010. Patterns of hybridization and asymmetrical gene flow in hybrid zones of the rare *Eucalyptus aggregata* and common *E. rubida*. Heredity 106:841–853.

Gadella, T. W. J. 1964. Cytotaxonomic studies in the genus Campanula. Wentia 11:1–104.

Galloway, L. F., and J. R. Etterson. 2005. Population differentiation and hybrid success in *Campanula americana*: geography and genome size. J. Evol. Biol. 18:81–89.

Galloway, L. F., J. R. Etterson, and J. L. Hamrick. 2003. Outcrossing rate and inbreeding depression in the herbaceous autotetraploid, *Campanula americana*. Heredity 90:308–315.

Hopkins, R., and M. D. Rausher. 2012. Pollinator-mediated selection on flower color allele drives reinforcement. Science 335:1090–1092.

Hülber, K., M. Sonnleitner, J. Suda, J. Krejčíková, P. Schönswetter, G. M. Schneeweiss, and M. Winkler. 2015. Ecological differentiation, lack of hybrids involving diploids, and asymmetric gene flow between polyploids in narrow contact zones of *Senecio carniolicus* (syn. *Jacobaea carniolica*, Asteraceae). Ecol Evol 5:1224-1234.

Irwin, D. E. 2005. Speciation by Distance in a Ring Species. Science 307:414–416.

Irwin, D. E., J. H. Irwin, and T. D. Price. 2001. Ring species as bridges between microevolution and speciation. Genetica 112-113:223–243.

Jakobsson, M., and N. A. Rosenberg. 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics 23:1801–1806.

Lepais, O., R. J. Petit, E. Guichoux, J. E. Lavabre, F. Alberto, A. Kremer, and S. Gerber. 2009. Species relative abundance and direction of introgression in oaks. Molecular Ecology 18:2228–2242.

Ley, A. C., and O. J. Hardy. 2017. Hybridization and asymmetric introgression after secondary contact in two tropical African climber species, *Haumania danckelmaniana* and *Haumania liebrechtsiana (*Marantaceae). Int. J. Plant Sci. 178:421–430.

Lindtke, D., and C. A. Buerkle. 2015. The genetic architecture of hybrid incompatibilities and their effect on barriers to introgression in secondary contact. Evolution 69:1987–2004.

Mandeville, E. G., T. L. Parchman, D. B. McDonald, and C. A. Buerkle. 2015. Highly variable reproductive isolation among pairs of *Catostomus* species. Molecular Ecology 24:1856–1872.

Mansion, G., G. Parolly, A. A. Crowl, E. Mavrodiev, N. Cellinese, M. Oganesian, K. Fraunhofer, G. Kamari, D. Phitos, R. Haberle, G. Akaydin, N. Ikinci, T. Raus, and T. Borsch. 2012. How to handle speciose clades? Mass taxon-sampling as a strategy towards illuminating the natural history of *Campanula* (Campanuloideae). PLOS One 7:e50076–24.

Matute, D. R. 2010. Reinforcement can overcome gene flow during speciation in *Drosophila*. Current Biology 20:2229–2233.

Matute, D. R., J. Gavin-Smyth, and G. Liu. 2014. Variable post-zygotic isolation in *Drosophila melanogaster*/*D. simulans* hybrids. J. Evol. Biol. 27:1691–1705.

McDermott, S. R., and M. A. F. Noor. 2011. Genetics of hybrid male sterility among strains and species in the *Drosophila pseudoobscura* species group. Evolution 65:1969–1978.

Muller, H. J. 1942. Isolating mechanisms, evolution, and temperature. Biological Symposia 6:71-125.

Natalis, L. C., and R. A. Wesselingh. 2012. Post-pollination barriers and their role in asymmetric hybridization in *Rhinanthus* (Orobanchaceae). American Journal of Botany 99:1847–1856.

Noor, M. A. F. 1999. Reinforcement and other consequences of sympatry. Heredity 83:503–508.

Paris, J. R., J. R. Stevens, and J. M. Catchen. 2017. Lost in parameter space: a road map for stacks. Methods Ecol Evol 188:799–14.

Patterson, N., P. Moorjani, Y. Luo, and S. Mallick. 2012. Ancient admixture in human history. Genetics 192:1065-1093.

Peter, B. M. 2016. Admixture, population structure, and F-statistics. Genetics 202:1485–1501.

Peterson, B. K., J. N. Weber, E. H. Kay, H. S. Fisher, and H. E. Hoekstra. 2012. Double Digest RADseq: an inexpensive method for se novo SNP discovery and genotyping in model and non-model species. PLOS One 7:e37135.

Pickrell, J. K., and J. K. Pritchard. 2012. Inference of population splits and mixtures from genome-wide allele frequency data. PLoS Genet 8:e1002967–17.

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

Rieseberg, L. H., and D. E. Soltis. 1991. Phylogenetic consequences of cytoplasmic gene flow in plants. Evolutionary Trends in Plants 5:65–84.

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

Seehausen, O., G. Takimoto, D. Roy, and J. Jokela. 2008. Speciation reversal and biodiversity dynamics with hybridization in changing environments. Molecular Ecology 17:30–44.

Stamatakis, A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30:1312–1313.

Tiffin, P., S. Olson, and L. C. Moyle. 2001. Asymmetrical crossing barriers in angiosperms. Proceedings of the Royal Society B: Biological Sciences 268:861–867.

Yukilevich, R. 2012. Asymmetrical patterns of speciation uniquely support reinforcement in Drosophila. Evolution 66:1430–1446. Blackwell Publishing Inc.

CHAPTER 3:

Prezygotic reproductive isolation varies among contact zones between incipient species

Abstract

Prezygotic reproductive isolating barriers are of particular importance to the process of speciation because during secondary contact they can either be reduced as lineages merge or increase via reinforcement. In *Campanula americana*, two partially isolated lineages experience low postzygotic reproductive isolation and high gene flow in a North Carolina contact zone, and high postzygotic reproductive isolation and low gene flow in a contact zone in Pennsylvania. In this study, I tested whether prezygotic isolating barriers in *C. americana* followed patterns consistent with lineages merging in North Carolina and with reinforcement in Pennsylvania. I found no evidence of prezygotic reproductive isolation in North Carolina; isolation found in allopatry due to pollen-style interactions is not present at that contact zone. Conversely, I find that all post-pollination reproductive barriers that are present in allopatry are also present in the Pennsylvania contact zone. In addition, the lineages are phenologically separated by two weeks in Pennsylvania, while they flower at the same time elsewhere across the range. In total, this study demonstrates differences in prezygotic isolation following secondary contact at two contact zones. Where lineages are merging, prezygotic isolation was not present, but where postzygotic isolation is high there was an increase in prezygotic isolation, showing that the differences across two contact zones between the same lineages can be important enough to change the outcome of contact.

Introduction

Early acting reproductive barriers can have the strongest effect in separating two groups of organisms because they filter the opportunity for hybridization at later stages (Coyne and Orr 2004). For this reason, prezygotic isolating barriers can be particularly important to the process of speciation. Prezygotic barriers are those that act before the male and female gametes join, and can range from geographic isolation to post-mating barriers that prevent the fusion of gametes from different groups (Coyne and Orr 2004). In the classic allopatric model of speciation, a geographic barrier arises, separating a single species into groups that then evolve as independent lineages. During geographic isolation, additional pre- or postzygotic genetic barriers may arise via drift or selection (Bateson 1909; Dobzhansky 1937; Muller 1942). If the geographic barrier disappears, or the lineages expand around it before genetic barriers completely isolate them, secondary contact between the lineages can allow them to either merge back together or to diverge more quickly, making secondary contact an important and dynamic stage in the process of speciation.

