Title

The Pinyon Pine Syngameon: Modeling the Past to Understand the Present and Predict the Future

Author

Buck, Ryan Collin

Publication Date

2022

UNIVERSITY OF CALIFORNIA
RIVERSIDE

AND

SAN DIEGO STATE UNIVERSITY

The Pinyon Pine Syngameon: Modeling the Past to Understand the Present and Predict the Future

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

Doctor of Philosophy

in

Evolutionary Biology

by

Ryan Collin Buck

June 2022

Dissertation Committee:
- Dr. Lluvia Flores-Rentería, Co-Chairperson
- Dr. Amy Litt, Co-Chairperson
- Dr. Marshal Hedin
- Dr. Kate Ostevik

The Dissertation of Ryan Collin Buck is approved:

Committee Co-Chairperson

Committee Co-Chairperson

University of California, Riverside
San Diego State University

ACKNOWLEDGEMENTS

The text of this dissertation, in part, is a reprint of the material as it appears in Buck *et al.*, 2020 and Buck and Flores-Rentería, 2022. The co-author Dr. Lluvia Flores-Rentería listed in those publications directed and supervised the research which forms the basis for this dissertation. Sandra Hyasat and Alice Hossfeld collected morphological data and helped write the Materials and Methods section of Buck *et al.*, 2020. Funding was provided by the Hispanic-Serving Institutions Education Grants (HSI) Program [grant no. 2018-38422-28614/project accession no. 1016839] from the USDA National Institute of Food and Agriculture, by the National Science Foundation [grant no. DEB-0816675] and by the University Grants Program of San Diego State University.

DEDICATION

This dissertation is dedicated to my supportive parents, Don and Karen, who have encouraged my academic interests, my inspiring sister, Kristin, whose own success motivated my intellectual pursuits, my loving boyfriend, Anthony, who navigated my stress and unusual work hours, my challenging high school AP Biology teacher, Mrs. Yang, who sparked my interest in biology, my college PIs, Steve Weller and Ann Sakai, who nurtured my love for research and introduced me to plant hybridization, and to my mentor Lluvia, whose drive and dedication inspired the hard work and accomplishments seen here.

ABSTRACT OF THE DISSERTATION

The Pinyon Pine Syngameon: Modeling the Past to Understand the Present and Predict the Future

by

Ryan Collin Buck

Doctor of Philosophy, Graduate Program in Evolutionary Biology
University of California, Riverside, and San Diego State University June 2022
Dr. Lluvia Flores-Rentería and Dr. Amy Litt, Co-Chairpersons

Syngameons are complex, multispecies hybridization networks made up of three or more species. The consequences of gene flow within a syngameon are largely unstudied, but participation may confer synergistic benefits that surpass those experienced by hybrid pairs. Climate change and habitat perturbation will likely increase the amount of interspecies interactions and could result in more hybridization events. The pinyon pines that occur throughout the southwestern United States and Baja California, Mexico are thought to hybridize due to their overlapping distributions, long-distance wind-mediated pollen dispersal mechanism, similar pollen dispersal times, observed intermediate morphology, and demonstrated lack of genetic incompatibility barriers. While the taxonomy of some species remains unresolved, drought tolerant traits have been observed within at least three taxa. If hybridization occurs in this complex, these traits could be introgressed across species barriers, which could be essential to the future survival of these species as high pinyon mortality has been observed after prolonged periods of drought. Using morphology, chloroplast haplotypes, next-generation sequencing, and climatic data, we address the presence of species barriers and hybridization, exploring how hybridization influenced the creation of the syngameon and how climate change may impact its future. Five taxa were genetically confirmed to be in this syngameon (*Pinus edulis*, *P. monophylla*, *P. quadrifolia*, *P. californiarum*, and *P.* x *fallax*), with extensive admixture among species. Species identities appear to be maintained at range cores with introgression only occurring where species distributions overlap, making this system a range-edge syngameon. Notably, sequential hybridization was detected where two species (*P. edulis* and either *P. monophylla* or *P. quadrifolia*) hybridized to create *P. californiarum*, which is currently hybridizing with *P. edulis* to produce the *P.* x *fallax* lineage. Populations composed exclusively of or mostly of admixed individuals were found in areas of tri-species sympatry, suggesting that genetic swamping could be occurring in these areas. Future niche models predict the loss of suitable habitat in all climate scenarios modeled, with minor expansions northwards and up in elevation. Populations of *P. californiarum* showed low structure and genetic diversity, with high levels of inbreeding, indicating that this rarely recognized species should be of conservation concern.

TABLE OF CONTENTS

LIST OF TABLES

Chapter I

Chapter II

Chapter III

Chapter IV

Chapter V

LIST OF FIGURES

Chapter I

Chapter II

Chapter III

Chapter IV

Chapter V

INTRODUCTION

The term syngameon is not widely known within the scientific community. As my research heavily focuses on multispecies hybridization, I first start with a background chapter (chapter one) reviewing the terminology, history, and formation hypotheses surrounding syngameons in both plants and animals. With that foundation set, the dissertation zooms in on a complex of pinyon pine species inhabiting the southwestern United States and Baja California, Mexico. Due to the taxonomic controversy surrounding the taxa in this system, the validity of each species must be investigated first. Previous phylogenetic studies did not provide definitive results as they left out taxa (mainly *P. juarezensis*, *californiarum*-type, and *fallax*-type), used paralogous loci, or encountered potential hybridization. Conclusive genetic studies were needed to determine what species of pinyon pines exist in the Southwest and to investigate the presence and role of hybridization in this complex. In chapter two, we begin by focusing on the proposed taxa in the western edge of the complex, addressing the hybrid origin of *P. quadrifolia* and the existence of *P. juarezensis*. It was discovered that three hybridizing species of pinyons exist across California and Baja California: *P. monophylla*, *P. californiarum*, and *P. quadrifolia*. In chapter three, we then include the two taxa from the eastern portion of the complex to assess if they are participating in the syngameon and if their drought tolerant traits could be introgressed across species barriers. The inclusion of all proposed taxa allowed us to assess the structure of the syngameon and the dynamics of its participants. With the participants and structure of the syngameon determined, bigger questions about the past and future of the syngameon could be explored. While examining the hybrid origin of *fallax*-type, we discover the hybrid origin of a second taxa, suggesting that syngameon participation facilitated the creation of two lineages and that it may have accelerated the speciation process. Hybridization has played a significant role in the evolution of this syngameon, but the future of these species remains uncertain, especially with climate change-induced droughts devastating forests worldwide. In chapter four, we attempt to understand how syngameon dynamics will change over time by modeling future habitat suitability. Substantial predicted habitat loss of the already restricted and newly delineated species, *P. californiarum*, warranted further investigations into its genetic conservation status. Lastly in chapter five, we explore the genetic health of *P. californiarum* populations and provide usable management recommendations to restore its genetic diversity.

CHAPTER I

The Syngameon Enigma

Ryan Buck and Lluvia Flores-Rentería

Department of Biology, San Diego State University, San Diego, CA, USA

Abstract: Despite their evolutionary relevance, multispecies networks or syngameons are rarely reported in the literature. Discovering how syngameons form and how they are maintained can give insight into processes such as adaptive radiations, island colonizations, and the creation of new hybrid lineages. Understanding these complex hybridization networks is even more pressing with anthropogenic climate change, as syngameons may have unique synergistic properties that will allow participating species to persist. The formation of a syngameon is not insurmountable, as several ways for a syngameon to form have been proposed, depending mostly on the magnitude and frequency of gene flow events, as well as the relatedness of its participants. Episodic hybridization with small amounts of introgression may keep syngameons stable and protect their participants from any detrimental effects of gene flow. As genomic sequencing becomes cheaper and more species are included in studies, the number of known syngameons is expected to increase. Syngameons must be considered in conservation efforts as the extinction of one participating species may have detrimental effects on the survival of all other species in the network.

Keywords: syngameon; hybridization; multispecies

Nearly all organisms, met with in nature as well as under cultivation, man included, are hybrids which were mistakenly considered to be specifically pure, so that their behaviour was unconsciously held to be that of specifically pure organisms, while it was that of hybrids; so it happened that segregation was mistaken for heredity.

- Lotsy 1916

Introduction

Interspecies hybridization is relatively common across taxa, with occurrence estimates of 25% in plants and 10% in animals (Mallet 2005), and is thought to be the cause of several major speciation events (Soltis and Soltis, 2009; Abbott *et al.*, 2013). When a group of otherwise distinct species are connected by hybridization, they form a syngameon, a copulative community (Lotsy 1925). The discovery of the first natural syngameon in birch trees by Gunnarsson and the multidirectional hybridization of cultivated *Saxifraga* by Lloyd (in Lotsy 1925) sparked the first serious investigations into hybridization itself, which continue to this day. From its inception, the word syngameon was used to collectively describe "a large number of different individuals [from different species], which are all apparently able to produce fertile offspring with one another; one very large pairing-community, one syngameon" (Lotsy 1925). However, shortly thereafter, Du Rietz (1930) used the term to describe a polymorphic hybrid collective in which "species have got more or less lost", precipitating an idea that syngameons were just a collection of taxonomic misfits unable to be classified. Later, other authors such as Cuenot (1951), Grant (1957), and Beaudry (1960) reclaimed the word to mean "species linked by frequent or occasional hybridization, a hybridizing group of species". Out of this definition was born the ecological species concept, which allowed for gene flow but separated species by their adaptations to particular niches, in an attempt to explain oak differentiation under gene flow. The now well-known oak syngameon was alternatively named a "multispecies" by Van Valen (1976). Templeton (1989) began a pattern of syngameon misinterpretation when he conflated the term with a hybridizing pair, the aftermath of which can be seen in recent hybridization papers that use syngameon to describe any two species that hybridize (e.g. Arduino *et al*., 1996; Holliday 2006; Barker 2007; Caujapé-Castells *et al*., 2017). According to Boecklen (2017) a syngameon is produced when a group of closely related species forms a complex set of hybrid combinations. We recommend this use of syngameon to define a breeding system with a three or more multispecies network. The minimum number of three participating species is important in this context as it distinguishes the commonly studied hybridizing species pairs from a more complex and possibly synergistic interacting multispecies system (Cannon and Petit, 2020). It is important to note that the current definition does not distinguish between fertile and sterile hybrids, or diploid and polyploid systems, both of which could have varying impacts on the structure of the syngameon. However, the term seems to be restricted to naturally occurring hybrids, whereas the term coenospecies would refer to artificial hybrids (Glossary) (Turesson 1922).

Despite misused language, the amount of described syngameons remains extremely low in comparison to the number of hybridizing species pairs and thus far is generally restricted to plant taxa (Table 1). Only recently have researchers begun to hypothesize how these rare complexes are able to overcome numerous reproductive barriers in the process of their formation (Cannon and Lerdau, 2015; Bog 2016; Cannon and Scher, 2017; Liu *et al*., 2017; Meier *et al*., 2017; Cronk and Suarez-Gonzalez, 2018; Suarez-Gonzalez et at., 2018; Cannon and Lerdau, 2019; Levi *et al*., 2019). Even less is known about how these interactions are maintained over time or if they are perpetually unstable. Additionally, the evolutionary consequences of sustained multispecies gene flow remain unexplored, leaving the future of syngameons speculative at best. In this review, we will explore three questions: 1) how do syngameons form, 2) how are they maintained over evolutionary time, and 3) why are they so rare? Lastly, we will discuss the future of syngameons in light of a changing world and provide some recommendations relevant to conservation.

Table 1. Known Syngameons. List of genera with known syngameons and their common names in parentheses. Their number of participants and the taxonomic kingdom they belong to are listed in the following columns.

Genera (common name)	Known Participants	Kingdom	Source
Acropora (Coral)	8	Animalia	van Oppen *et al*., 2001*; Ladner 2012*
Anser + *Branta* (Geese)	15	Animalia	Ottenburghs *et al.*, 2016
Artibeus (bats)	3	Animalia	Larsen *et al*., 2010
Callithrix (marmosets)	3	Animalia	Ackermann *et al.*, 2019
Canis	3	Animalia	Wilson *et al*., 2009; Rutledge *et al.*, 2010*
Carabus (Carabid beetles)	6	Animalia	Kubota and Sota, 1998
Catostomus + *Chasmistes* + *Deltistes* (catostomid fish)	4	Animalia	Dowling *et al.*, 2016*
Cerion (snail)	not specified	Animalia	Gould and Woodruff, 1990*
Colias (sulfur butterflies)	3	Animalia	Wheat and Watt, 2008*
Daphnia (plankton)	5	Animalia	Colbourne *et al*., 2006 (but see Hebert and Wilson, 2006)
Desmognathus (Dusky Salamanders)	3	Animalia	Pyron *et al*., 2020
Drosophila	at least 3	Animalia	Nunes *et al*., 2010; Matute and Ayroles, 2014
Eueides (butterflies)	5	Animalia	Mallet *et al*., 2007
Geospiza (Darwin's finches)	two sets of 3	Animalia	Grant and Grant, 1989, 1992, 2010
Habronattus (jumping spiders)	at least 3	Animalia	Leduc-Robert and Maddison, 2018
Heliconius (butterflies)	one set of 3; one set of 4; one set of 9	Animalia	Mallet *et al*., 2007
Homo	3	Animalia	Hammer *et al*., 2011; Ackermann *et al*., 2019
Konia + *Myaka* + *Pungu* + *Sarotherodon* (Cichlid fish)	8	Animalia	Schliewen and Klee, 2004*
Liolaemus (lizard)	4	Animalia	Olave *et al*., 2018