Secondary contact between partially isolated lineages can have disparate outcomes. In some cases, selection can favor individuals without the alleles that cause incompatibility (Felsenstein 1981; Barton and Bengtsson 1986; Taylor et al. 2005; Kearns et al. 2018), purging reproductive isolation from contact zones and causing the lineages to merge together. In other cases selection favors traits that prevent hybridization and the production of low-quality offspring. This development of prezygotic barriers may occur in one or both of the lineages, resulting in accelerated divergence at contact zones (Dobzhansky 1937; Noor 1999; Matute 2010; Hopkins 2013). The process of accumulating additional barriers at contact zones is termed reinforcement. For plants, these barriers can include phenological differences (McNeilly and

Antonovics 1968; Silvertown et al. 2005), pollinator-mediated floral differences (Grossenbacher and Whittall 2011; Hopkins and Rausher 2012), increased selfing rates (Fishman and Wyatt 1999), and post-pollination prezygotic barriers such as pollen-style incompatibilities (Kay and Schemske 2008). If reinforcement is occurring, prezygotic barriers are expected to be stronger at the contact zone than elsewhere in the range (Hopkins 2013).

Campanula americana is an herbaceous plant growing in eastern North America. It is divided into two lineages, an Appalachian and a Western lineage, that began diverging 2-7 million years ago (Barnard-Kubow et al. 2015). During the Pleistocene, the Appalachian lineage was confined to small microrefugia within the Appalachian mountains, and the Western lineage was confined to refugia along the Gulf and Atlantic coasts (Barnard-Kubow et al. 2015). Over the past 20,000 years, the Appalachian lineage has spread throughout the Appalachian Mountains while the Western lineage has expanded north and west. These lineages are isolated from each other by reductions in hybrid germination, survival, and fruit production (Barnard-Kubow et al. 2016; Barnard-Kubow and Galloway 2017). Today, the lineages are in contact in North Carolina, Pennsylvania, and Virginia. At the North Carolina contact zone, gene flow is high (Chapter 2) and postzygotic reproductive isolation is lower than it is in allopatry (Chapter 1). In the Pennsylvania contact zone, gene flow is low (Chapter 2) and postzygotic reproductive isolation is similar to allopatry (Chapter 1). However, little is known about prezygotic reproductive isolation between these lineages.

Postzygotic reproductive isolation between Appalachian and Western lineages of *C. americana* is low in a North Carolina contact zone, but remains high in a Pennsylvania contact zone (Chapter 1, 2). If prezygotic isolation is also present in *C. americana*, I predict that in North Carolina it will follow a similar pattern to postzygotic reproductive isolation and be lower

relative to allopatry. In Pennsylvania, I expect that prezygotic reproductive isolation will be equal to or higher than it is in allopatry. In this study, I test whether flowering phenology, floral phenotype, pollinator constancy, and pollen-style interactions contribute to prezygotic isolation between the Appalachian and Western lineages across differentiated contact zones in *C. americana*.

Methods

Study system

Campanula americana is an herbaceous autotetraploid that grows in eastern North America. It has an annual/biennial life history, in which seeds germinate in the spring or fall and rosettes must overwinter before bolting and flowering in June-August the following summer. It is pollinated by generalist bees, including bumblebees, megachilids, and other small bees, with bumblebees serving as the most effective pollinators (Lau and Galloway 2004; Koski et al. 2017; 2018). *Campanula americana* is comprised of two main lineages: one present throughout the Appalachian Mountains (hereafter, the Appalachian lineage), and the other present throughout its range west of the Appalachians (hereafter, the Western lineage). These groups are reproductively isolated by low germination and fruit production of hybrids, and low hybrid survival when the Western lineage is the mother (Barnard-Kubow et al. 2016). The lineages are in contact in two areas of the range: North Carolina and Pennsylvania (Fig. 1; Chapter 2 Fig. 1A). However, the germination incompatibility between the lineages is not present at the North Carolina contact zone (Chapter 1). The nuclear genome of the lineages is highly mixed in North Carolina, indicating gene flow between the lineages there, but gene flow is minimal at the Pennsylvania contact zone (Chapter 2).

To test whether prezygotic barriers contribute to isolation between the lineages of *C. americana*, I grew two cohorts of plants: one in 2016 and one in 2018. For each cohort, I selected several populations from each lineage (Appalachian and Western) in each of three regions: allopatry (4 populations of each lineage), North Carolina contact zone (6 populations of each lineage), and Pennsylvania contact zone (3 populations of each lineage; Table S1). Seeds were grown in a controlled environment. I planted at least 20 seeds per population (705 in 2016, 818 in 2018), representing as many families as possible (mean 21 for 2016, mean 19.5 for 2018). Additional seeds were planted if the germination rate of a population was low. Seeds were planted singly in a 3:1 mixture of Metromix and Turface, and germinated in a growth chamber set at 21°C day, 14°C night with 12 hour days for 44 days for the 2016 cohort, and 56 days for the 2018 cohort. I then moved plants to a cold room to vernalize for 82 days for the 2016 cohort, and 56 days for the 2018 cohort at 5°C with 12 hour days. After vernalization, I transplanted rosettes into conetainers, arranged them in random order, and moved them into the greenhouse where lights extended day length to 16 hours. Plants were fertilized every other week until bolting, then weekly, and watered as needed. Plants from the 2016 cohort were crossed with other plants within their population, then offspring of those crosses were germinated and vernalized as stated above, and transplanted to a common garden in the field near a natural population of *C. americana*.

Phenology

I used plants from the 2016 cohort to determine whether flowering time differences contribute to reproductive isolation between the lineages. I scored days until flowering of the first generation of plants in the greenhouse as the number of days between when the plants were removed from the cold until the first day a flower opened on a plant, to approximate the time

until flowering after the end of winter. For the second generation of 2016 plants that were transplanted into the field, I used reproductive phenology as a proxy for day of first flower. I counted the number of flowers and fruits on each plant in mid-August when almost all plants had flowered, and estimated relative phenology by calculating the proportion of reproductive structures (flowers + fruits) that were fruits. A higher proportion of fruits per reproductive structure means that the plant began flowering relatively earlier. I analyzed day of first flower and relative phenology data using a generalized linear mixed model with a Poisson distribution and a log link, using lineage, region, and their interaction as fixed effects, and population nested within the interaction of lineage and region as a random effect.

Floral phenotype

Floral phenotypes can influence prezygotic reproductive isolation via pollinator attraction. I tested for differences in floral phenotype between lineages and among regions by measuring petal color, flower size, pollen darkness, and style color. Petal color was measured using a spectrometer to record the wavelengths of light reflected by petals from the 2016 and 2018 cohorts. I extracted UV chroma and blue chroma values from spectra using the software CLR (version 1.05, Mongomerie 2008). Pollen darkness was measured by applying pollen from 2018 cohort plants to a clean piece of black felt, and categorizing the color as white, tan, light purple, purple, or dark purple, then converting these to ordinal values from 1-5 (cf. Ison et al. 2018). Flower diameter was measured on 2018 cohort plants as the distance across a flattened flower from petal tip to petal tip. Style darkness was measured by removing all pollen from the style, then scoring the style as white, light purple, or purple, then converting these to ordinal values from 0-2. Presence of yellow pigment on the style was measured by removing all pollen from the style, then scoring the presence of yellow pigment as 1 and its absence as 0. I ran linear

mixed models on flower diameter, pollen darkness, and style darkness, using lineage, region, and their interaction as fixed effects, and population nested within the interaction of lineage and region as a random effect. I ran generalized linear mixed models on petal UV chroma and Blue chroma with a beta distribution and a logit link, using the same factors as the models for the other floral phenotypes. Style yellow did not have enough variation for a model to converge; nearly all Appalachian plants had yellow styles, and almost no Western plants did.

Pollination arrays

Pollination arrays were created using pairs of Western and Appalachian lineage *C. americana* populations. Nine population pairs were tested, three from each region. In each array, one lineage was in the majority (75%) and the other in the minority (25%), such that arrays had 12 plants of one lineage and four of the other. This was done to mimic contact zones where members of one lineage are typically in the majority. For each pair of populations, six arrays were created, three in which one population was in the majority and three in which that same populations was in the minority. The resulted in 54 arrays (9 pairs of populations *2 frequencies/population pair * 3 arrays/frequency). Within each array Appalachian and Western lineages were distributed evenly throughout the array (Fig. S1).

Arrays were exposed to pollinators at Mountain Lake Biological Station, within five km of natural *C. americana* populations. Arrays were placed outside in the afternoon and then left for approximately 48 hours. The number of flowers on each plant was counted each day. Observations began the second day to give visiting insects time to discover them. Arrays were then observed daily for at least 15 minutes from 1-4pm. Pollinators were classified as small, medium, halictid, megachilid, bumblebee, or other (Fig. S3A). One pollinator was watched at a time. Transitions were observed between flowers on a plant and between plants for up to 10 plant

transitions or until the pollinator left the array. At that point, a new pollinator was observed. Representative samples of pollinators were collected for identification. Plant to plant transitions were analyzed to determine if pollinator preference could lead to prezygotic isolation through assortative mating. I observed a total of 582 pollinators that performed 3,408 transitions between plants.

The data were analyzed using a generalized linear mixed model with a binary distribution and a logit link on only the transitions starting on the majority lineage, using transition type (within or between lineages) as the response variable, and region (allopatric, North Carolina contact zone, Pennsylvania contact zone), starting lineage (Appalachian or Western), pollinator type (bumblebee or other), and all interactions thereof as main effects. Starting population (nested within the interaction of starting lineage and region) and array were included as random effects. In addition, since bumblebees are the most effective pollinators (Koski et al. 2018) and performed the most transitions by far (Fig. S3A), I ran a model on only the transitions performed by bumblebees with the same model except omitting pollinator type and all its interactions as factors. I also ran a model with the same settings as the bumblebee-only model, that included the proportion of flowers on other plants in the array that were in the same lineage as the starting plant as a covariate.

Post-pollination barriers

I tested for post-pollination prezygotic barriers to reproduction in two ways using the 2018 cohort of plants. First, I compared how quickly within- and between-lineage pollen fertilized ovules. Second, I determined how many seeds per fruit were produced by within- and between-lineage crosses. To do this, I emasculated flowers on “maternal” plants, then applied either within-population or between-lineage, within-region pollen to the stigma. To compare the

speed of fertilization, I cut the styles after six hours approximately 8mm below the point at which the stigmatic lobes separate. Cutting the style meant that only pollen tubes that had grown beyond that point in six hours could fertilize ovules. I left some styles uncut to determine potential seed production. Fruits were harvested when mature and the number of seeds counted. Three crosses were conducted per population pair. Seed number was analyzed using a generalized linear mixed model with a Poisson distribution and a log link, using region, lineage, cross type (between- or within-lineage), and all interactions as fixed effects, and maternal population as a random effect.

Results

Phenology

Both greenhouse-grown and field-planted individuals from Western and Appalachian lineages flowered at similar times in populations from allopatry and from North Carolina, but Western lineage plants from Pennsylvania populations flowered earlier than Appalachian lineage plants (Table 1, Fig. 2). In the greenhouse, Appalachian lineage plants from Pennsylvania flowered an average of 77 days after being moved from the cold room to the greenhouse, while Western lineage plants began flowering an average of two weeks earlier, 63 days after being moved to the greenhouse (Fig. 2A). In the field, though not significant, the same pattern was found with Appalachian lineage plants from Pennsylvania having 49.4% fewer fruits per reproductive structure in mid-August than Western lineage plants from the same region (Table 1, Fig. 2B). In contrast, Western and Appalachian lineage plants from North Carolina and from allopatry had similar reproductive phenology (Table 1).

Floral phenotype

Most floral phenotypes varied across regions, lineages, or both (Table 1). Western lineage flowers were larger than Appalachian lineage flowers, and Allopatric flowers were larger than North Carolina flowers (Table 1, Fig. S2). Pollen from Appalachian lineage populations was darker in than Western lineage populations in allopatry and Pennsylvannia, while in North Carolina pollen from Western lineage populations was darker (Table 1, Fig. S2). Western lineage styles were darker than Appalachian lineage styles, however Appalachian lineage styles in allopatric populations were darker than Appalachian lineage styles elsewhere (Table 1, Fig. S2). Appalachian lineage populations almost always had yellow or green pigment on their styles, while Western lineage plants almost never did (Fig. S2). Flower petals of Appalachian lineage populations reflected more UV light and less blue light than Western lineage populations (Table 1, Fig. S2). However, petals reflected similar amounts of UV and blue light across the range (Table 1, Fig. S2).

Pollination arrays

In Allopatry and Pennsylvania, *Bombus* transitioned between plants at random, and were slightly more likely to remain on Appalachian lineage plants than transition to Western lineage plants than would be expected based on overall array plant proportion or neighboring plant proportion (Table 2, Fig. 3A). However, in populations from North Carolina, pollinators were more likely to transition away from Appalachian lineage plants when they were in the majority than expected based on overall array plant proportion or neighboring plant proportion. Similarly, pollinators were more likely to remain within the Western lineage when they were in the majority. These results indicate a preference of pollinators for Western lineage plants from North Carolina relative to Appalachian lineage plants (Table 2, Fig. 3A). Similar overall patterns were found when using all pollinators (Table S2, Fig. S3B). However, this preference is likely

associated with flower number. Appalachian lineage plants produced 17% more flowers than Western lineage plants in allopatry and Pennsylvania but in in North Carolina, Appalachian lineage plants produce 50% fewer flowers than Western lineage plants (Table 2, Fig. 3B). When the proportion of flowers in an array that were within-lineage was used as a covariate, it accounts for flower number contributions to pollinator preference, and there is no longer evidence of preference (Table 2, Fig. 3C).

Pollen-style interactions

Styles cut at six hours produced fewer seeds per fruit than those that were not cut (Fig. 4), therefore seeds produced in the cut treatment are referred as "early" seeds. More early seeds were produced for crosses within the same lineage than crosses between lineages on Allopatric and Pennsylvania Western lineage plants (Table 3, Fig. 4A). Appalachian lineage plants from North Carolina populations produced more early seeds when crossed with local Western lineage plants than when crossed with plants of their same lineage (Table 3, Fig. 4A). In all other combinations, within-lineage crosses had similar numbers of early seeds as between-lineage crosses (Table 3, Fig. 4A).

When styles were not cut, within-lineage crosses set more seeds than between-lineage crosses (Table 3, Fig. 4B), indicating overall reproductive isolation in seed set due to pollen-style interactions. However, the magnitude of this pattern differed among lineages and regions. Within-lineage crosses produced significantly more seeds than between-lineage crosses in Appalachian lineage Allopatric populations and Western lineage Pennsylvania populations (Table 3B). The difference in seed set in Appalachian lineage populations in Pennsylvania between within- and between-lineage crosses was marginally significant ($p = 0.0605$).