Montastraea (coral)	3	Animalia	Budd 2010* (but see Fukami *et al*., 2004)
Pacifigorgia (octocorals)	not specified	Animalia	Granados Cifuentes 2008; Manrique Rodríguez 2008*
Papio + *Theropithecus* (baboons)	at least 3	Animalia	Godfrey and Marks, 1991*
Psammocora (Indo-Pacific corals)	3	Animalia	Stefani *et al*., 2008
Pseudophryne (frogs)	3	Animalia	Woodruff 1981
Steatocranus (cichlid fish)	18	Animalia	Schwarzer *et al*., 2012*
Stylophora (Red Sea coral)	not specified	Animalia	Arrigoni *et al*., 2016*
Sus (wild pigs)	4	Animalia	Frantz *et al*., 2013
Ursa (bears)	6	Animalia	Kumar *et al*., 2017
Xiphophorus (fishes)	5	Animalia	Cui *et al*., 2013
Abies (fir)	3	Plantae	Cignet *et al*., 2015 (in Beddows and Rose, 2018*)
Actinidia (Kiwi)	9	Plantae	Liu *et al*., 2010*, 2017
A8esculus (buckeye)	3	Plantae	dePamphilis and Wyatt, 1989 (in Beddows and Rose, 2018*)
Ajuga (bugleherb) or *Amaranthus* (amaranths)	5	Plantae	Beddows and Rose, 2018*
Ambrosia (ragweed)	3	Plantae	Beddows and Rose, 2018*
Amelanchier (serviceberry)	5	Plantae	Nielsen 1939 (in Beddows and Rose, 2018*)
Aquilegia (Columbines)	not specified	Plantae	Grant 1971*
Arbutus (madrones)	5	Plantae	McElwee-Adame (pers. comm.)
Arctostaphylos (manzanita)	at least 3	Plantae	Gottlieb 1968; Schmid *et al*., 1968
Asclepias (milkweed)	4	Plantae	Beddows and Rose, 2018*
Asplenium (spleenworts)	16	Plantae	Brownsey 1977 (in Boecklen 2017*)
Betula (birch)	one set of 4; one set of 6	Plantae	Gunnarsson (in Lotsy 1925*); Barnes and Dancik, 1985 (in Beddows and Rose, 2018*); Walters and Yawney, 2004 (in Beddows and Rose, 2018*)

Boechera (rockcress)	58	Plantae	Alexander *et al*., 2015; D. Bailey (in Boecklen 2017*)
Carex (true sedges)	three sets of 3; two sets of 4	Plantae	Hedrén 2004*
Castanea (chestnut)	not specified	Plantae	Chen *et al*., 2014*
Ceanothus (California lilac)	not specified	Plantae	Grant 1981*
Cirsium (plume thistle)	17	Plantae	Bureš *et al*., 2010*
Citrus	8	Plantae	Wu *et al*., 2018; Butelli *et al*., 2019*
Coprosma (stinkwood)	6	Plantae	Papadopulos *et al*., 2013*
Cornus (dogwood)	4	Plantae	Beddows and Rose, 2018*
Corybas (helmet orchid)	3	Plantae	Wagner *et al*., 2021
Cyperus	3	Plantae	Marcks 1974 (in Beddows and Rose, 2018*)
Dichanthelium (rosette grass)	two sets of 3; one set of 4	Plantae	Voss and Reznicek, 2012 (in Beddows and Rose, 2018); Beddows and Rose, 2018*;
Diospyros (ebonies)	one set of 3, one set of 4	Plantae	Linan *et al*., 2021*
Diplacus (monkey flower)	5	Plantae	Beeks 1962; Grant 1981*
Drosera (sundew)	4	Plantae	Wood 1955 (in Beddows and Rose, 2018*); Crowder *et al*., 1990 (in Beddows and Rose, 2018*)
Dryopteris (woodfern)	4	Plantae	Rünk *et al*., 2012 (in Beddows and Rose, 2018*)
Dubatia	6	Plantae	Carr 1978
Elymus (wildrye)	3	Plantae	Beddows and Rose, 2018*
Encelia (brittlebush)	11	Plantae	Singhal *et al*., 2021*
Equisetum (horsetail)	3	Plantae	Beddows and Rose, 2018*
Eschweilera	3	Plantae	Schmitt *et al*., 2020*
Espeletia (frailejones)	3	Plantae	Pineda Torres 2019
Eucalyptus (Green ashes)	4	Plantae	Rutherford *et al*., 2018
Eucalyptus (Boxes)	~10	Plantae	Flores-Rentería *et al*., 2017
Ficus (figs)	13	Plantae	Wang *et al*., 2021
Gentiana	4	Plantae	Beddows and Rose, 2018*
Geum (avens)	not specified	Plantae	Du Rietz 1930 (in Grant 1981*)

Gymnocarpium (oak fern)	4	Plantae	Pryer and Haufler, 1993, in Beddows and Rose, 2018*
Helianthus (sunflower)	one set of 4; one set of 6	Plantae	Heiser 1949, 1951a, 1951b; Rieseberg 1991b; Beddows and Rose, 2018*
Hieracium (hawkweed)	one set of 4; one set of 5	Plantae	Beddows and Rose, 2018*
Huperzia (firmosses)	3	Plantae	Beddows and Rose, 2018*
Hypericum (St. John's wort)	3	Plantae	Beddows and Rose, 2018*
Iris (California irises)	12	Plantae	Lenz 1959*; Grant 1981; Young 1998 (in Boecklen 2017*)
Juncus (rushes)	7	Plantae	Lint 1977*
Juniperus (junipers)	3	Plantae	Flake 1978; Palma-Otal *et al*., 1983 (in Beddows and Rose, 2018*)
Lantana	6	Plantae	Urban *et al*., 2011*
Lespedeza (bush clovers)	5	Plantae	Beddows and Rose, 2018*
Ligularia (leopard plants)	3	Plantae	Zhang *et al*., 2018
Lycopodiella (bog clubmosses)	4	Plantae	Barkworth and Adams, 1993 (in Beddows and Rose, 2018*)
Lycopus	4	Plantae	Beddows and Rose, 2018*
Lysimachia	3	Plantae	Beddows and Rose, 2018*
Melandrium/*Silene* (campion)	not specified	Plantae	Du Rietz 1930 (in Grant 1981*)
Micromeria	20	Plantae	Curto *et al*., 2017*
Nothofagus (southern beeches)	at least 3	Plantae	Du Rietz 1930 (in Grant 1981*); Stecconi *et al*., 2004; Soliani *et al*., 2012
Opuntia (prickly pear cactus)	at least 16	Plantae	Grant and Grant, 1979*
Phaseolus (bean)	3	Plantae	Debouck 1992*
Phlox	3	Plantae	Wyatt 1981 (in Beddows and Rose, 2018*)
Picea (spruces)	3	Plantae	Hamilton *et al*., 2015
Pinus (Southwestern pinyon pines)	4	Plantae	Buck *et al*., 2020*, 2022 in review*
Platanthera (butterfly orchids)	two sets of 3	Plantae	Wallace 2003 (in Beddows and Rose, 2018*)

Populus (cottonwood)	three sets of 3	Plantae	Seybold 2009 (in Beddows and Rose, 2018*); Chhatre *et al*., 2018 (but see Cronk and Suarez-Gonzalez, 2018*)
Potamogeton (pondweed)	19	Plantae	Grant 1981 (in Boecklen 2017*); Clapham *et al*., 1990
Prosopis (mesquite)	7	Plantae	Saidman and Valardi, 1987*; Torales *et al*., 2013*
Prunus (plums)	18	Plantae	Shaw 2005*
Pycnanthemum (mountain mints)	3	Plantae	Beddows and Rose, 2018*
Quercus (Chinese oaks)	4	Plantae	Li *et al*., 2021
Quercus (Eastern white oaks)	14	Plantae	Grant 1981*; Hardin 1975*
Quercus (Southwestern white oaks)	16	Plantae	Grant 1981*; R. Spellenberg (in Boecklen 2017*)
Rosa (rose)	3	Plantae	Lewis 2008; Beddows and Rose, 2018*
Rubus (brambles)	3	Plantae	Beddows and Rose, 2018*
Salix (willow)	two sets of 3; one set of 6	Plantae	Brunsfeld *et al*., 1992; Voss and Reznicek, 2012 (in Beddows and Rose, 2018*)
Saxifraga (saxifrages)	3	Plantae	Lloyd (in Lotsy 1925*)
Schiedea	4	Plantae	Weller *et al*., 2001
Scirpus (club-rush)	3	Plantae	Beddows and Rose, 2018*
Senecio	5	Plantae	Bog 2016*
Solidago (goldenrods)	one set of 4; one set of 5	Plantae	Beddows and Rose, 2018* (but see Laureto and Barkman, 2011)
Sphaeralcea (globemallows)	not specified	Plantae	Dreher 2014
Stipa	two sets of 3	Plantae	Love 1954; Baiakhemtov *et al*., 2020
Symphonia	3	Plantae	Schmitt *et al*., 2021*
Symphyotrichum	8	Plantae	Brouillet *et al*., 2006 (in Beddows and Rose, 2018*)
Thalictrum (meadow-rue)	3	Plantae	Beddows and Rose, 2018*
Tolumnia (Dancing-lady orchid)	4	Plantae	Sauleda and Hamilton, 2006*
Tragopogon (salsifies)	5	Plantae	Ownbey 1950; Lipman *et al*., 2013

Trillium	one set of 3; one set of 4	Plantae	Stoehrel 2010* (but see Case and Case1997); Beddows and Rose, 2018*;
Tripsacum (gamagrass)	7	Plantae	Randolph 1970*
Verbascum (mullein)	4	Plantae	Beddows and Rose, 2018*
Verbena (vervain)	4	Plantae	Beddows and Rose, 2018*
Viola	one set of 4; one set of 5; one set of 7	Plantae	Beddows and Rose, 2018*

*Cited study used the term "syngameon".

How do Syngameons Form and Collapse?

The Origin of Syngameons

For hybridization to occur, species must overcome any existing barriers to gene flow, which include pre- and postzygotic reproductive isolating mechanisms. Prezygotic barriers can consist of temporal, geographic, mechanical, behavioral, and genetic mechanisms, while postzygotic barriers can consist of hybrid sterility, hybrid inviability, and F_2 breakdown (Widmer *et al*., 2009). While not always initially present, some of these barriers can form after the secondary contact of two lineages to prevent further hybridization and reinforce species boundaries (Garner *et al*., 2018). Conversely, the initial lack of reproductive isolating mechanisms or the failure of reinforcement (Glossary) can lead to stable hybrid zones. Despite the numerous obstacles faced, hybridization is not rare (Mallet 2005; Moran *et al*., 2021). So why, then, are syngameons so rarely reported? After all, syngameons are just hybridization events between three or more species. What makes adding this third species interaction so difficult? The answer may lie in how syngameons form and collapse.

The Birth and Death Hypotheses

The Rapid Radiation Hypothesis

The rapid radiation hypothesis (Seehausen 2004; Liu *et al.*, 2017) postulates that rapid radiations, or in other words, relatively quick and numerous speciation events, allowed for the repeated origins of hybrid lineages. Syngameons are able to form among the newly radiated species because reproductive isolating mechanisms have yet to develop (Figure 1). In turn, this gene flow can act as a catalyst for additional radiation by replenishing standing genetic variation, aiding in the consumption of unexploited resources and occupation of new niches (Kagawa and Seehausen, 2020). Further, these hybrid lineages could speciate themselves, becoming hybrid species; however, this often requires the formation of reproductive barriers, which would exclude the newly formed species from the syngameon. Seehausen (2004) used Heliconius butterflies to exemplify syngameons providing new adaptive traits and promoting ecological diversification. Using kiwifruit as an example, Liu (2017) showed how syngameons developed during early radiation, but later collapsed as species diversified into new ecological opportunities to reduce contact and competition (but see Yang *et al.*, 2019). The classic examples of radiations, including Heliconius butterflies (Beltrán *et al.*, 2002; Gilbert 2003), Darwin’s finches (Grant and Grant, 1996; Freeland and Boag, 1999), and African cichlid fish (Seehausen *et al.*, 1997), showed a similar pattern as species numbers rose and underutilized resources became scarce, stabilizing selection occurred and species began to accumulate genomic incompatibilities (Seehausen 2004). While the concepts behind the hypothesis remain valid, it is difficult to prove if ancient syngameons formed during radiation events, as many participating species may have since gone extinct and the rapid timeline of diversification would make a transient hybridization event hard to detect. Current simplified methods to detect ancient hybridization (e.g. ABBA-BABA) can fail to distinguish population structure from actual introgression when population sizes are large (Martin *et al*., 2015), as could happen in syngameons and rapid radiations. However, with improving molecular and coalescent techniques, ghost lineages and ancient introgression events are becoming easier to trace (Twyford and Ennos, 2012; Hey *et al*., 2018; Ottenburghs 2020).

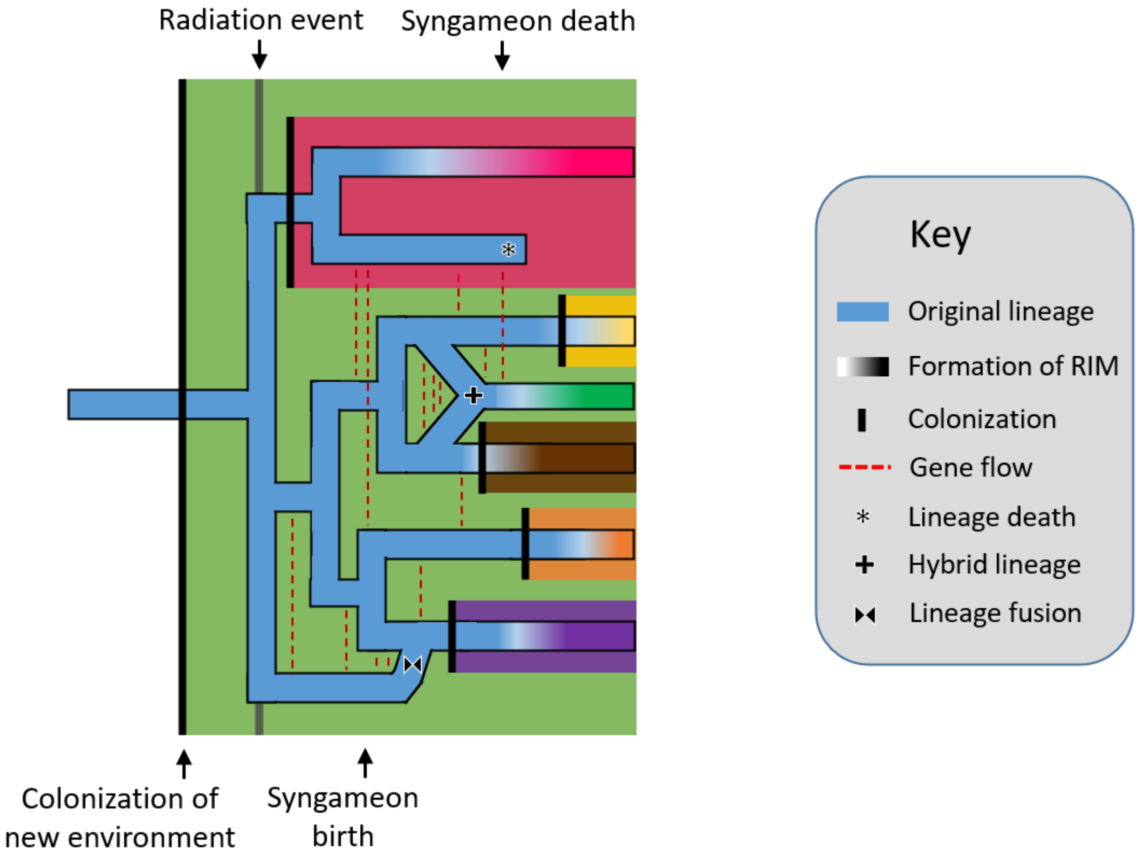


Figure 1. Rapid radiation hypothesis showing a lineage (blue horizontal line) colonizing a new environment (black vertical line), which eventually triggers a rapid radiation event. Speciation is followed by gene flow events (red dashed lines) which form a syngameon. The eventual collapse of the syngameon occurs when reproductive isolating barriers form among species, usually after the colonization of new environments, leaving two or no species with interspecific gene flow. Several potential outcomes are shown including hybrid speciation (plus symbol), extinction (asterisk), and fusion (bowtie symbol). RIM= Reproductive Isolation Mechanism.