Discussion

Early acting barriers have the greatest effect on reproductive isolation (Coyne and Orr 2004). In this study, I found that there is prezygotic reproductive isolation between the Appalachian and Western lineages of *Campanula americana* due to pre-pollination barriers (flowering phenology, Table 1, Fig. 2) and post-pollination barriers (pollen-style interactions, Table 3, Fig. 4). However, I found limited evidence that pollinators contribute to assortative or disassortative mating which could influence hybridization (Table 2, Fig. 3). Patterns of prezygotic reproductive isolation vary across the range. The North Carolina contact zone has fewer prezygotic isolating barriers than allopatric populations; it lacks reproductive isolation due to post-pollination barriers (Fig. 4, Table 3B). In contrast, the Pennsylvania contact zone has more prezygotic barriers than allopatry; Western plants flower two weeks earlier than Appalachian plants (Fig. 2, Table 1B). The difference between the contact zones supports previous results (Chapters 1 and 2) that the lineages may be merging in North Carolina, and that reinforcement may accelerate divergence between lineages in Pennsylvania.

Flowering phenology is one of the earliest stages of isolation and in *C. americana* is a likely prezygotic barrier in the Pennsylvania contact zone. Flowering phenology does not differ between Western and Appalachian lineages in allopatry or in a contact zone in North Carolina. However, there were substantial differences between the lineages in populations from a contact zone in Pennsylvania. Specifically, plants from the Western lineage began flowering two weeks earlier on average than plants from the Appalachian lineage. The difference in flowering time is present under controlled greenhouse conditions as well as in a field common garden, and is therefore likely controlled genetically. *Campanula americana* performs the majority of its flowering over the course of approximately four weeks (Haggerty and Galloway 2010), so a two

week shift in initial date of flowering does not completely isolate the two groups, but it does reduce temporal overlap of flowering by half. It allows Western lineage plants, which have chloroplasts incompatible with Appalachian lineage plants (Barnard-Kubow et al. 2016), to flower in the absence of Appalachian lineage pollen at the beginning of their flowering season. Flowering phenology is a common mechanism for reinforcement (McNeilly and Antonovics 1968; Silvertown et al. 2005; Hopkins 2013), likely because it acts early to prevent any chance of hybridization. It is therefore an especially strong isolating barrier. It is possible that the divergence in flowering time in Pennsylvania is due to reinforcement.

Floral traits and pollinator response to any differentiation in those traits do not appear to contribute to reproductive isolation between lineages of *C. americana*. Although most traits measured were similar between Appalachian and Western lineage populations, petal UV reflectance, style darkness, and the presence of green or yellow pigment on the style differ between lineages. Differences in floral phenotype between lineages can mediate reproductive isolation by influencing pollinator preference (Kay and Sargent 2009; Hopkins and Rausher 2012; Brothers and Atwell 2014). However, the only pollinator preference detected here was in populations from the North Carolina contact zone where pollinators preferred Western lineage plants over Appalachian lineage plants. This preference appeared to be mediated by a larger number of flowers per plant in North Carolina Western lineage than Appalachian lineage plants. Preference for Western plants could contribute to prezygotic reproductive isolation if Appalachian lineage plants were rare, because pollinators would infrequently transition from the common Western lineage plants to the rare Appalachian lineage plants. However, for the same reason, increased gene flow may result if Appalachian lineage plants are common and Western lineage plants are rare. Pollination by generalists such as bumblebees can make it less likely that

pollinators will drive divergence (Waser et al. 1996), which may be one explanation why I find no prezygotic isolation due to pollinators in *C. americana*. In total, differences in floral traits between Western and Appalachian lineages did not appear to contribute to prezygotic isolation.

Pollen-style interactions contribute to prezygotic isolation between Appalachian and Western lineages of *C. americana*. Plants from allopatric populations make more seeds per fruit if the pollen came from the same lineage when their styles were cut six hours post-pollination. This indicates that same-lineage pollen germinates or grows more quickly than between-lineage pollen. Pollen competition is frequently important in the wild (Stephenson and Bertin 1983; Delph 2019), and would result in more ovules being pollinated by same-lineage pollen than between-lineage pollen if both types are deposited at the same time. Therefore, there pollen-style interactions contribute to prezygotic isolation in allopatric *C. americana* populations. This isolation is exaggerated in Western-lineage plants from a Pennsylvania contact zone. However, it disappears in Western-lineage plants from a North Carolina contact zone, and reverses in Appalachian-lineage plants from both contact zones. This indicates that pollen-style interactions do not contribute to isolation and may even enhance hybridization in Appalachian lineage populations in North Carolina. These populations have high hybrid fitness because of increased fruit production (Chapter 1), so the enhancement of hybridization due to the pollen-style interactions investigated in this study may facilitate advantageous hybridization in North Carolina. When pollen competition is not important (e.g. Koski et al. 2017), comparable to the uncut styles, plants from both lineages make more seeds per fruit with same-lineage pollen than they do with between-lineage pollen, indicating that there is some prezygotic isolation in seed production across the range. Pollen-style incompatibility has been demonstrated to facilitate reinforcement in other systems (Kay 2006; Kay and Schemske 2008), and the increase of

reproductive isolation due to pollen-style interactions in Pennsylvania relative to allopatry may indicate that it is also facilitating reinforcement in *C. americana*.

Concordant with my predictions, prezygotic reproductive isolation is lower in the North Carolina contact zone than it is in allopatry. This contact zone has lower postzygotic reproductive isolation relative to allopatry (Chapter1) and higher gene flow (Chapter 2). Together with my findings here, this suggests that the Appalachian and Western lineages are merging together in North Carolina by purging reproductive isolation, and thereby increasing gene flow. Prezygotic reproductive isolation is high in Pennsylvania relative to allopatry. In this contact zone, postzygotic reproductive isolation is also high (Chapter 1) and gene flow is low (Chapter 2). Indeed, all prezygotic isolating barriers that are present in allopatry are also present in Pennsylvania, and an additional barrier due to flowering time is found at the contact zone. This suggests that reinforcement may be acting in Pennsylvania, mediated by flowering time differences. The variation I found in outcome of secondary contact between the contact zones of *C. americana* suggests that the amount of variation found within species can be important enough to change the trajectory of the speciation process.

Acknowledgements

I thank Karen Barnard-Kubow, Rebecca Risser, Judy Debban, Paul Debban, Jeanette Engel, Susan Munch, and Marge Van Tassel for help finding populations of *Campanula americana* for this study. I also thank Jeffrey and Rebecca Risser for assistance planting the common garden experiment for the second generation of the 2016 cohort plants. I thank Linnea Fraser for her assistance with pollination arrays and crossing. This work was supported by an

NSF GRFP, the Margaret Walton Scholarship for Mountain Lake, and the NSF REU-Site grant to Mountain Lake Biological Station.

Table 1. (A) Analysis of variance comparing phenology and floral phenotype of populations from Appalachian and Western lineages sampled from three regions: allopatry and contact zones in North Carolina and Pennsylvania. Population nested within the interaction of region and lineage was included as a random effect (not shown). (B) Lineages compared within contact zones to understand Region*Lineage effects. F-values are given with * p<0.05, ** p<0.01, *** p<0.001.