Surfing Syngameon Hypothesis

Rather than rapid radiation events causing syngameons, the surfing syngameon hypothesis (Caujapé-Castells and Bramwell, 2011; Caujapé-Castells *et al*., 2017) suggests syngameons that form during island colonization events can both cause and prevent rapid radiations. Distinct colonizing genotypes (referred to by Caujapé-Castells as morpho-species or incipient species) that were previously isolated on the mainland but are phylogenetically close enough to have gene flow can form a syngameon during a colonizing event of a new island (Figure 2A & 2B). The increase in genetic diversity would be enough to overcome selective pressures and founder effects, promoting colonization of syngameon participants. Using species in the Canarian archipelago, Caujapé-Castells and Bramwell (2011) indicated that this type of event could be detected through the level of endemic species, with low levels of endemism resulting from syngameon colonization and high levels from the formation of incompatibility barriers. In low-complexity islands, syngameons would stall evolutionary change due to high levels of gene flow homogenizing genomes, thus preventing rapid diversification. In high-complexity islands, syngameons would promote adaptations due to the high genetic diversity hybridization provides; therefore resulting in rapid radiations and the eventual collapse of the syngameon (Figure 2C) due to competition and the formation of reproductive barriers (Caujapé-Castells *et al*., 2017). Future studies should test the validity of this hypothesis beyond the Canary Islands and examine if it is broadly applicable to other allopatric dispersal scenarios such as the colonization of nearby niches or mosaic hybrid zones.

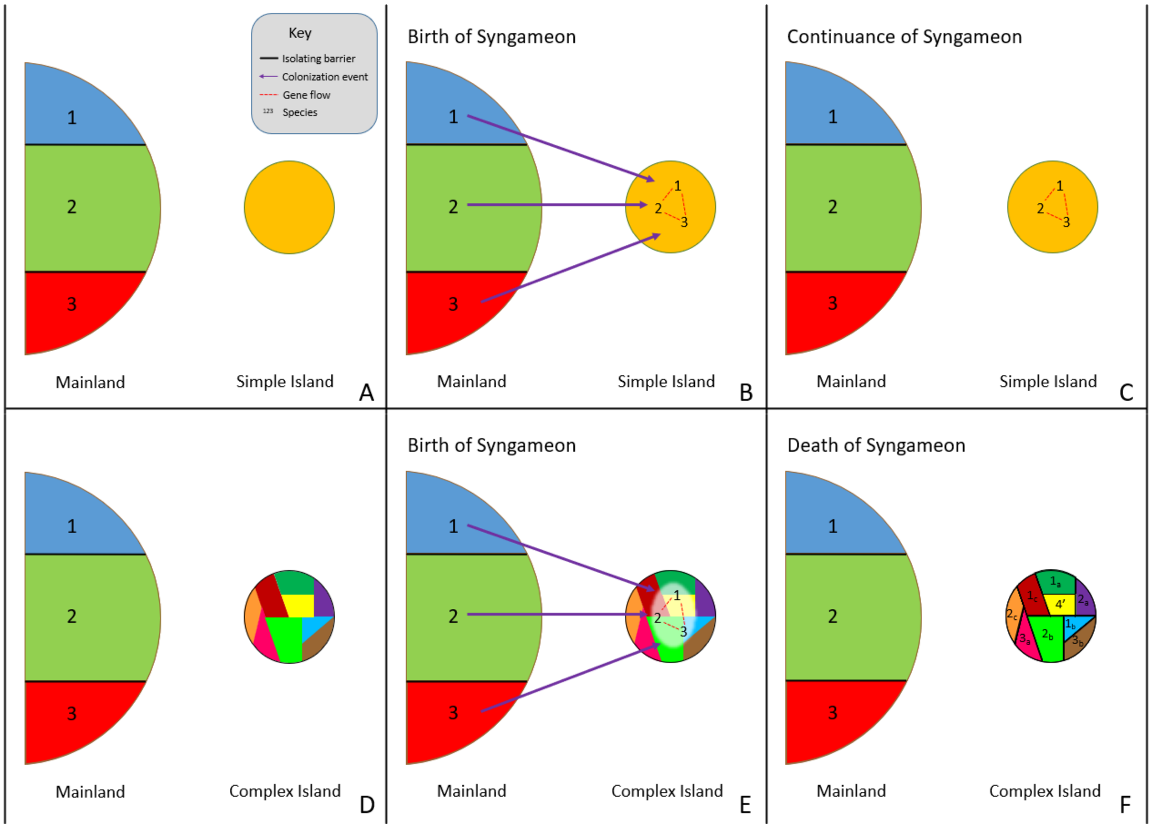


Figure 2. Surfing syngameon hypothesis, in which previously isolated species (1, 2, and 3) come into contact during the colonization of a low-complexity island (B) and high-complexity island (E), resulting in hybridization and the formation of a syngameon (B and E). The syngameon increases genetic diversity and reduces the effects of bottleneck events, resulting in the successful colonization of an island. If the island is open and uniform (A-C), with little to no ecological and geographical complexity (simple island), then evolutionary change is slowed down by syngameonic introgression/gene flow,

resulting in homogenization of traits and the continuation of the syngameon (C). If the island is geographically and ecologically complex (D-F), then selection, adaptation, and competition eventually drive divergence and the formation of reproductive isolating barriers, resulting in the eventual collapse of the syngameon (F). Participation could even result in the creation of a new hybrid lineage (F, shown as lineage 4').

Edge Range Hypothesis

Syngameons may form at the edges of species ranges, where multiple species can overlap in distribution (Figure 3A) (Cronk and Suarez-Gonzalez, 2018). Typically, range edges are seen as population sinks because the species is unable to adapt to the new, local environment beyond the current distribution boundary (Pfennig *et al.*, 2016). However, hybridization at a species' range edge may facilitate survival by introducing locally adapted or novel traits through introgression (Pfennig *et al.*, 2016). Cronk and Suarez-Gonzalez (2018) used a poplar syngameon to show how a tri-species interaction allows for the increased survival of hybrids at the edge of species boundaries. They also illustrated how as ranges expand and contract, these gene flow events could be episodic, explaining patterns of ancient introgression followed by divergence, then introgression again. However, Ottenburghs (2021) pointed out that these "merging-and-diverging cycles" could result in the build up of genetic divergence during allopatric phases, leading to lower levels of introgression during the following sympatric phase, eventually ending with a collapse of the syngameon. Additionally, Cronk and Suarez-Gonzalez (2018) failed to consider the stability of syngameons at range edges because these interactions could lead to the formation of species barriers and thus the collapse of the syngameon, or even the formation of a new hybrid species with higher fitness than its parental species. Hybrid speciation could lead to the collapse of the syngameon and possible extinction of the parent species via genetic swamping (Glossary) or hybrid superiority (Todesco *et al*., 2016). Moreover, there are examples of syngameons that do not form at the range edges, such as in Quercus in which some species overlap in wide ranges of the distribution (Cronk and Suarez-Gonzalez, 2018). As studies expand their scope beyond hybrid pairs to include more hybridizing species, range overlaps should be closely investigated in order to revisit this hypothesis under more scrutiny.

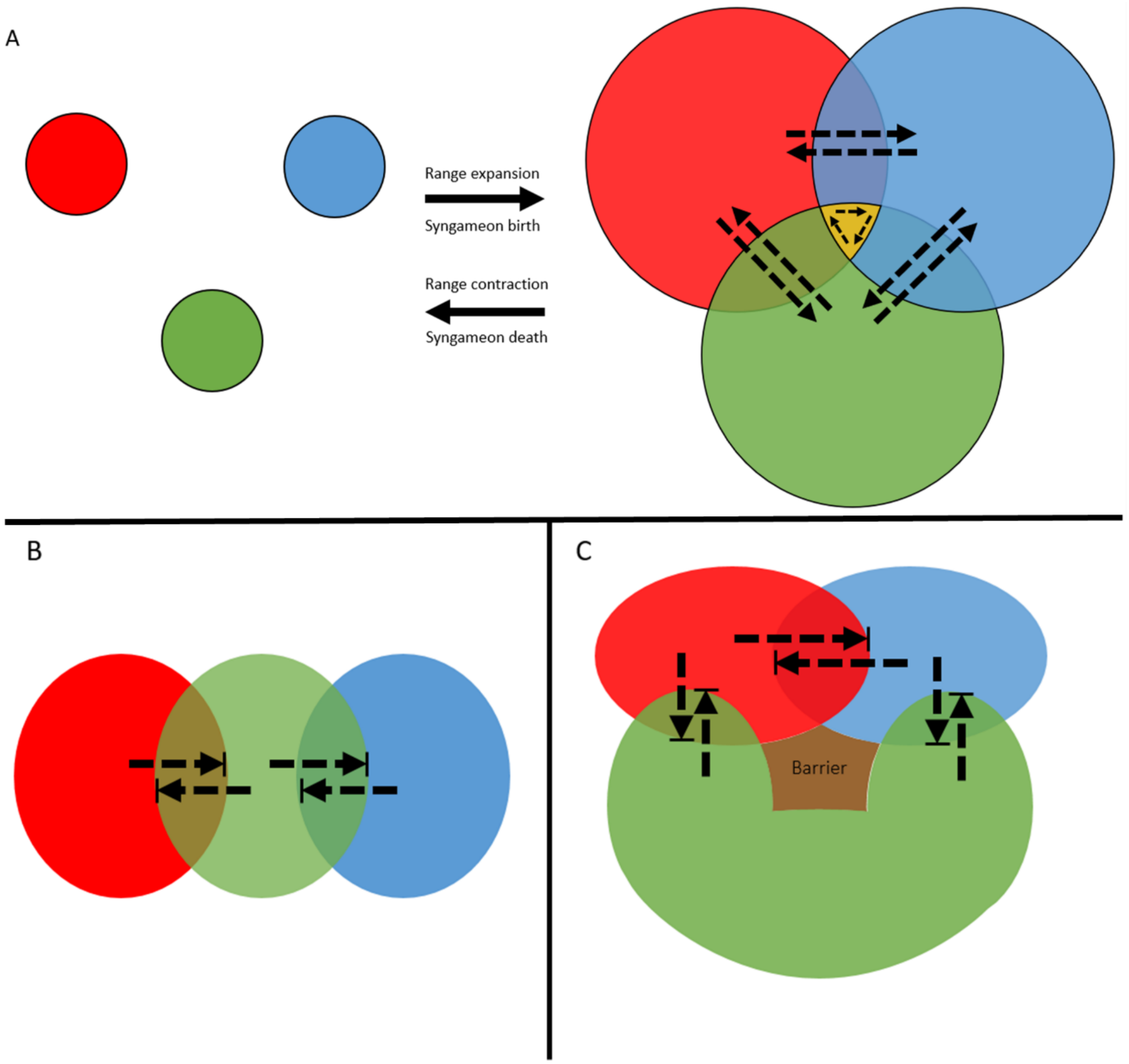


Figure 3. Edge-range hypothesis whereby the expansion and contraction of species' ranges (A) over time makes gene flow within the syngameon episodic. This allows for the retention of species' identities while still allowing for the exchange of adaptive alleles (dashed arrows). A caveat to the edge-range hypothesis is that all three species' ranges rarely overlap (shown in gold). More probable scenarios are shown in B and C, where species' distributions overlap independently. While still technically syngameons, the scenarios represented in B and C may result in introgression not extending past the hybrid zones (bounded box), resulting in local admixture directly between hybrid pairs but no genes are shared indirectly through introgression via a third species.

Genomic Mutualist Hypothesis

Lastly, Cannon and Lerdau (2015) hypothesized that species form syngameons by acting as genomic mutualists. In their scenario, multiple species remain partially interfertile with each other but experience divergent selection on portions of their genome, while low levels of neutral or adaptive gene flow occur in other parts of the genome. This creates a balance between purifying selection within species for specific phenotypes and diversifying selection among species for novel phenotypes. To avoid the negative consequences of extensive gene flow, species would develop a reduced but persistent capacity for interspecific mating, making periods of gene flow infrequent, episodic, and often unidirectional; however, in some systems syngameons are multidirectional and often reciprocal gene flow occurs in different magnitudes (Chhatre *et al*., 2018; Cronk and Suarez-Gonzalez, 2018; Buck *et al*., 2020). These mating interactions are largely controlled by the quantity and quality of pollen or sperm, so interspecific gene flow would often be triggered by the decline of one species (Figure 4A to B), resulting in an overabundance of heterospecific gamete landing on the rare species (Figure 4B to C). This in turn allows for the rare species to avoid local extinction and inbreeding depression through the maintenance of diversity, a process called genetic rescue (Glossary; Figure 4C to A) (Ingvarsson 2001; Carlson *et al*., 2014). However, demographic swamping (Glossary; Figure 4D), or genetic swamping (Figure 4E), where rare species are replaced by hybrids (Todesco *et al*., 2016) is often used to counter this hypothesis as too much gamete swamping could instead result in the proliferation of hybrids and extinction of the rare species. Cannon and Scher (2017) suggested that Mendelian segregation and pollen competition allow for the formation of genetic bridges among species and thus the participation in syngameons. They argue that small proportions of the gametophytes produced by F_1 hybrids would be 80-90% identical to a gametophyte produced by one of the parental species. That small portion (which could total millions of gametophytes in a heavily producing system like oaks) coupled with conspecific pollen advantage, could result in a backcross generation nearly identical to the parental types, making introgression possible without the erosion of genetic coherence. Although they used simulations based on a real oak syngameon, they limit their hypothesis to organisms with low chromosome numbers, copious gamete production, conserved genomic structure, and conspecific gamete advantage. More syngameons are being uncovered that do not follow these strict assumptions, thus future simulation studies will need to broaden their scopes and reassess the genomic mutualist hypothesis.