A

		Phenology		Floral phenotype				
Effect	**Num DF**	**Greenhouse**	**Common Garden**	**Flower diameter**	**Pollen darkness**	**Style darkness**	**UV Chroma**	**Blue Chroma**
Region	2	4.2*	0.1	6.54*	7.18**	4.75*	2.25	1.44
Lineage	1	7.19*	0.03	6.93*	0.1	130.69***	7.35*	4.02*
Region*Lineage	2	5.09*	2	1.29	9.58**	4.62*	0.63	1.34
Den DF		23	15	12	12	12	24	24

B

Phenology Comparison	**Num DF**	**Greenhouse**	**Common Garden**
PA-West vs PA-App	1	13.89**	2.72
NC-West vs NC-App	1	0.03	0.32
Allo-West vs Allo-App	1	0.27	0.96
Den DF		23	15

Table 2. Analysis of variance of the frequency of between-lineage bumblebee transitions starting from the lineage that was the most common in an array. Arrays were composed of a mix of plants from the Appalachian and Western lineages with a 75% frequency of one lineage or the other. Arrays included populations sampled from allopatry or contact zones in North Carolina and Pennsylvania. A pair of populations nested within the interaction of region and lineage was included as a random effect (not shown). Flower number per plant in arrays was analyzed with a comparable model. Finally, the frequency of between-lineage bumblebee transitions was analyzed with the proportion of flowers available for transitions that were within lineage as a covariate to account for differences in the number of flowers per plant between populations in the two lineages. F-values are given with * $p<0.05$, ** $p<0.01$, *** $p<0.001$.

Effect	Num DF	Flower Number	Without Covariate	With Covariate
Starting lineage[a]	1	14.28***	2.48	0.05
Region[a]	2	6.4*	0.29	0.15
Starting lineage*Region[a]	2	36.74***	7.47**	0.12
Within-lineage flower proportion (WLFP)[b]	1			8.1**
WLFP*Starting lineage[b]	1			0.04
WLFP*Region[b]	2			0.15
WLFP*Starting lineage*Region[b]	2			0.17
Den DF[a]		418	11	11
Den DF[b]				1369

Table 3. Analysis of variance to evaluate prezygotic reproductive isolation due to pollen-style interactions. Seed production in flowers with styles cut six hours after pollination and intact flowers that had been pollinated with a plant from its same population (Within) or a population from the other lineage (Between) were compared. (A) Plants were from populations in the Appalachian or Western lineages sampled from three regions including allopatry and contact zones in North Carolina and Pennsylvania. Population nested within the interaction of region and lineage was included as a random effect (not shown). (B) Exploration of Region*Lineage*Cross Type effects by comparing cross type (within or between) within contact zone and lineage. F-values are given with + p<0.1, * p<0.05, ** p<0.01, *** p<0.001.

A

Effect	**Num DF**	**6 hour cut**	**Uncut**
Region[b]	2	0.84	0.12
Maternal lineage[b]	1	0.24	2.04
Cross Type[a]	1	13.51***	5.02*
Region*Maternal lineage[b]	2	1.67	0.12
Region*Cross Type[a]	2	76.34***	0.65
Maternal lineage*Cross Type[a]	1	184.87***	0.28
Region*Maternal lineage*Cross Type[a]	2	5.68**	4.3*
Den Df[a]		219	39
Den Df[b]		11	8

B

Comparison: Within vs Between	**Num DF**	**6 hour cut**	**Uncut**
Allopatry, Appalachian	1	3.97*	32.41***
Allopatry, Western	1	53.03***	2.82
NC, Appalachian	1	101.55***	0.21
NC, Western	1	1.79	1.54
PA, Appalachian	1	9.23**	3.74+
PA, Western	1	165.99***	6.57*
Den DF		219	39

Figure 1. (A) Locations of *Campanula americana* populations used to evaluate prezygotic isolation. Array testing location marked with a star.

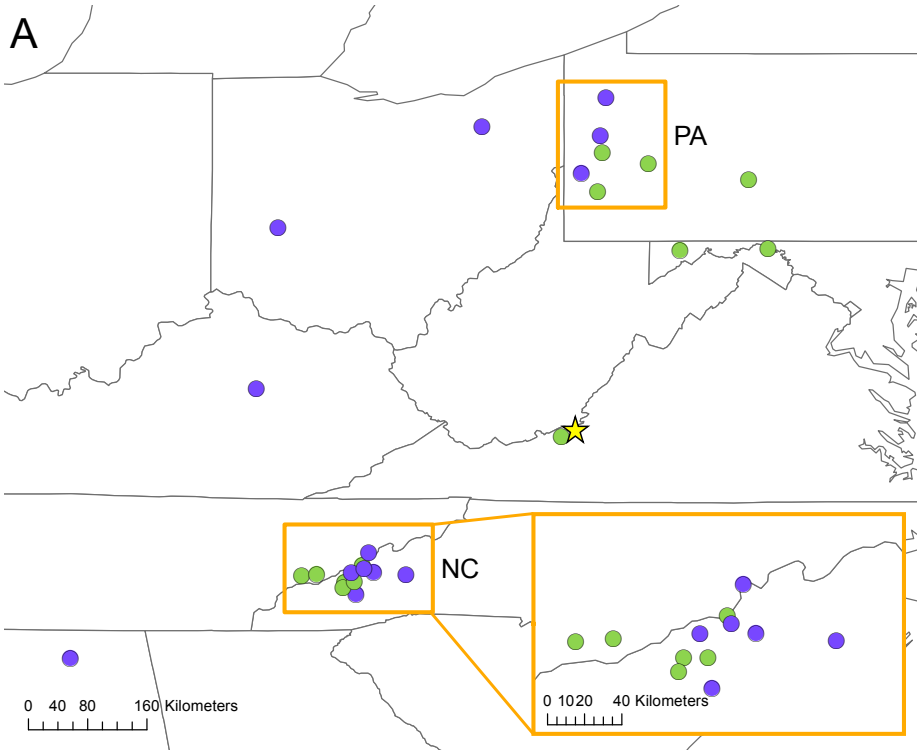


Figure 2. Least squares means of floral phenology in greenhouse (A) and field common garden (B). In the greenhouse, phenology was measured as the number of days between when the plants were vernalized and the first day a plant produced a flower. Common garden phenology was approximated by calculating the proportion of reproductive structures (flowers and fruit) that were fruit; higher numbers indicate early flowering and low numbers indicate late flowering. Contrasts between neighboring bars are given with ** $p<0.01$.

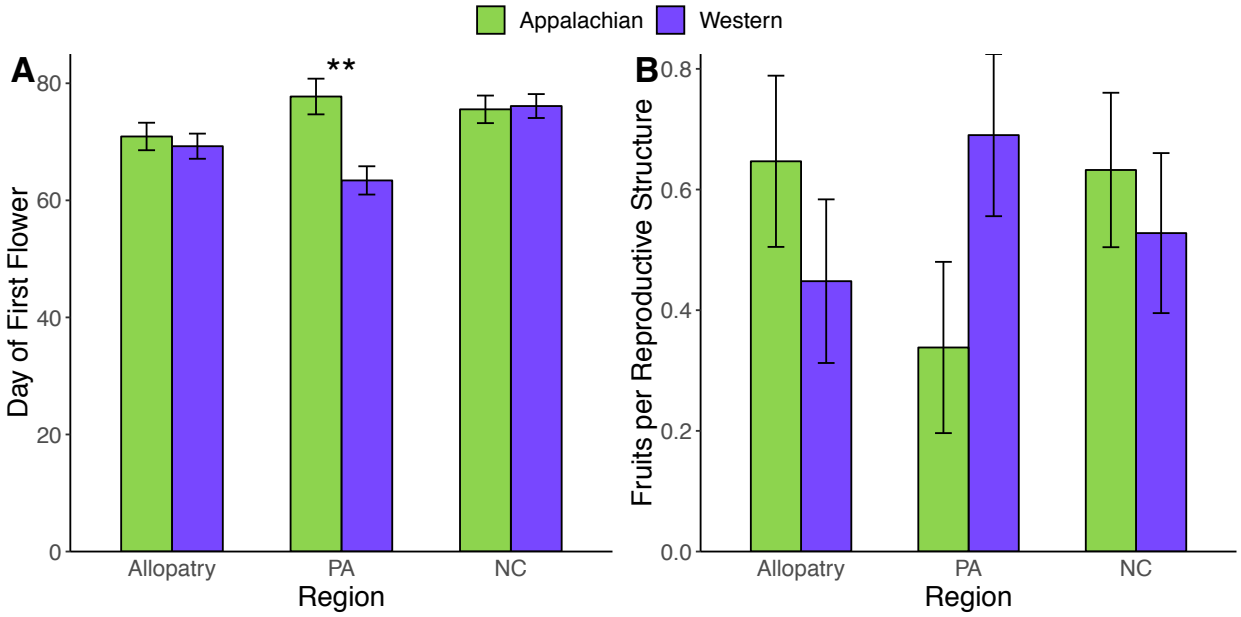


Figure 3. (A) Lsmeans of within-lineage pollinator transitions by bumblebees from populations in the majority (75%) of experimental arrays. Arrays had either Appalachian or Western lineage populations in the majority and included populations from both lineages sampled from allopatry, North Carolina and Pennsylvania. Top dotted line represents array proportion of within-lineage potential partners; bottom dotted line represents nearest-neighbor proportion of within-lineage potential partners. (B) Lsmeans of flower numbers for plants from Appalachian or Western lineage populations from all three regions on the days they were used in the experimental arrays. (C) Lsmeans of within-lineage transitions by bumblebees from populations in the majority (75%) of experimental arrays when the proportion of flowers in the array that were within-lineage was included as a covariate.

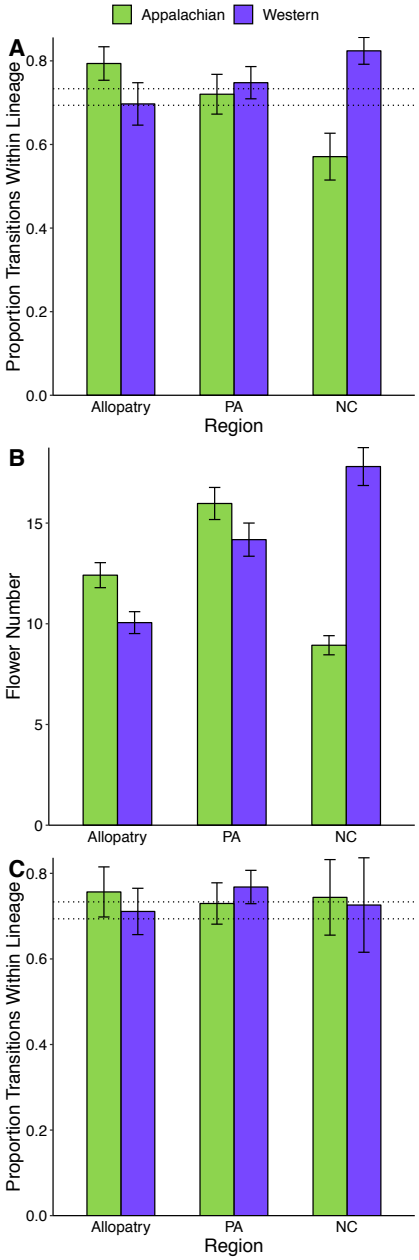


Figure 4. Lsmeans of number of seeds produced by each fruit by crosses within the same lineage (solid bars) and crosses between lineages (hollow bars) from allopatry and two contact zones (NC, PA). Styles were cut at 6 hours post-pollination (A) or left uncut (B). Green bars indicate crosses with Appalachian mothers, and purple bars indicate crosses with Western mothers. Contrasts between neighboring bars are given with + p<0.1, * p<0.05, ** p<0.01, *** p<0.001.

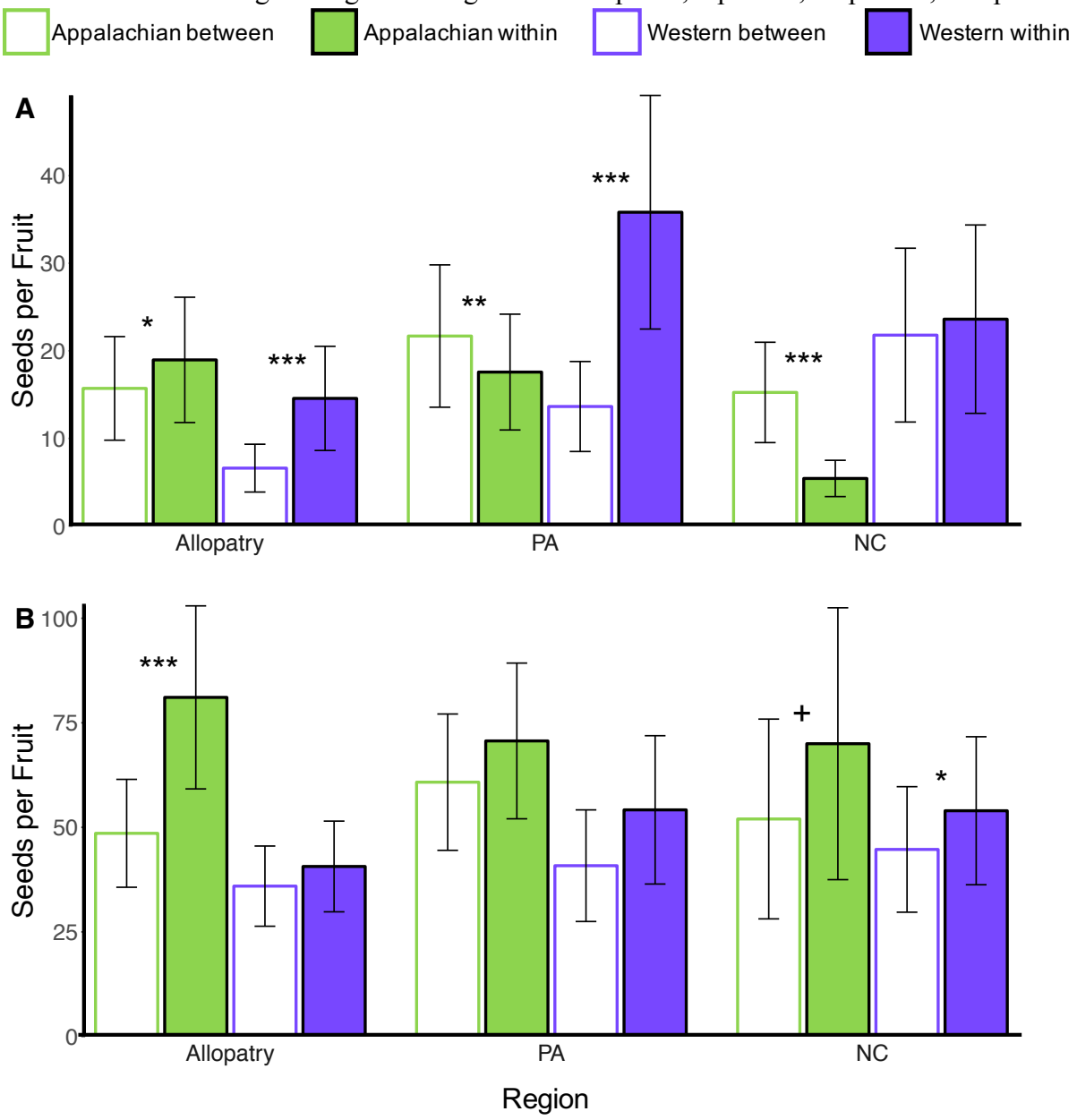


Table S1. Populations of *Campanula americana* were sampled from three regions, including allopatry and contact zones in North Carolina and Pennsylvania. Populations from both the Appalachian and Western lineages were included from each zone.