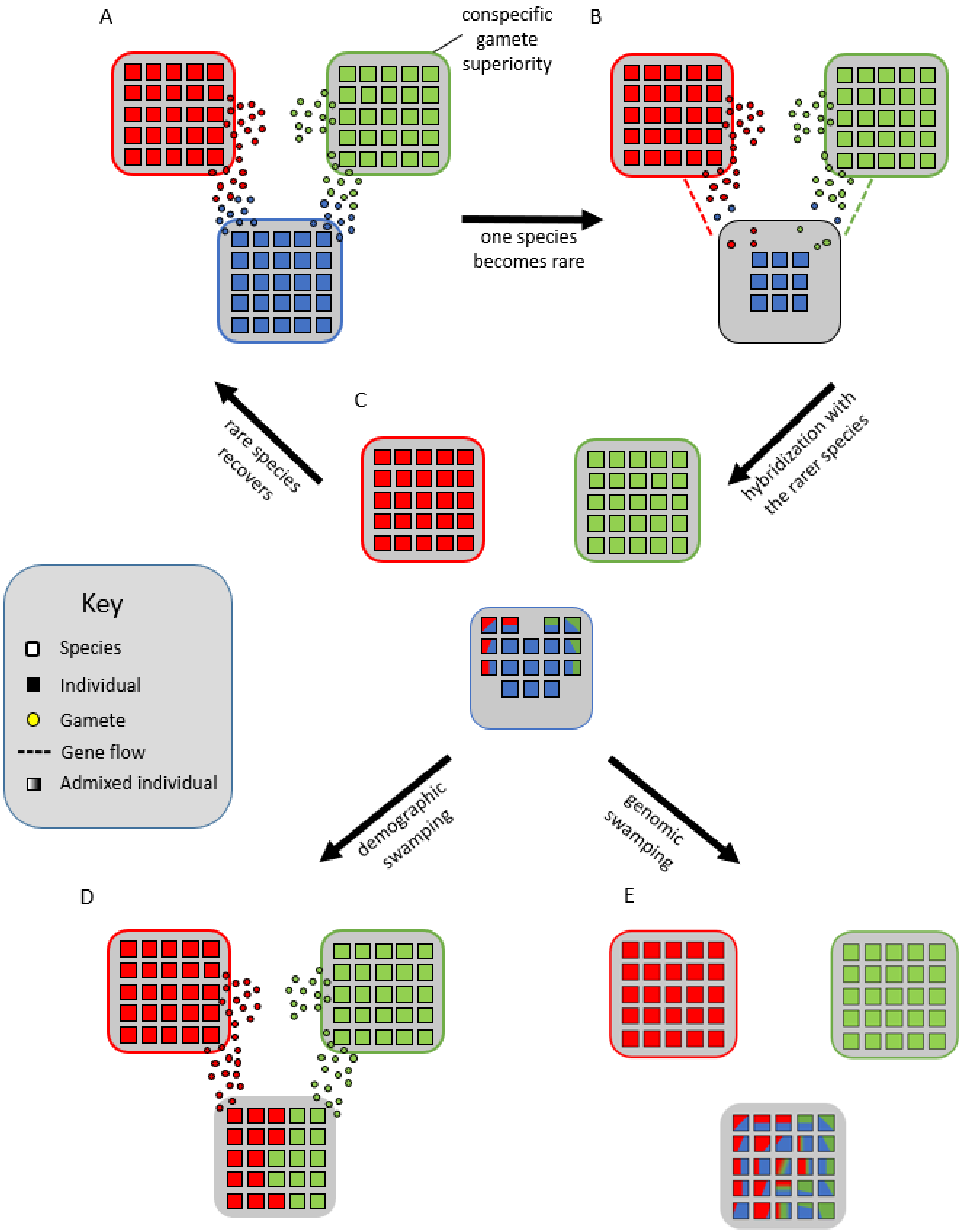


Figure 4. Genomic mutualist hypothesis in which there is a reproductive barrier favoring conspecific gametes (A) until one species becomes rare (B), wherein the gamete load from interspecific donors forces the rarer species to hybridize (C). This could lead to the rarer species benefiting from the increased genetic variation and effective population size, allowing it to overcome inbreeding depression and recover, a process known as genetic rescue (A). Alternatively, the rarity could lead to demographic swamping, where the rare species is replaced by the more abundant species through the purging of maladaptive hybrids (D), or genetic swamping, in which the rare species is replaced by admixed individuals (E).

Spatial Limitations

With the various ways syngameons are thought to form, it seems that there should be an abundance of syngameons. Perhaps the limiting factor is species distribution, meaning that, despite the number of species pairs with overlapping distributions reported to hybridize, the chances to have multiple hybridizable species with overlapping distributions is limited. In describing competition among highly diverse tropical tree communities, Cannon and Lerdau (2015, 2019) found that direct spatial proximity with close relatives was infrequent, so even in complex ecological landscapes, the chances of overlapping with a congeneric species is low. Yet even if direct spatial overlap does not frequently occur, pollen and seed could still disperse into adjacent habitats and trigger syngameonic behavior. If, however, sympatry does occur, species in a syngameon could coexist and avoid competition by diversifying into microhabitats, as demonstrated by Schmitt *et al*. (2021) in a Neotropical syngameon. Similarly, differing patterns of speciation may also play a role in limiting syngameon formation because allopatric species coming into secondary contact could be less likely to share a large enough portion of their range to overlap with more than one species. Further, the narrow hybrid zones that can result from secondary contact do not allow for the introgression of genes beyond the hybrid zone itself, which is usually at the edge of species' ranges. Alternatively, F_1s could form but reproductive barriers could prevent any backcrossing with the parental species, thus introgression would not occur, as seen in *Ligularia* (Zhang *et al*., 2018). While many syngameon participants defined here hybridize with the same single species, genes are not necessarily introgressed across all species' ranges, especially if the species hybridize at opposite ends of a range (Figure 3B,C). While this would still technically be considered a syngameon, the participants are not receiving all the benefits of the network-like structure of more sympatric syngameons. Cases of sympatric speciation may create more opportunities for multiple species to have overlapping distributions, however these scenarios are rarer (Barraclough and Vogler, 2000; Foote 2018) and usually result in the formation of a reproductive isolating barrier (Smith 1966; Schliewen *et al.*, 1994; Ottenburghs 2021), which would likely prevent any further hybridization. While the above syngameon formation hypotheses are not necessarily mutually exclusive, maintaining hybridization in multiple species at once can have compounding complications, with genetic swamping, lineage collapse, and the formation of reproductive barriers, all challenging the stability of a syngameon. If syngameons constantly fight to exist, then how are they maintained over time? The structure of known syngameons may shed light on this perplexing question.

How are Syngameons Maintained Over Evolutionary Time?

Most of the formation hypotheses above mention the episodic occurrences of gene flow within a syngameon and the limited amount of gene flow that must occur to stabilize the interactions. Yet most known examples of syngameons show extensive and constant gene flow among numerous participants (Table 1). This discrepancy in theory and practice may be due to the varying hubs of introgression (Glossary), where some species contribute more genetic information than they receive and are connected to a large number of other species through gene flow (Ladner 2012). The number of participating species can vary over geographical space and evolutionary time, with a single species (referred to as a hub) that has direct contact with multiple species, and as a result genes passively introgress through the various pathways radiating from the hub (Figure 5). In these hub-based networks, if one pathway collapses, gene flow can still be maintained through the numerous other pathways connecting the species together, as long as there are no geographic or intrinsic barriers that act to contain alleles to one hybridizing species pair. However, if a hub disappears, that will likely have a larger effect on the entire network, but the extent of that effect is not currently known.

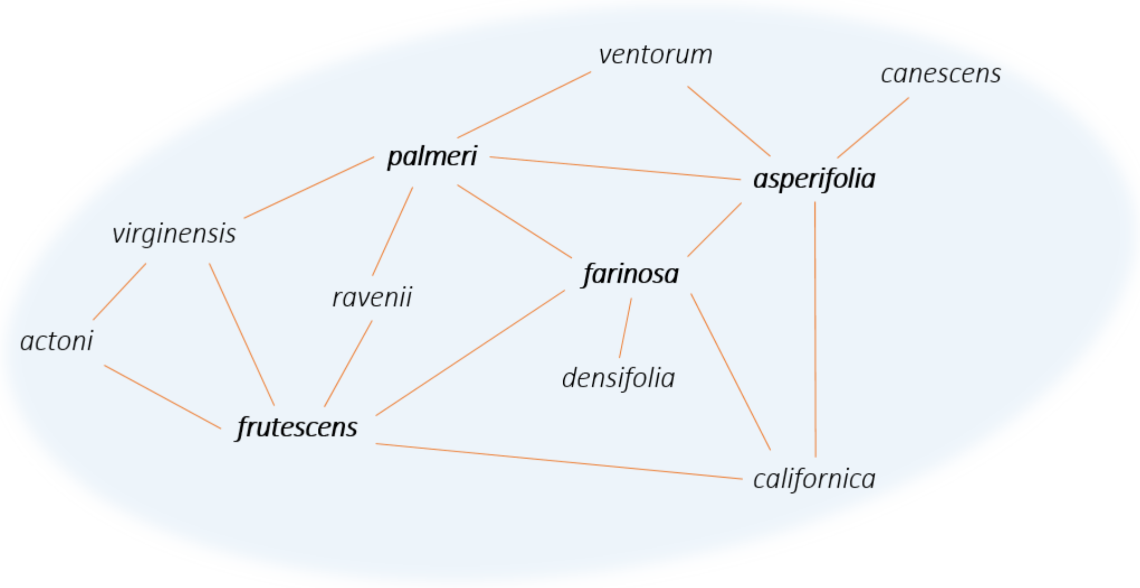


Figure 5. Network diagram of the *Encelia* syngameon (adapted from Singhal *et al*., 2021). Lines connecting species represent gene flow, with species such as *E. frutescens*, *E. farinosa*, *E. palmeri*, and *E. asperifolia* exemplifying hubs of introgression (in bold).

The direction and magnitude of these introgression pathways are rarely uniform. Boecklen (2017) used simulations to test the structure of nine natural and four artificial syngameons (Glossary), finding that a majority exhibit a nonrandom structure, with a few species dominating the patterns of introgression. He concluded that geographically widespread species would have more opportunities to hybridize than restricted ones, with the *Boechera* syngameon demonstrating a positive relationship between geographic range and number of mating combinations. The same is seen in North American white oaks, with the widespread *Quercus alba* mating with 11 out of 14 species in the syngameon (Cronk and Suarez-Gonzalez, 2018). It seems that the distribution of a species has a large impact on its ability to participate in the syngameon, with large contributors maintaining the structure of syngameons across geographical space. However, geographically widespread species may encounter more geographic and ecological barriers that could lead to population structuring or barriers to gene flow (Nikolakis *et al*., 2021). These could ultimately prevent the species, or at least certain populations, from participating in the syngameon or could result in reduced introgression beyond the hybrid zone. Additionally, the propensity to hybridize was unequal, even when species had equal opportunities to hybridize (Boecklen 2017). This suggests that there are other factors beyond range that affect the direction and magnitude of introgression within a syngameon. Genetic distance (Glossary) may be the largest of these factors, with closer related species hybridizing more readily than distant ones (Larcombe *et al*., 2015; Boecklen 2017). This would mean that the structure of syngameons is partially dependent on the relatedness of the species participating. Hypothetically, as time passes, species would become more distinct, compromising the structure and putting the maintenance of the syngameon at risk. However, the occasional gene flow events among syngameon members would counteract divergence and keep genomic distance smaller.

There are several cases (e.g. coral (Ladner 2012) and pinyon pines (Buck *et al.*, 2020)) where gene flow can favor one direction within a syngameon. The reasons for unidirectionality are numerous but include postzygotic barriers that prevent one parent from backcrossing with the hybrid offspring, such as hybrid inviability, hybrid sterility, and F_2 breakdown (Trucco *et al.*, 2009; Álvarez and Garcia-Vazquez, 2011; Yang *et al.*, 2019; Del-Rio *et al.*, 2021). This could promote a stable syngameon by preventing maladaptive hybrid derivatives from forming and only allowing the viable and fertile backcrossed individuals to proliferate. In this sense, the formation of reproductive barriers can actually maintain syngameons rather than collapsing them by preventing hybridization. On a genetic level, the uneven exchange rate of loci may represent regions that maintain functional differences between species (Ladner 2012). In corals, large sections of non-introgressing genes were found among species with high levels of overall gene flow (Ladner 2012). This suggests that loci responsible for differentiating species may be linked to loci that contribute to reproductive isolation, creating gene regions that maintain individual lineages in a syngameon, while still allowing for some gene flow. In hybridizing species of *Drosophila*, recombination rates may be reduced while chromosomal inversion rates are increased to promote divergence under gene flow, yet maintain high diversity in the rest of the genome (Barton 2020). However, selection could reduce diversity in genomic regions and result in a similar, but misleading pattern as non-introgressing loci

(Cruickshank and Hahn, 2014). Future studies will need to take both mechanisms into account by examining diversity across the whole genome, especially when taxa have recently diverged (Cruickshank and Hahn, 2014).