Population	Lineage	Region	Latitude	Longitude
MD5	Appalachian	Allopatry	39.614	-79.116
PA95	Appalachian	Allopatry	40.475	-78.281
VA73	Appalachian	Allopatry	37.353	-80.552
WV98	Appalachian	Allopatry	39.632	-78.043
AL_BG	Western	Allopatry	34.656	-86.517
KY51	Western	Allopatry	37.934	-84.259
OH119	Western	Allopatry	39.885	-83.997
OH64	Western	Allopatry	41.115	-81.518
NC109E	Appalachian	North Carolina	35.787	-82.973
NC110	Appalachian	North Carolina	35.582	-83.186
NC130	Appalachian	North Carolina	35.516	-83.210
NC91	Appalachian	North Carolina	35.586	-83.066
TN113	Appalachian	North Carolina	35.660	-83.710
TN92	Appalachian	North Carolina	35.676	-83.526
NC105	Western	North Carolina	35.703	-82.833
NC106	Western	North Carolina	35.667	-82.443
NC107	Western	North Carolina	35.943	-82.895
NC108	Western	North Carolina	35.701	-83.106
NC109A	Western	North Carolina	35.748	-82.954
NC114	Western	North Carolina	35.436	-83.048
PA101	Appalachian	Pennsylvania	40.664	-79.501
PA102	Appalachian	Pennsylvania	40.322	-80.111
PA104	Appalachian	Pennsylvania	40.802	-80.055
PA103	Western	Pennsylvania	40.550	-80.311
PA27	Western	Pennsylvania	41.008	-80.083
PA94	Western	Pennsylvania	41.467	-80.011

Table S2. Analysis of variance of the frequency of between-lineage pollinator transitions starting from the lineage that was the most common in its array. Pollinators were grouped into bumblebee and non-bumblebee Pollinator type categories (See Fig. S3A). Arrays were composed of a mix of plants from the Appalachian or the Western lineage with a 75% frequency of one lineage or the other. Arrays included populations sampled from allopatry or contact zones in North Carolina and Pennsylvania. A pair of populations nested within the interaction of region and lineage was included as a random effect (not shown). F-values are given with * $p<0.05$, *** $p<0.001$.

Effect	**Num DF**	**F Value**
Starting lineage[a]	1	1.78
Region[b]	2	0.09
Pollinator Type[a]	1	5.96*
Starting lineage*Region[a]	2	11.18***
Starting lineage*Pollinator Type[a]	1	5.62*
Region*Pollinator Type[a]	2	0.18
Starting lineage*Region*Pollinator Type[a]	2	1.03
Den DF[a]		2023
Den DF[b]		15

Figure S1. Pollination array layouts. Appalachian plants are marked with green dots, and Western plants are marked with purple dots.

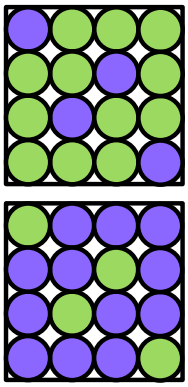


Figure S2. Lsmeans of floral phenotypes of Appalachian (green bars) and Western (purple bars) lineages of *Campanula americana*, across allopatry and two contact zones (NC & PA). (A) Flower diameter; (B) pollen darkness; (C) Style darkness; (D) Style yellow (means); (E) Petal UV chroma; (F) Petal Blue chroma

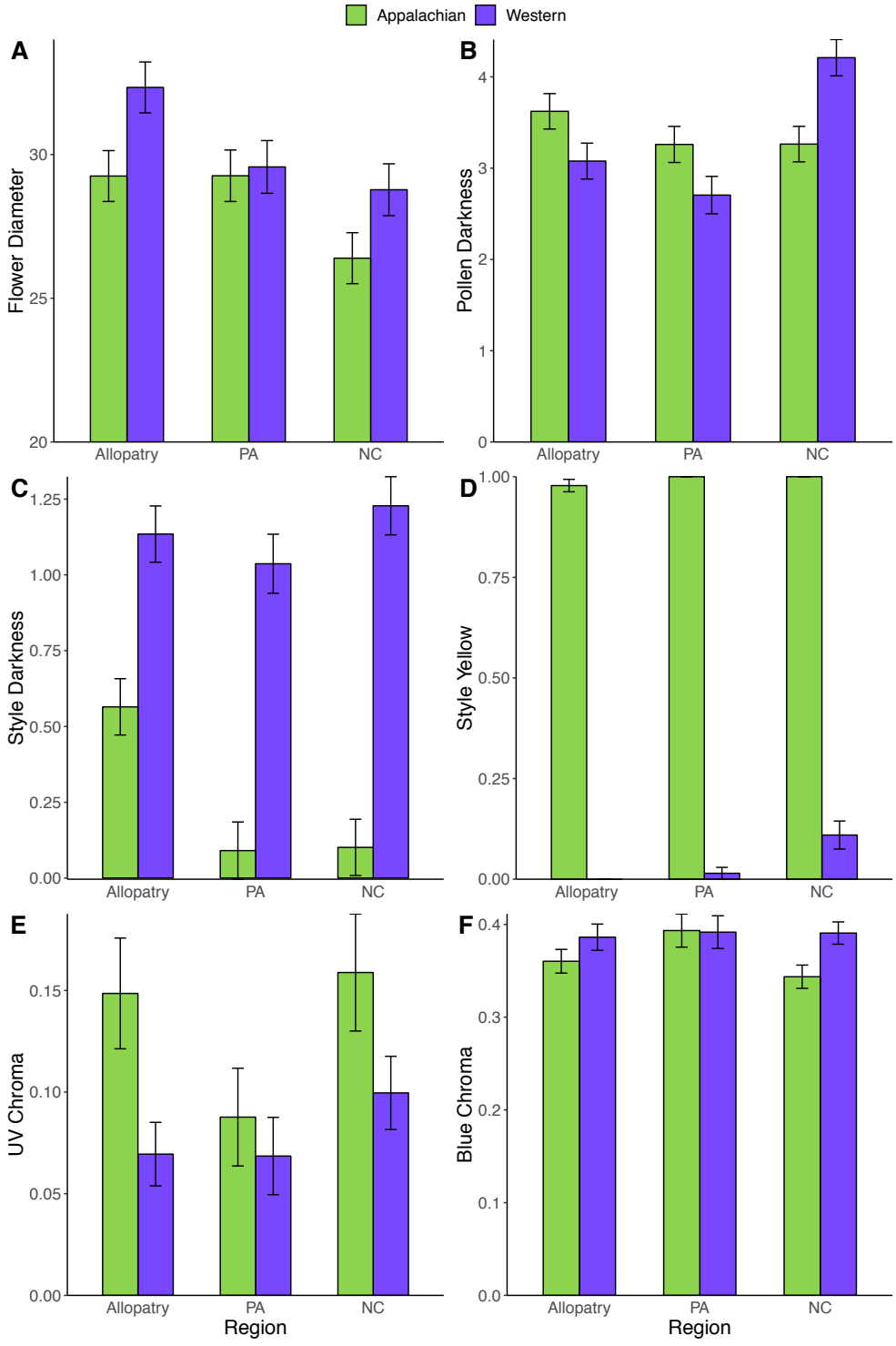
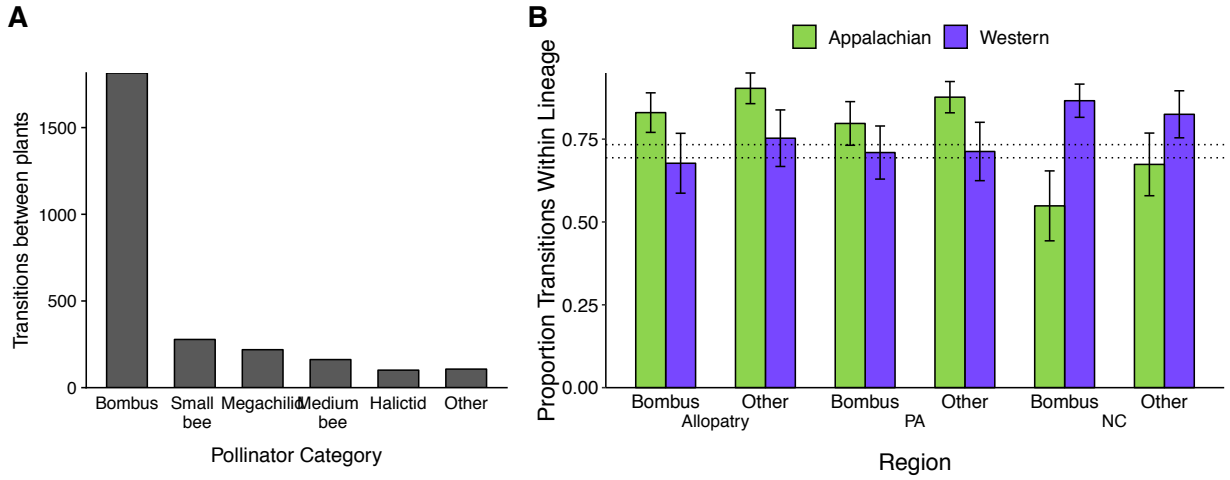