The evolutionary advantages and disadvantages of interspecific gene flow are well understood (Rieseberg and Wendel, 1993; Tallmon *et al*., 2004; Widmer *et al*., 2009; Abbott *et al*., 2013), but it is not known if these consequences are the same in these multispecies networks. Cannon and Petit (2020) suggest that syngameons have synergistic properties, with network-like benefits that total more than just the sum of individual species pairs. Schmitt *et al*. (2020, 2021) suggest that two contrasting evolutionary pressures are constantly acting on a syngameon, one at the species level to maximize individual species' fitness and reduce competition among species, and one at the syngameon level to increase genus survival and maximize population size. In syngameons, adaptive introgression can maintain hybrid zones through the sharing of beneficial alleles (Suarez-Gonzalez *et al*., 2018). Natural selection plays a role in maintaining the poplar syngameon when adaptive alleles are episodically exchanged across species boundaries (Chhatre *et al*., 2018). Syngameons can also have increased heterozygosity, while maintaining partial infertility among species (Chen *et al*., 2014). In the Fabaceae family, this partial infertility prevents genomes from fully merging, while still allowing gene flow to increase heterozygosity (Chen *et al*., 2014). Levi *et al.* (2019) suggested syngameons could help fuel the Red Queen arms race (Glossary) in tropical trees by increasing heterozygosity and introducing novel phenotypes. These beneficial outcomes of gene flow help maintain syngameons and can counter the negative complications that arise with hybridization. While the current definition does not differentiate between fertile hybrids that can backcross with their parental species and sterile hybrids that would prevent introgression, the hypothesized synergistic effects would likely only exist in the former situation where adaptive traits can be passed across species barriers. Further, the creation of infertile hybrids would more likely result in demographic swamping (Glossary; Figure 4D) and be detrimental to the syngameon as a whole.

There are many ways that syngameons can remain stable over long periods of evolutionary time including uneven participation, geographic distribution, genetic distance, and direction of gene flow within a syngameon. These factors can allow gene flow to occur episodically or in minute amounts, preserving the beneficial aspects of hybridization while avoiding the detrimental ones. A common misconception with hybridization is that it is ephemeral and only a stopping point on the way to reproductive isolation (Cannon 2021). While time since divergence is positively correlated with the strength of reproductive barriers (Rothfels *et al*., 2015), classic two-species hybrid zones can be stable over evolutionary time through the balance between selection and dispersal (Barton 1979), so it is reasonable that multispecies hybrid zones, while more complex, can be maintained in the same way. Without strong selection for the formation of reproduction barriers and with occasional gene flow partially homogenizing genomes, isolating barriers may take even longer to form within a syngameon, if at all. Cannon and Petit (2020) argue that syngameons do not have to be transitional or incipient phases on the way to complete speciation because reproductive isolation is not a requirement for speciation in the first place. We assert that while syngameons can be ephemeral and collapse if reproductive barriers form, they can also last for as long as species themselves, constantly fluctuating and evolving. With the potential stability of known syngameons over time, why are we just now discovering syngameons and why have we not detected more? Both the past and future of science hold the answer.

Why are Syngameons So Rare?

Hybridization in general was overlooked for decades. Considered infrequent and not important to evolution, it remained unexplored for the better half of the 1900s. Although extensive efforts were eventually made, studies of hybridization were limited to phenotypic comparisons (Anderson 1949). This initial lack of genomic data could be the main reason so few syngameons have been uncovered. With the incorporation of next generation sequencing (NGS) and whole genome data, more hybridization events are being discovered, and thus more syngameons are being uncovered (Table 1, Figure 6). Likewise, scientists are starting to recognize the importance of hybridization events and are able to describe patterns of reticulated evolution, so it is only a matter of time before more syngameons are reported.

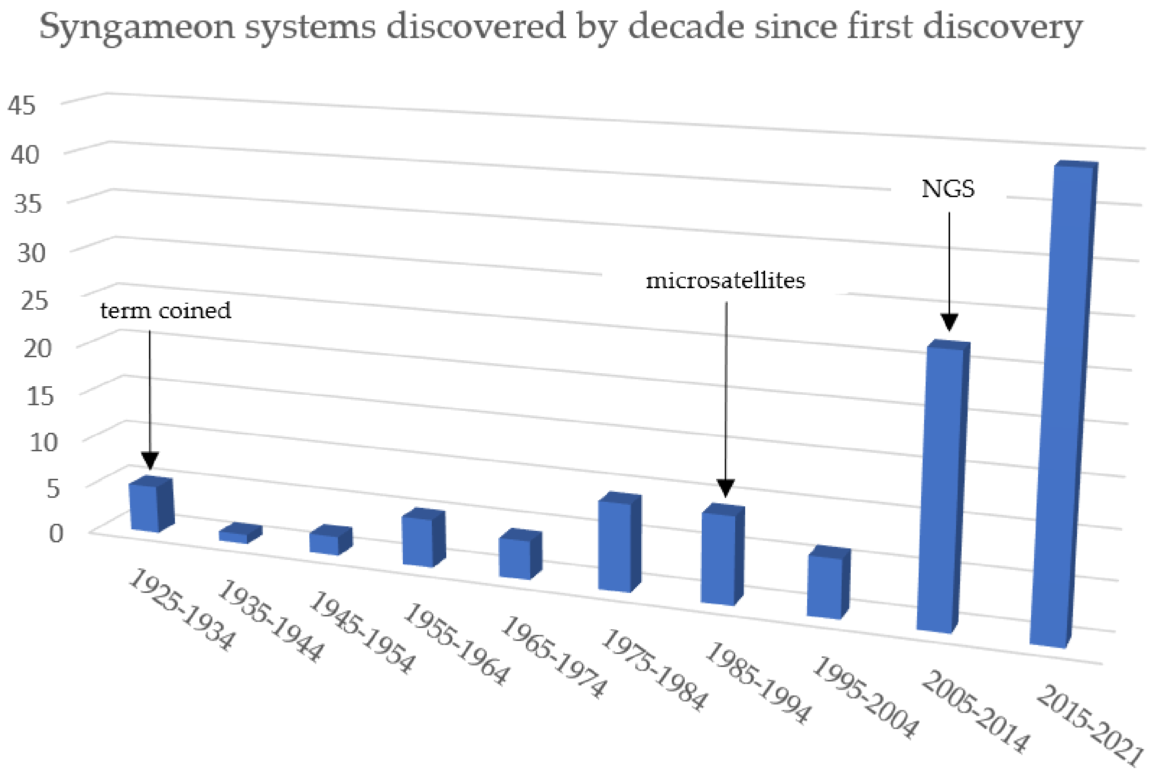


Figure 6. The number of syngameon systems discovered by decade (as of December, 2021), showing an increase in the number of discoveries, especially in the 2010s.

While our overall detection methods are improving, several factors can still prevent the discovery of syngameons. Cryptic hybrids, which are genetically of hybrid origin but morphologically appear identical to one parental species, are one such preventing factor. Ladner (2012) found cryptic hybrids in corals and Buck *et al.* (2020) found cryptic hybrids in pinyon pines, both cases exemplifying the issue that syngameons cannot be detected if hybrid individuals are not known to exist. It makes one wonder how many other systems have individuals of hybrid origin hidden among their parental species. The increased use of a combined morphological and genetic approach should help reveal cryptic hybrids in the future. Unexplored hybrid pathways are another limit on syngameon detection. Most studies explore hybrid zones by looking at two parental species and their resulting offspring without considering the potential for multispecies introgression. As studies expand to incorporate more species, we may find that hybridization extends beyond species pairs into syngameons. In pinyon pines, for example, Buck *et al.* (2020) found that a complex originally thought to be composed of two hybridizing species actually consisted of three species undergoing tridirectional gene flow. The same pattern was found in poplars (Chhatre *et al.*, 2018). The genetic bridge hypothesis (Cannon and Scher, 2017) postulates another reason why syngameons may go undetected. The minute amounts of genetic information that are introgressed from a F_1 hybrid back into a parent species through backcrossing may result in gametes that are indistinguishable from the parental species. This is similar to the cryptic hybrids problem, except that the genetic bridge between species is undetectable, while the hybrids rarely pass the backcrossed F_1 generation. Finally, as brought up in the rapid radiation hypothesis, it is difficult to prove that syngameons occurred in the past so they may remain hidden by time until molecular, coalescent, and ancient introgression techniques improve (Ottenburghs 2021).

While the potential for future syngameon discovery could increase with the incorporation of new technology and more species, the global climate is rapidly changing due to anthropogenic activities (Oreskes 2004). What does this mean for the future of syngameons? Can syngameons generate the right combination of genes to save the member species from extinction? How will climate change affect syngameons and how can we conserve species that participate in gene flow networks?

The Conservation and Future of Syngameons

With climate change, several species ranges are shifting polewards or disappearing altogether (Parmesan and Yohe, 2003, Williams *et al.*, 2007) and novel interspecies interactions are being established (Chunco 2014). These migrations and novel interactions could lead to new hybridization events between previously isolated species (Ottenburghs 2021) and could result in the formation of syngameons, especially at range edges (Caujapé-Castells and Bramwell, 2011) and during colonization events (Cronk and Suarez-Gonzalez, 2018). Additionally, anthropogenic introductions and disturbed habitats can create novel niches and allow hybrids to thrive (Stanford 1995; Ottenburghs 2021). The incorporation of adaptive alleles, heterozygosity, and an increase in effective population size via participation in a syngameon could be critical to the survival of species in a quickly changing climate (Cannon and Lerdau, 2015; Cannon and Petit, 2020). Conversely, contracting ranges and increasing extinction rates could result in the collapse of syngameons if participating species begin to disappear or become allopatric. However, the degree to which one species affects a syngameon as a whole remains unknown.

The focal unit of conservation is a species. The definition of what constitutes a species is widely debated (Wheeler and Meier, 2000). A "species" under the Endangered Species Act (1973) (but see 1978 amendment) includes “any subspecies of fish, wildlife, or plants and any other group of fish or wildlife of the same species or smaller taxa in common spatial arrangement that interbreed when mature”. However, this loose definition does not take hybridization into consideration. Gene flow has always been a taxonomic issue since the early species concept debates. Hybridization, especially at a multispecies level, went against the standing concept of a biological species (Mayr 1942). Thus, the discovery of syngameons made taxonomists rethink species concepts, leading (Grant 1971) who studied an oak syngameon, to create the ecological species concept. However, arbitrary cutoffs must be made to distinguish niches and some syngameon participants may not occupy different niches. Therefore, it is more appropriate to consider the whole syngameon as a biological conservation unit (Cronk and Suarez-Gonzalez, 2018). While individual members of a syngameon are not reproductively isolated from each other, syngameons are isolated from other syngameons and non-participating species (Stanford 1995). However, conservation efforts should not necessarily treat syngameons as they would a single species, because doing so would essentially collapse all the lineages into one and reduce the conservation importance of the individual species. Each participating species should be conserved with the assumption that individual contributions have widespread effects across the whole multispecies network (Cannon 2021). This is particularly important as it has been suggested that in some instances this multidirectional and recurrent hybridization has created new hybrid species (Rieseberg *et al*., 1990, 1991; Rieseberg 1991a; Linan *et al*., 2021; Buck *et al*., in review).

With limited funds, conservationists often find they cannot protect every species but must focus their efforts. In the case of a syngameon, the structure must be taken into account, with a priority on hub species which have a larger effect on the network as a whole. As many hub species encompass larger ranges, they do not usually represent a conservation concern and protecting them might require a large amount of resources. However, we argue they need to be considered in conservation genetic plans as they harbor important genetic diversity needed for the evolution of the complex. As an individual's effects on the structure of a syngameon are still largely unknown, it is hard to predict how the loss of one species could affect the syngameon as a whole. The decline of a single species may result in the collapse of the whole syngameon and potentially lead to the extinction of the remaining species. Alternatively, if one population is participating in the syngameon but the rest are not (Dowling *et al*., 2016), as is possible in edge-range syngameons, limited conservation efforts can equally focus on that population and core populations to preserve syngameon structure. Lastly, gene flow should be considered as a potential tool for conservation because the immediate increase in heterozygosity and the adaptive introgression of beneficial alleles could be critical to the survival of endangered species, with some authors arguing that the benefits outweigh the potential dangers (Hamilton and Miller, 2016; Chan *et al*., 2019; Quilodrán *et al*., 2020). However, human-induced hybridization events should be carefully planned and controlled to avoid the outbreeding depression effects seen in unregulated anthropogenic gene flow events (Ottenburghs 2021). This highlights the need for a systematic change in the legal framework of conservation policy. Current conservation efforts are reserved for well defined species, while hybrids are largely ignored and discounted as "genetic erosion" or "pollution" (Ottenburghs 2021). The Endangered Species Act should expand its protection to not only hybridizing pairs, but also complexes like syngameons.

Conclusion

Almost a century ago, Lotsy (1925) recognized the complexity of a syngameon as species that readily mate among several species but also recognized the difficulty of detecting multidirectionally hybridization by his statement, “Can a careful study in nature… reveal the true relationship between the various individuals within the genus, can it decide which of the forms are hybrids, which species and from which combination of the latter the hybrids arose? To my way of thinking, not”. The advancement of next generation sequencing has opened the possibility to carefully and precisely answer these questions. Not only has this technology enabled us to detect multidirectional hybridization, the magnitude of gene flow and the percentage of the parental ancestry, but it has also demonstrated that syngameons are not as rare as previously thought. In our comprehensive literature review, we found that over the past century, reports of syngameons have increased in relation to the use of genetic markers. Just over the past decade the numbers have increased (Figure 6) and we predict they may keep rising. Future syngameon studies should focus on understanding how syngameons form and remain stable over long periods of evolutionary time. As more syngameons are discovered, formation hypotheses can be tested and compared. Combining biogeographic and population-level genetic data may give insight into ancient introgression events that coincide with range contacts, colonizations, and rapid radiations. More simulations can be run to detect the structure of syngameons, which may shed light on individual species’ roles in these multispecies networks. Discovering how the individual species affects the structure of a syngameon as a whole remains the largest enigma of the syngameon. If ranges contract out of sympatry or species go extinct, researchers can examine the resulting effects on the other participating species.

Funding: This research received no external funding.

Acknowledgments: We’d like to thank Polly Campbell and Leonard Nunney at UCR for their feedback to the early stages of the manuscript. We’d also like to thank the Flores-Rentería lab, Arun Sethuraman, and the SDSU EB Journal Club for their feedback and constructive criticisms.

Author Contributions: Conceptualization, R.B. and L.F.R.; writing—original draft preparation, R.B.; writing—review and editing, R.B. and L.F.R. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Glossary

- Artificial syngameon - A syngameon created through manual/human-induced hybridization
- Demographic swamping - A potential consequence of hybridization in which hybrid fitness is significantly lower than parental fitness, resulting in the decline of one or both parental lineages due to the wasteful production of unfit hybrids.
- Genetic distance - A measure of genetic divergence between species that indicates how closely related those species are.
- Genetic rescue - An increase in fitness due to hybridization resulting in the recovery from inbreeding depression.
- Genetic swamping - A potential consequence of hybridization where one or both parental lineages are replaced by equally or more fit admixed individuals.
- Hub species - A species in a syngameon that hybridizes with several other species thus connecting multiple participants through gene flow.
- Red Queen’s hypothesis - An evolutionary hypothesis about the coevolution of competing species wherein one species must constantly adapt to compete against or evade their opposing species that is also constantly adapting to defeat them.
- Reinforcement - A process where selection against unfit hybrids leads to the formation of reproductive barriers and fortifies species boundaries.
- Syngameon - A complex of three or more species connected through gene flow

References

Abbott, R., Albach, D., Ansell, S., Arntzen, J.W., Baird, S.J., Bierne, N., Boughman, J., Brelsford, A., Buerkle, C.A., Buggs, R. and Butlin, R.K., 2013. Hybridization and speciation. *Journal of evolutionary biology*, *26*(2), pp.229-246.

Ackermann, R.R., Arnold, M.L., Baiz, M.D., Cahill, J.A., Cortés-Ortiz, L., Evans, B.J., Grant, B.R., Grant, P.R., Hallgrimsson, B., Humphreys, R.A. and Jolly, C.J., 2019. Hybridization in human evolution: Insights from other organisms. *Evolutionary Anthropology: Issues, News, and Reviews*, *28*(4), pp.189-209.

Alexander, Patrick J., Michael D. Windham, James B. Beck, Ihsan A. Al-Shehbaz, Loreen Allphin, and C. Donovan Bailey. "Weaving a tangled web: divergent and reticulate speciation in *Boechera fendleri sensu lato* (Brassicaceae: Boechereae)." *Systematic Botany* 40, no. 2 (2015): 572-596.

Álvarez, D. and Garcia-Vazquez, E., 2011. Maintenance of asymmetric hybridization between Atlantic salmon (Salmo salar) and brown trout (Salmo trutta) via postzygotic barriers and paternal effects. *Canadian Journal of Fisheries and Aquatic Sciences*, *68*(4), pp.593-602.

Anderson, E., 1949. Introgressive hybridization. *Introgressive hybridization.*

Arduino, P., Verra, F., Cianchi, R., Rossi, W., Corrias, B. and Bullini, L., 1996. Genetic variation and natural hybridization between *Orchis laxiflora* and *Orchis palustris* (Orchidaceae). *Plant Systematics and Evolution*, *202*(1), pp.87-109.

Arrigoni, R., Benzoni, F., Terraneo, T.I., Caragnano, A. and Berumen, M.L., 2016. Recent origin and semi-permeable species boundaries in the scleractinian coral genus *Stylophora* from the Red Sea. *Scientific Reports*, *6*(1), pp.1-13.

Baiakhmetov, E., Nowak, A., Gudkova, P.D. and Nobis, M., 2020. Morphological and genome-wide evidence for natural hybridisation within the genus *Stipa* (Poaceae). *Scientific reports*, *10*(1), pp.1-14.

Barker, M.J., 2007. The empirical inadequacy of species cohesion by gene flow. *Philosophy of Science*, *74*(5), pp.654-665.

Barkworth, M.E. and Adams, R.P., 1993. *Flora of North America: Volume 2: Pteridophytes and Gymnosperms* (Vol. 2). Oxford University Press on Demand.

Barnes, B.V. and Dancik, B.P., 1985. Characteristics and origin of a new birch species, *Betula murrayana*, from southeastern Michigan. *Canadian journal of botany*, *63*(2), pp.223-226.

Barraclough, T.G. and Vogler, A.P., 2000. Detecting the geographical pattern of speciation from species-level phylogenies. *The American Naturalist*, *155*(4), pp.419-434.

Barton, N.H., 1979. The dynamics of hybrid zones. *Heredity*, *43*(3), pp.341-359.

Barton, N.H., 2020. On the completion of speciation. *Philosophical Transactions of the Royal Society B*, *375*(1806), p.20190530.

Beaudry, J.R., 1960. The species concept: its evolution and present status. *The species concept: its evolution and present status.*

Beddows, I. and Rose, L.E., 2018. Factors determining hybridization rate in plants: A case study in Michigan. *Perspectives in Plant Ecology, Evolution and Systematics*, *34*, pp.51-60.

Beeks, R.M., 1962. Variation and hybridization in southern California populations of *Diplacus* (Scrophulariaceae). *Aliso: A Journal of Systematic and Floristic Botany*, *5*(2), pp.83-122.

Beltrán, M., Jiggins, C.D., Bull, V., Linares, M., Mallet, J., McMillan, W.O. and Bermingham, E., 2002. Phylogenetic discordance at the species boundary: comparative gene genealogies among rapidly radiating *Heliconius* butterflies. *Molecular biology and evolution*, *19*(12), pp.2176-2190.

Boecklen, W.J., 2017. Topology of syngameons. *Ecology and evolution*, *7*(24), pp.10486-10491.

Bog, M., 2016. *Hybridisation and its consequences: Population genomics, herbivory, and phytochemistry in the* Senecio nemorensis *syngameon*. Doctoral dissertation, Universität Regensburg.

Brownsey, P.J., 1977. A taxonomic revision of the New Zealand species of *Asplenium*. *New Zealand journal of botany*, *15*(1), pp.39-86.

Brouillet, L., Semple, J.C., Allen, G.A., Chambers, K.L. and Sundberg, S.D., 2006. *Symphyotrichum*. *Flora of North America*, *20*, pp.465-539.

Brunsfeld, S.J., Soltis, D.E. and Soltis, P.S., 1992. Evolutionary patterns and processes in *Salix* sect. *Longifoliae*: evidence from chloroplast DNA. *Systematic Botany*, pp.239-256.

Buck, R., Hyasat, S., Hossfeld, A. and Flores-Rentería, L., 2020. Patterns of hybridization and cryptic introgression among one-and four-needled pinyon pines. *Annals of Botany*, *126*(3), pp.401-411.

Buck, R.; Ortega-Del Vecchyo, D.; Whipple, A.V.; Gehring, C.; Michelson, R.; Flores-Rentería, D.; Klein, B.; Flores-Rentería, L. 2022. Sequential hybridization may have facilitated ecological transitions in the Southwestern pinyon pine syngameon. *New Phytologist*, *in review*.

Budd, A.F., 2010. Tracing the long-term evolution of a species complex: Examples from the *Montastraea* "*annularis*" complex. *Palaeoworld*, *19*(3-4), pp.348-356.

Bureš, P., Šmarda, P., Rotreklova, O., Oberreiter, M., Burešová, M., Konečný, J., Knoll, A., Fajmon, K. and Šmerda, J., 2010. Pollen viability and natural hybridization of Central European species of *Cirsium*. *Preslia*, *82*(4), pp.391-422.

Butelli, E., Licciardello, C., Ramadugu, C., Durand-Hulak, M., Celant, A., Recupero, G.R., Froelicher, Y. and Martin, C., 2019. Noemi controls production of flavonoid pigments and fruit acidity and illustrates the domestication routes of modern citrus varieties. *Current Biology*, *29*(1), pp.158-164.

Cannon, C.H., 2021. Is speciation an inexorable march to reproductive isolation?. *Molecular Ecology*, *30*(18), pp.4349-4352.

Cannon, C.H. and Lerdau, M., 2015. Variable mating behaviors and the maintenance of tropical biodiversity. *Frontiers in Genetics*, *6*, p.183.

Cannon, C.H. and Lerdau, M.T., 2019. Demography and destiny: The syngameon in hyperdiverse systems. *Proceedings of the National Academy of Sciences*, *116*(17), pp.8105-8105.

Cannon, C.H. and Petit, R.J., 2020. The oak syngameon: more than the sum of its parts. *New Phytologist*, *226*(4), pp.978-983.

Cannon, C.H. and Scher, C.L., 2017. Exploring the potential of gametic reconstruction of parental genotypes by F1 hybrids as a bridge for rapid introgression. *Genome*, *60*(9), pp.713-719.

Carlson, S.M., Cunningham, C.J. and Westley, P.A., 2014. Evolutionary rescue in a changing world. *Trends in Ecology & Evolution*, *29*(9), pp.521-530.

Carr, G.D., 1978. Hybridization of the Hawaiian Silversword complex. In Proceedings of the Second Conference in Natural Sciences, Honolulu, HI, June 1-3.

Case, F.J.; Case, R., 1997. *Trilliums*. Timber Press, p.285.

Caujapé-Castells, J. and Bramwell, D., 2011. Jesters, red queens, boomerangs and surfers: a molecular outlook on the diversity of the Canarian endemic flora. *The biology of island floras*, pp.284-324.

Caujapé-Castells, J., García-Verdugo, C., Marrero-Rodríguez, Á., Fernández-Palacios, J.M., Crawford, D.J. and Mort, M.E., 2017. Island ontogenies, syngameons, and the origins and evolution of genetic diversity in the Canarian endemic flora. *Perspectives in Plant Ecology, Evolution and Systematics*, *27*, pp.9-22.

Chan, W.Y., Hoffmann, A.A., and van Oppen, M.J.H., 2019. Hybridization as a conservation management tool. *Conserv Letters, 12,* pp.1–11.

Chen, S.C., Cannon, C.H., Kua, C.S., Liu, J.J. and Galbraith, D.W., 2014. Genome size variation in the Fagaceae and its implications for trees. *Tree genetics & genomes*, *10*(4), pp.977-988.

Chhatre, V.E., Evans, L.M., DiFazio, S.P. and Keller, S.R., 2018. Adaptive introgression and maintenance of a trispecies hybrid complex in range-edge populations of *Populus*. *Molecular ecology*, *27*(23), pp.4820-4838.

Chunco, A.J., 2014. Hybridization in a warmer world. *Ecology and evolution*, *4*(10), pp.2019-2031.

Cinget, B., de Lafontaine, G., Gérardi, S. and Bousquet, J., 2015. Integrating phylogeography and paleoecology to investigate the origin and dynamics of hybrid zones: insights from two widespread North American firs. *Molecular Ecology*, *24*(11), pp.2856-2870.

Clapham, A.R., Tutin, T.G. and Moore, D.M., 1990. *Flora of the British isles*. CUP Archive.

Colbourne, J.K., Wilson, C.C. and Hebert, P.D.N., 2006. The systematics of Australian *Daphnia* and *Daphniopsis* (Crustacea: Cladocera): a shared phylogenetic history transformed by habitat-specific rates of evolution. *Biological Journal of the Linnean Society*, *89*(3), pp.469-488.

Cronk, Q.C. and Suarez-Gonzalez, A., 2018. The role of interspecific hybridization in adaptive potential at range margins. *Molecular Ecology, 27*, pp.4653-4656.

Crowder, A.A., Pearson, M.C., Grubb, P.J. and Langlois, P.H., 1990. Drosera L. *Journal of Ecology*, *78*(1), pp.233-267.

Cruickshank, T.E. and Hahn, M.W., 2014. Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. *Molecular ecology*, *23*(13), pp.3133-3157.

Cuénot, L., 1951. *L'Evolution Biologique: Les Faits, Les Incertitudes*. Masson et Cie: Paris, France, p.594.

Cui, R., Schumer, M., Kruesi, K., Walter, R., Andolfatto, P. and Rosenthal, G.G., 2013. Phylogenomics reveals extensive reticulate evolution in *Xiphophorus* fishes. *Evolution*, *67*(8), pp.2166-2179.

Curto, M., Puppo, P., Kratschmer, S. and Meimberg, H., 2017. Genetic diversity and differentiation patterns in *Micromeria* from the Canary Islands are congruent with multiple colonization dynamics and the establishment of species syngameons. *BMC Evolutionary Biology*, *17*(1), pp.1-16.

Debouck, D.G., 1992. Views on variability in *Phaseolus* beans. *Ann. Rep. Bean Improv. Coop*, *35*, pp.9-10.

Del-Rio, G., Rego, M.A., Whitney, B.M., Schunck, F., Silveira, L.F., Faircloth, B.C. and Brumfield, R.T., 2022. Displaced clines in an avian hybrid zone (Thamnophilidae: Rhegmatorhina) within an Amazonian interfluve. *Evolution*, *76*(3), pp.455-475.

dePamphilis, C.W. and Wyatt, R., 1989. Hybridization and introgression in buckeyes (Aesculus: Hippocastanaceae): a review of the evidence and a hypothesis to explain long-distance gene flow. *Systematic Botany*, pp.593-611.

Du Rietz, G.E., 1930. *The fundamental units of biological taxonomy*. Uppsala, Sweden: Svensk Botaniska Foreningen.

Dowling, T.E., Markle, D.F., Tranah, G.J., Carson, E.W., Wagman, D.W. and May, B.P., 2016. Introgressive hybridization and the evolution of lake-adapted catostomid fishes. *PloS one*, *11*(3), p.e0149884.

Dreher, S.E., 2014. *Interspecific Hybridization in Sphaeralcea (Malvaceae)*. The Claremont Graduate University.

Endangered Species Act of 1973, D.O.T. Interior, Editor. 1973. p. 44.

Flake, R.H., Urbatsch, L. and Turner, B.L., 1978. Chemical documentation of allopatric introgression in Juniperus. *Systematic Botany*, pp.129-144.

Flores-Rentería, L., Rymer, P.D. and Riegler, M., 2017. Unpacking boxes: integration of molecular, morphological and ecological approaches reveals extensive patterns of reticulate evolution in box eucalypts. *Molecular Phylogenetics and Evolution*, *108*, pp.70-87.

Foote, A.D., 2018. Sympatric speciation in the genomic era. *Trends in Ecology & Evolution*, *33*(2), pp.85-95.

Frantz, L.A., Schraiber, J.G., Madsen, O., Megens, H.J., Bosse, M., Paudel, Y., Semiadi, G., Meijaard, E., Li, N., Crooijmans, R.P. and Archibald, A.L., 2013. Genome sequencing reveals fine scale diversification and reticulation history during speciation in *Sus*. *Genome biology*, *14*(9), pp.1-12.

Freeland, J.R. and Boag, P.T., 1999. The mitochondrial and nuclear genetic homogeneity of the phenotypically diverse Darwin's ground finches. *Evolution*, *53*(5), pp.1553-1563.

Fukami, H., Budd, A.F., Levitan, D.R., Jara, J., Kersanach, R. and Knowlton, N., 2004. Geographic differences in species boundaries among members of the *Montastraea annularis* complex based on molecular and morphological markers. *Evolution*, *58*(2), pp.324-337.

Garner, A.G., Goulet, B.E., Farnitano, M.C., Molina-Henao, Y.F. and Hopkins, R., 2018. Genomic signatures of reinforcement. *Genes*, *9*(4), p.191.