Figure S3. (A) Number of transitions performed by each type of pollinator recorded. (B) Lsmeans of within-lineage pollinator transitions from populations in the majority (75%) of experimental arrays. Arrays had either Appalachian or Western lineage populations in the majority and included populations from both lineages sampled from allopatry, North Carolina and Pennsylvania. Top dotted line represents array proportion of within-lineage potential partners; bottom dotted line represents nearest-neighbor proportion of within-lineage potential partners.



Citations

Barnard-Kubow, K. B., and L. F. Galloway. 2017. Variation in reproductive isolation across a species range. Ecol Evol 7:9347–9357.

Barnard-Kubow, K. B., C. L. Debban, and L. F. Galloway. 2015. Multiple glacial refugia lead to genetic structuring and the potential for reproductive isolation in a herbaceous plant. American Journal of Botany 102:1842–1853.

Barnard-Kubow, K. B., N. So, and L. F. Galloway. 2016. Cytonuclear incompatibility contributes to the early stages of speciation. Evolution 70:2752–2766.

Barton, N., and B. O. Bengtsson. 1986. The barrier to genetic exchange between hybridising populations. Heredity 56:357–376.

Bateson, W. 1909. Heredity and variation in modern lights. Pp. 85–101 *in* A. C. Seward, ed. Darwin and Modern Science. Cambridge, MA.

Brothers, A. N., and J. W. Atwell. 2014. The role of pollinator-mediated selection in the divergence of floral traits between two closely related plant species. Int. J. Plant Sci. 175:287–295.

Coyne, J. A., and H. A. Orr. 2004. Speciation. Sinauer Associates, Inc., Sunderland, MA, USA.

Delph, L. F. 2019. Pollen competition is the mechanism underlying a variety of evolutionary phenomena in dioecious plants. New Phytol 2:1–5.

Dobzhansky, T. 1937. Genetics and the origin of species. Columbia University Press, New York.

Felsenstein, J. 1981. Skepticism towards Santa Rosalia, or Why are there so few kinds of animals? Evolution 35:124–138.

Fishman, L., and R. Wyatt. 1999. Pollinator-mediated competition, reproductive character displacement, and the evolution of selfing in *Arenaria uniflora* (Caryophyllaceae). Evolution 53:1723–1733.

Grossenbacher, D. L., and J. B. Whittall. 2011. Increased floral divergence in sympatric monkeyflowers. Evolution 65:2712–2718.

Haggerty, B. P., and L. F. Galloway. 2010. Response of individual components of reproductive phenology to growing season length in a monocarpic herb. Journal of Ecology 99:242–253.

Hopkins, R. 2013. Reinforcement in plants. New Phytol 197:1095–1103.

Hopkins, R., and M. D. Rausher. 2012. Pollinator-mediated selection on flower color allele drives reinforcement. Science 335:1090–1092.

Ison, J. L., E. S. L. Tuan, M. H. Koski, J. S. Whalen, and L. F. Galloway. 2018. The role of pollinator preference in the maintenance of pollen colour variation. Annals of Botany 99:926–10.

Kay, K. M. 2006. Reproductive isolation between two closely related hummingbird-pollinated neotropical gingers. Evolution 60:538–552.

Kay, K. M., and D. W. Schemske. 2008. Natural selection reinforces speciation in a radiation of neotropical rainforest plants. Evolution 62:2628–2642.

Kay, K. M., and R. D. Sargent. 2009. The role of animal pollination in plant speciation: Integrating ecology, geography, and genetics. Annu. Rev. Ecol. Evol. Syst. 40:637–656.

Kearns, A. M., M. Restani, I. Szabo, A. Schrøder-Nielsen, J. A. Kim, H. M. Richardson, J. M. Marzluff, R. C. Fleischer, A. Johnsen, and K. E. Omland. 2018. Genomic evidence of speciation reversal in ravens. Nature Communications 9:906.

Koski, M. H., D. L. Grossenbacher, J. W. Busch, and L. F. Galloway. 2017. A geographic cline in the ability to self-fertilize is unrelated to the pollination environment. Ecology 98:2930–2939.

Koski, M. H., J. L. Ison, A. Padilla, A. Q. Pham, and L. F. Galloway. 2018. Linking pollinator efficiency to patterns of pollen limitation: small bees exploit the plant–pollinator mutualism. Proceedings of the Royal Society B: Biological Sciences 285:20180635–9.

Lau, J. A., and L. F. Galloway. 2004. Effects of low-efficiency pollinators on plant fitness and floral trait evolution in *Campanula americana* (Campanulaceae). Oecologia 141:577–583.

Matute, D. R. 2010. Reinforcement can overcome gene flow during speciation in *Drosophila*. Current Biology 20:2229–2233.

McNeilly, T., and J. Antonovics. 1968. Evolution in closely adjacent plant populations. Heredity 23:205–218.

Muller, H. J. 1942. Isolating mechanisms, evolution, and temperature. Biological Symposia 6:71-125.

Noor, M. A. F. 1999. Reinforcement and other consequences of sympatry. Heredity 83:503–508.

Silvertown, J., C. Servaes, P. Biss, and D. Macleod. 2005. Reinforcement of reproductive isolation between adjacent populations in the Park Grass Experiment. Heredity 95:198–205.

Stephenson, A. G., and R. I. Bertin. 1983. Male competition, female choice, and sexual selection in plants. Pp. 109–151 *in* Pollination Biology. New York, NY, USA.

Taylor, E. B., J. W. Boughman, M. Groenenboom, M. Sniatynski, D. Schluter, and J. L. Gow. 2005. Speciation in reverse: morphological and genetic evidence of the collapse of a three-spined stickleback (*Gasterosteus aculeatus*) species pair. Molecular Ecology 15:343–355.

Waser, N. M., L. Chittka, M. V. Price, N. M. Williams, and J. Ollerton. 1996. Generalization in pollination systems, and why it matters. Ecology 77:1043–1060.