Gilbert, L.E., 2003. Adaptive novelty through introgression in *Heliconius* wing patterns: evidence for shared genetic “tool box” from synthetic hybrid zones and a theory of diversification. *Ecology and evolution taking flight: butterflies as model systems*, pp. 281-318.

Godfrey, L. and Marks, J., 1991. The nature and origins of primate species. *American Journal of Physical Anthropology*, *34*(S13), pp.39-68.

Gottlieb, L.D., 1968. Hybridization between *Arctostaphylos viscida* and *A. canescens* in Oregon. *Brittonia*, *20*(1), pp.83-93.

Gould, S.J. and Woodruff, D.S., 1990. History as a cause of area effects: an illustration from *Cerion* on Great Inagua, Bahamas. *Biological Journal of the Linnean Society*, *40*(1), pp.67-98.

Granados Cifuentes, C.A., 2008. Interspecific and intragenomic variation of the internal transcribed spacer2 (ITS2, rDNA) and their consequence in the evolution of eastern pacific octocorals.

Grant, B.R. and Grant, P.R., 1989. *Evolutionary dynamics of a natural population: the large cactus finch of the Galápagos*. University of Chicago Press.

Grant, B.R. and Grant, P.R., 1996. High survival of Darwin's finch hybrids: effects of beak morphology and diets. *Ecology*, *77*(2), pp.500-509.

Grant, P.R. and Grant, B.R., 1992. Hybridization of bird species. *Science*, *256*(5054), pp.193-197.

Grant, P.R. and Grant, B.R., 2020. Triad hybridization via a conduit species. *Proceedings of the National Academy of Sciences*, *117*(14), pp.7888-7896.

Grant, V., 1971. *Plant speciation.* Columbian University Press: New York, NY.

Grant, V., 1981. *Plant speciation*. 2nd ed., Columbia University Press, New York, NY.

Grant, V. 1957. The plant species in theory and practice. In: *The species problem,* Mayr, E., ed. American Association for the Advancement of Science, Washington D.C., pp. 39-80

.

Grant, V. and Grant, K.A., 1979. Systematics of the *Opuntia phaeacantha* group in Texas. *Botanical Gazette*, *140*(2), pp.199-207.

Hamilton, J.A., De la Torre, A.R. and Aitken, S.N., 2015. Fine-scale environmental variation contributes to introgression in a three-species spruce hybrid complex. *Tree Genetics & Genomes*, *11*(1), pp.1-14.

Hamilton, J.A. and Miller, J.M., 2016. Adaptive introgression as a resource for management and genetic conservation in a changing climate. *Conservation Biology*, *30*(1), pp.33-41.

Hammer, M.F., Woerner, A.E., Mendez, F.L., Watkins, J.C. and Wall, J.D., 2011. Genetic evidence for archaic admixture in Africa. *Proceedings of the National Academy of Sciences*, *108*(37), pp.15123-15128.

Hardin, J.W., 1975. Hybridization and introgression in *Quercus alba*. *Journal of the Arnold Arboretum*, *56*(3), pp.336-363.

Hebert, P.D. and Wilson, C.C., 1994. Provincialism in plankton: endemism and allopatric speciation in Australian *Daphnia*. *Evolution*, *48*(4), pp.1333-1349.

Hedrén, M., 2004. Species delimitation and the partitioning of genetic diversity–an example from the *Carex flava* complex (Cyperaceae). *Biodiversity & Conservation*, *13*(2), pp.293-316.

Heiser, C., 1949. Study in the evolution of the sunflower species *Helianthus annuus* and *H. bolanderi*. *Bot.*, *23*, pp.157-208.

Heiser Jr, C.B., 1951a. Hybridization in the annual sunflowers: *Helianthus annuus* x *H. debilis* var. *cucumerifolius*. *Evolution*, pp.42-51.

Heiser Jr, C.B., 1951b. Hybridization in the annual sunflowers: *Helianthus annuus* x *H. argophyllus*. *The American Naturalist*, *85*(820), pp.65-72.

Heiser Jr, C.B., 1947. Hybridization between the sunflower species *Helianthus annuus* and *H. petiolaris*. *Evolution*, pp.249-262.

Hey, J., Chung, Y., Sethuraman, A., Lachance, J., Tishkoff, S., Sousa, V.C. and Wang, Y., 2018. Phylogeny estimation by integration over isolation with migration models. *Molecular biology and evolution*, *35*(11), pp.2805-2818.

Holliday, T.W., 2006. Neanderthals and modern humans: an example of a mammalian syngameon?. In *Neanderthals revisited: New approaches and perspectives* (pp. 281-297). Springer, Dordrecht.

Ingvarsson, P.K., 2001. Restoration of genetic variation lost–the genetic rescue hypothesis. *Trends in Ecology & Evolution*, *16*(2), pp.62-63.

Kagawa, K. and Seehausen, O., 2020. The propagation of admixture-derived adaptive radiation potential. *Proceedings of the Royal Society B*, *287*(1934), p.20200941.

Kubota, K. and Sota, T., 1998. Hybridization and speciation in the carabid beetles of the subgenus *Ohomopterus* (Coleoptera, Carabidae, genus *Carabus*). *Researches on Population Ecology*, *40*(2), pp.213-222.

Kumar, V., Lammers, F., Bidon, T., Pfenninger, M., Kolter, L., Nilsson, M.A. and Janke, A., 2017. The evolutionary history of bears is characterized by gene flow across species. *Scientific Reports*, *7*(1), pp.1-10.

Ladner, J.T., 2012. *Hidden diversity in corals and their endosymbionts*. Doctoral dissertation, Stanford University.

Larsen, P.A., Marchán-Rivadeneira, M.R. and Baker, R.J., 2010. Natural hybridization generates mammalian lineage with species characteristics. *Proceedings of the National Academy of Sciences*, *107*(25), pp.11447-11452.

Larcombe, M.J., Holland, B., Steane, D.A., Jones, R.C., Nicolle, D., Vaillancourt, R.E. and Potts, B.M., 2015. Patterns of reproductive isolation in *Eucalyptus*—a phylogenetic perspective. *Molecular Biology and Evolution*, *32*(7), pp.1833-1846.

Laureto, P.J. and Barkman, T.J., 2011. Nuclear and chloroplast DNA suggest a complex single origin for the threatened allopolyploid Solidago houghtonii (Asteraceae) involving reticulate evolution and introgression. *Systematic Botany*, *36*(1), pp.209-226..

Leduc-Robert, G. and Maddison, W.P., 2018. Phylogeny with introgression in Habronattus jumping spiders (Araneae: Salticidae). *BMC evolutionary biology*, *18*(1), pp.1-23.

Lenz, L.W., 1959. Hybridization and speciation in the Pacific Coast irises. *Aliso: A Journal of Systematic and Floristic Botany*, *4*(2), pp.237-309.

Levi, T., Barfield, M., Holt, R.D. and Terborgh, J., 2019. Reply to Cannon and Lerdau: Maintenance of tropical forest tree diversity. *Proceedings of the National Academy of Sciences*, *116*(17), pp.8106-8106.

Lewis, W.H., 2008. *Rosa carolina* (Rosaceae) subspecies and hybrids in eastern and midwestern United States, Canada, and Mexico. *Novon: A Journal for Botanical Nomenclature*, *18*(2), pp.192-198.

Li, X., Wei, G., El-Kassaby, Y.A. and Fang, Y., 2021. Hybridization and introgression in sympatric and allopatric populations of four oak species. *BMC plant biology*, *21*(1), pp.1-14.

Lipman, M.J., Chester, M., Soltis, P.S. and Soltis, D.E., 2013. Natural hybrids between *Tragopogon mirus* and *T. miscellus* (Asteraceae): a new perspective on karyotypic changes following hybridization at the polyploid level. *American Journal of Botany*, *100*(10), pp.2016-2022.

Linan, A.G., Lowry, P.P., Miller, A.J., Schatz, G.E., Sevathian, J.C. and Edwards, C.E., 2021. RAD-sequencing reveals patterns of diversification and hybridization, and the accumulation of reproductive isolation in a clade of partially sympatric, tropical island trees. *Molecular Ecology*, *30*(18), pp.4520-4537.

Lint, H.L., 1977. *A revision of* Juncus *subgenus* Genuini *(Juncaceae) in the Pacific States*. Doctoral dissertation, Oregon State University.

Liu, Y., Li, D., Zhang, Q., Song, C., Zhong, C., Zhang, X., Wang, Y., Yao, X., Wang, Z., Zeng, S. and Wang, Y., 2017. Rapid radiations of both kiwifruit hybrid lineages and their parents shed light on a two-layer mode of species diversification. *New Phytologist*, *215*(2), pp.877-890.

Liu, Y., Liu, Y. and Huang, H., 2010. Genetic variation and natural hybridization among sympatric *Actinidia* species and the implications for introgression breeding of kiwifruit. *Tree genetics & genomes*, *6*(5), pp.801-813.

Lotsy, J.P., 1925. Species or linneon. *Genetica*, *7*(5), pp.487-506.

Love, R.M., 1954. Interspecific hybridization in Stipa. II. Hybrids of *S. cernua*, *S. lepida*, and *S. pulchra*. *American Journal of Botany*, pp.107-110.

Mallet, J., 2005. Hybridization as an invasion of the genome. *Trends in ecology & evolution*, *20*(5), pp.229-237.

Mallet, J., Beltrán, M., Neukirchen, W. and Linares, M., 2007. Natural hybridization in heliconiine butterflies: the species boundary as a continuum. *BMC evolutionary biology*, *7*(1), pp.1-16.

Manrique Rodríguez, N.A., 2008. Phylogenetic reconstruction of the genus *pacifigorgia* (Gorgoniidae: octocorallia) using the internal transcribed spacer ITS2. Master's thesis, Universidad de Los Andes.

Marcks, B.G., 1974. Preliminary reports on the flora of Wisconsin. 66. Cyperaceae II--Sedge family II. the genus Cyperus: The umbrella sedges. *Transactions of the Wisconsin Academy of Sciences, Arts, and Letters*.

Martin, S.H., Davey, J.W. and Jiggins, C.D., 2015. Evaluating the use of ABBA–BABA statistics to locate introgressed loci. *Molecular biology and evolution*, *32*(1), pp.244-257.

Matute, D.R. and Ayroles, J.F., 2014. Hybridization occurs between *Drosophila simulans* and *D. sechellia* in the Seychelles archipelago. *Journal of evolutionary biology*, *27*(6), pp.1057-1068.

Mayr, E., 1942. Systematics and the origin of species. Columbia University Press, New York, NY.

Meier, J.I., Marques, D.A., Mwaiko, S., Wagner, C.E., Excoffier, L. and Seehausen, O., 2017. Ancient hybridization fuels rapid cichlid fish adaptive radiations. *Nature communications*, *8*(1), pp.1-11.

McElwee-Adame, A., 2020. Personal Communication, San Diego State University, San Diego, CA.

Moran, B.M., Payne, C., Langdon, Q., Powell, D.L., Brandvain, Y. and Schumer, M., 2021. The genomic consequences of hybridization. *Elife*, *10*, p.e69016.

Nielsen, E.L., 1939. A taxonomic study of the genus *Amelanchier* in Minnesota. *The American Midland Naturalist*, *22*(1), pp.160-206.

Nikolakis, Z.L., Orton, R. and Crother, B.I., 2021. Fine scale population structure and extensive gene flow within an Eastern Nearctic snake complex (*Pituophis melanoleucus*). *bioRxiv*. *51*, pp.133-146.

Nunes, M.D., Orozco‐ter Wengel, P., Kreissl, M. and Schlötterer, C., 2010. Multiple hybridization events between *Drosophila simulans* and *Drosophila mauritiana* are supported by mtDNA introgression. *Molecular ecology*, *19*(21), pp.4695-4707.

Olave, M., Avila, L.J., Sites Jr, J.W. and Morando, M., 2018. Hybridization could be a common phenomenon within the highly diverse lizard genus *Liolaemus*. *Journal of evolutionary biology*, *31*(6), pp.893-903.

Oreskes, N., 2004. The scientific consensus on climate change. *Science*, *306*(5702), pp.1686-1686.

Ottenburghs, J., 2020. Ghost introgression: spooky gene flow in the distant past. *Bioessays*, *42*(6), p.2000012.

Ottenburghs, J., 2021. The genic view of hybridization in the Anthropocene. *Evolutionary Applications*, *14*(10), pp.2342-2360.

Ottenburghs, J., van Hooft, P., van Wieren, S.E., Ydenberg, R.C. and Prins, H.H., 2016. Hybridization in geese: a review. *Frontiers in Zoology*, *13*(1), pp.1-9.

Ownbey, M., 1950. Natural hybridization and amphiploidy in the genus *Tragopogon*. *American Journal of Botany*, pp.487-499.

Palma-Otal, M., Moore, W.S., Adams, R.P. and Joswiak, G.R., 1983. Morphological, chemical, and biogeographical analyses of a hybrid zone involving *Juniperus virginiana* and *J. horizontalis* in Wisconsin. *Canadian journal of botany*, *61*(10), pp.2733-2746.

Papadopulos, A.S.T., Price, Z., Devaux, C., Hipperson, H., Smadja, C.M., Hutton, I., Baker, W.J., Butlin, R.K. and Savolainen, V., 2013. A comparative analysis of the mechanisms underlying speciation on Lord Howe Island. *Journal of Evolutionary Biology*, *26*(4), pp.733-745.

Parmesan, C. and Yohe, G., 2003. A globally coherent fingerprint of climate change impacts across natural systems. *Nature*, *421*(6918), pp.37-42.

Pfennig, K.S., Kelly, A.L. and Pierce, A.A., 2016. Hybridization as a facilitator of species range expansion. *Proceedings of the Royal Society B: Biological Sciences*, *283*(1839), p.20161329.

Pineda Torres, Y.M., 2019. The nature of *Espeletia* species. Master's Thesis, Universidad de Los Andes.

Pryer, K.M. and Haufler, C.H., 1993. Isozymic and chromosomal evidence for the allotetraploid origin of *Gymnocarpium dryopteris* (Dryopteridaceae). *Systematic Botany*, pp.150-172.

Pyron, R.A., O'Connell, K.A., Lemmon, E.M., Lemmon, A.R. and Beamer, D.A., 2020. Phylogenomic data reveal reticulation and incongruence among mitochondrial candidate species in Dusky Salamanders (*Desmognathus*). *Molecular Phylogenetics and Evolution*, *146*, p.106751.

Quilodrán, C.S., Montoya-Burgos, J.I. and Currat, M., 2020. Harmonizing hybridization dissonance in conservation. *Communications Biology*, *3*(1), pp.1-10.

Randolph, L.F., 1970. Variation among Tripsacum populations of Mexico and Guatemala. *Brittonia*, *22*(4), pp.305-337.

Rieseberg, L.H., 1991a. Homoploid reticulate evolution in Helianthus (Asteraceae): evidence from ribosomal genes. *American Journal of Botany*, *78*(9), pp.1218-1237.

Rieseberg, L.H., 1991b. Hybridization in rare plants: insights from case studies in Cercocarpus and Helianthus. *Genetics and conservation of rare plants*, pp.171-181.

Rieseberg, L.H., Beckstrom-Sternberg, S.M., Liston, A. and Arias, D.M., 1991. Phylogenetic and systematic inferences from chloroplast DNA and isozyme variation in *Helianthus* sect. *Helianthus* (Asteraceae). *Systematic Botany*, pp.50-76.

Rieseberg, L.H., Carter, R. and Zona, S., 1990. Molecular tests of the hypothesized hybrid origin of two diploid *Helianthus* species (Asteraceae). *Evolution*, *44*(6), pp.1498-1511.

Rieseberg, L.H. and Wendel, J.F., 1993. Introgression and its consequences in plants. *Hybrid zones and the evolutionary process*, *70*, p.109.

Rothfels, C.J., Johnson, A.K., Hovenkamp, P.H., Swofford, D.L., Roskam, H.C., Fraser-Jenkins, C.R., Windham, M.D. and Pryer, K.M., 2015. Natural hybridization between genera that diverged from each other approximately 60 million years ago. *The American Naturalist*, *185*(3), pp.433-442.

Rünk, K., Zobel, M. and Zobel, K., 2012. Biological Flora of the British Isles*: Dryopteris carthusiana*, *D. dilatata* and *D. expansa*. *Journal of ecology*, *100*(4), pp.1039-1063.

Rutherford, S., Rossetto, M., Bragg, J.G., McPherson, H., Benson, D., Bonser, S.P. and Wilson, P.G., 2018. Speciation in the presence of gene flow: population genomics of closely related and diverging *Eucalyptus* species. *Heredity*, *121*(2), pp.126-141.

Rutledge, L.Y., Garroway, C.J., Loveless, K.M. and Patterson, B.R., 2010. Genetic differentiation of eastern wolves in Algonquin Park despite bridging gene flow between coyotes and grey wolves. *Heredity*, *105*(6), pp.520-531.

Saidman, B.O. and Vilardi, J.C., 1987. Analysis of the genetic similarities among seven species of *Prosopis* (Leguminosae: Mimosoideae). *Theoretical and Applied Genetics*, *75*(1), pp.109-116.

Sauleda, R.P. and Hamilton, C.W., 2019. *Tolumnia xpulchella* (Hooker) Rafinesque (Orchidaceae) is Established as the Proper Name for a Jamaican Hybrid Swarm (Syngameon) and a New Species, One of The Parents of the Hybrid swarm, *Tolumnia hamiltonii* Sauleda is described.

Schliewen, U.K. and Klee, B., 2004. Reticulate sympatric speciation in Cameroonian crater lake cichlids. *Frontiers in zoology*, *1*(1), pp.1-12.

Schliewen, U.K., Tautz, D. and Pääbo, S., 1994. Sympatric speciation suggested by monophyly of crater lake cichlids. *Nature*, *368*(6472), pp.629-632.

Schmid, R., Mallory, T.E. and Tucker, J.M., 1968. Biosystematic evidence for hybridization between *Arctostaphylos nissenana* and *A. viscida*. *Brittonia*, *20*(1), pp.34-43.

Schmitt, S., Hérault, B., Ducouret, É., Baranger, A., Tysklind, N., Heuertz, M., Marcon, É., Cazal, S.O. and Derroire, G., 2020. Topography consistently drives intra-and inter-specific leaf trait variation within tree species complexes in a Neotropical forest. *Oikos*, *129*(10), pp.1521-1530.

Schmitt, S., Tysklind, N., Hérault, B. and Heuertz, M., 2021. Topography drives microgeographic adaptations of closely related species in two tropical tree species complexes. *Molecular Ecology*, *30*(20), pp.5080-5093.

Schwarzer, J., Misof, B. and Schliewen, U., 2012. Speciation within genomic networks: a case study based on *Steatocranus* cichlids of the lower Congo rapids. *Journal of Evolutionary Biology*, *25*(1), pp.138-148.

Seehausen, O., 2004. Hybridization and adaptive radiation. *Trends in ecology & evolution*, *19*(4), pp.198-207.

Seehausen, O., Alphen, J.J.V. and Witte, F., 1997. Cichlid fish diversity threatened by eutrophication that curbs sexual selection. *Science*, *277*(5333), pp.1808-1811.

Seybold, S., 2009. Schmeil-Fitschen: Flora von Deutschland und angrenzender Länder. *Quelle and Meyer, Heidelberg, Wiesbaden*.

Shaw, J.T., 2005. *Systematics of the North American Plums (*Prunus *subgenus* Prunus *section* Prunocerasus*; Rosaceae)*. Doctoral dissertation, The University of Tennessee.

Smith, J.M., 1966. Sympatric speciation. *The American Naturalist*, *100*(916), pp.637-650.

Singhal, S., Roddy, A.B., DiVittorio, C., Sanchez-Amaya, A., Henriquez, C.L., Brodersen, C.R., Fehlberg, S. and Zapata, F., 2021. Diversification, disparification and hybridization in the desert shrubs *Encelia*. *New Phytologist*, *230*(3), pp.1228-1241.

Soliani, C., Gallo, L. and Marchelli, P., 2012. Phylogeography of two hybridizing southern beeches (*Nothofagus* spp.) with different adaptive abilities. *Tree Genetics & Genomes*, *8*(4), pp.659-673.

Soltis, P.S. and Soltis, D.E., 2009. The role of hybridization in plant speciation. *Annual review of plant biology*, *60*, pp.561-588.

Stanford, P.K., 1995. For pluralism and against realism about species. *Philosophy of Science*, *62*(1), pp.70-91.

Stecconi, M., Marchelli, P., Puntieri, J., Picca, P. and Gallo, L., 2004. Natural hybridization between a deciduous (*Nothofagus antarctica*, Nothofagaceae) and an evergreen (*N. dombeyi*) forest tree species: evidence from morphological and isoenzymatic traits. *Annals of Botany*, *94*(6), pp.775-786.

Stefani, F., Benzoni, F., Pichon, M., Cancelliere, C. and Galli, P., 2008. A multidisciplinary approach to the definition of species boundaries in branching species of the coral genus *Psammocora* (Cnidaria, Scleractinia). *Zoologica Scripta*, *37*(1), pp.71-91.

Stoehrel, C.P., 2010. *A study of the systematic relationships between members of the* Trillium erectum *complex*. Master's thesis, Western Carolina University.

Suarez-Gonzalez, A., Lexer, C. and Cronk, Q.C., 2018. Adaptive introgression: a plant perspective. *Biology letters*, *14*(3), p.20170688.

Tallmon, D.A., Luikart, G. and Waples, R.S., 2004. The alluring simplicity and complex reality of genetic rescue. *Trends in Ecology & Evolution*, *19*(9), pp.489-496.

Templeton, A.R., 1989. The meaning of species and speciation: a genetic perspective. *Speciation and its Consequences*, pp.3-27.

Todesco, M., Pascual, M.A., Owens, G.L., Ostevik, K.L., Moyers, B.T., Hübner, S., Heredia, S.M., Hahn, M.A., Caseys, C., Bock, D.G. and Rieseberg, L.H., 2016. Hybridization and extinction. *Evolutionary applications*, *9*(7), pp.892-908.

Torales, S.L., Rivarola, M., Pomponio, M.F., Gonzalez, S., Acuña, C.V., Fernández, P., Lauenstein, D.L., Verga, A.R., Hopp, H.E., Paniego, N.B. and Poltri, S.N.M., 2013. De novo assembly and characterization of leaf transcriptome for the development of functional molecular markers of the extremophile multipurpose tree species *Prosopis alba*. *Bmc Genomics*, *14*(1), pp.1-14.

Trucco, F., Tatum, T., Rayburn, A.L. and Tranel, P.J., 2009. Out of the swamp: unidirectional hybridization with weedy species may explain the prevalence of *Amaranthus tuberculatus* as a weed. *New Phytologist*, *184*(4), pp.819-827.

Turesson, G., 1922. The species and the variety as ecological units. *Hereditas*, *3*(1), pp.100-113.

Twyford, A.D. and Ennos, R.A., 2012. Next-generation hybridization and introgression. *Heredity*, *108*(3), pp.179-189.

Urban, A.J., Simelane, D.O., Retief, E., Heystek, F., Williams, H.E. and Madire, L.G., 2011. The invasive '*Lantana camara* L.' hybrid complex (Verbenaceae): a review of research into its identity and biological control in South Africa. *African Entomology*, *19*(1), pp.315-348.

van Oppen, M.J., McDonald, B.J., Willis, B. and Miller, D.J., 2001. The evolutionary history of the coral genus *Acropora* (Scleractinia, Cnidaria) based on a mitochondrial and a nuclear marker: reticulation, incomplete lineage sorting, or morphological convergence?. *Molecular biology and evolution*, *18*(7), pp.1315-1329.

Van Valen, L., 1976. Ecological species, multispecies, and oaks. *Taxon*, pp.233-239.

Voss, E.G. and Reznicek, A.A., 2012. *Field manual of Michigan flora*. University of Michigan Press.

Wagner, N.D., Clements, M.A., Simpson, L. and Nargar, K., 2021. Conservation in the face of hybridisation: genome-wide study to evaluate taxonomic delimitation and conservation status of a threatened orchid species. *Conservation Genetics*, *22*(1), pp.151-168.

Wallace, L.E., 2003. Molecular evidence for allopolyploid speciation and recurrent origins in *Platanthera huronensis* (Orchidaceae). *International Journal of Plant Sciences*, *164*(6), pp.907-916.

Walters, R.S. and Yawney, H.W., 2004. Silvics Manual: Volume 2: Hardwoods. *United States Department of Agriculture, Forest Service, Washington, DC*.

Wang, G., Zhang, X., Herre, E.A., McKey, D., Machado, C.A., Yu, W.B., Cannon, C.H., Arnold, M.L., Pereira, R.A., Ming, R. and Liu, Y.F., 2021. Genomic evidence of prevalent hybridization throughout the evolutionary history of the fig-wasp pollination mutualism. *Nature communications*, *12*(1), pp.1-14.

Weller, S.G., Sakai, A.K. and Wagner, W.L., 2001. Artificial and natural hybridization in *Schiedea* and *Alsinidendron* (Caryophyllaceae: Alsinoideae): the importance of phylogeny, genetic divergence, breeding system, and population size. *Systematic Botany*, *26*(3), pp.571-584.

Wheat, C.W. and Watt, W.B., 2008. A mitochondrial-DNA-based phylogeny for some evolutionary-genetic model species of *Colias* butterflies (Lepidoptera, Pieridae). *Molecular Phylogenetics and Evolution*, *47*(3), pp.893-902.

Wheeler, Q.D. and Meier, R. eds., 2000. *Species concepts and phylogenetic theory: a debate*. Columbia University Press.

Widmer, A., Lexer, C. and Cozzolino, S., 2009. Evolution of reproductive isolation in plants. *Heredity*, *102*(1), pp.31-38.

Williams, J.W., Jackson, S.T. and Kutzbach, J.E., 2007. Projected distributions of novel and disappearing climates by 2100 AD. *Proceedings of the National Academy of Sciences*, *104*(14), pp.5738-5742.

Wilson, P.J., Grewal, S.K., Mallory, F.F. and White, B.N., 2009. Genetic characterization of hybrid wolves across Ontario. *Journal of Heredity*, *100*(suppl_1), pp.S80-S89.

Wood, C.E., 1955. Evidence for the hybrid origin of *Drosera anglica*. *Rhodora*, *57*(676), pp.105-130.

Woodruff, D.S., 1981. Towards a genodynamics of hybrid zones: studies of Australian frogs and West Indian land snails. *Evolution and speciation*, pp.171-197.

Wu, G.A., Terol, J., Ibanez, V., López-García, A., Pérez-Román, E., Borredá, C., Domingo, C., Tadeo, F.R., Carbonell-Caballero, J., Alonso, R. and Curk, F., 2018. Genomics of the origin and evolution of Citrus. *Nature*, *554*(7692), pp.311-316.

Wyatt, R. and Antonovics, J., 1981. Butterflyweed re-revisited: spatial and temporal patterns of leaf shape variation in *Asclepias tuberosa*. *Evolution*, pp.529-542.

Yang, R., Folk, R., Zhang, N. and Gong, X., 2019. Homoploid hybridization of plants in the Hengduan mountains region. *Ecology and Evolution*, *9*(14), pp.8399-8410.

Young, N.D., 1998. Pacific Coast Iris species delimitation using three species definitions: biological, phylogenetic and genealogical. *Biological Journal of the Linnean Society*, *63*(1), pp.99-120.

Zhang, N.N., Yu, J.J., Wang, Y.H. and Gong, X., 2018. Molecular evidence for asymmetric hybridization in three closely related sympatric species. *AoB Plants*, *10*(1), p.ply011.

CHAPTER II

Patterns of hybridization and cryptic introgression among one and four needle pinyon pines

Ryan Buck, Sandra Hyasat, Alice Hossfeld, Lluvia Flores-Rentería*
Department of Biology, San Diego State University, San Diego, CA, United States

Abstract

Pinyon pine hybridization is widely acknowledged but the frequency of and contributors to such interspecific mating remain largely unstudied. *Pinus quadrifolia* has three to four needles per fascicle, suggesting it is a result of hybridization between the five-needled *P. juarezensis* and the single-needled *P. monophylla*. In this study we address the taxonomic validity of *P. juarezensis*, the hybrid origin of *P. quadrifolia*, and the presence of hybridization and intermediate morphology as a result of interspecific hybridization in this complex. We addressed these questions by combining a genomic and morphological approach. We generated 1,868 SNPs to detect genetic clusters using PCoA, DAPC, fastSTRUCTURE, and ADMIXTURE analyses and performed a morphological analysis of the leaves. We found that the five-needled pinyons did not differ genetically from the four-needled *P. quadrifolia*, reducing *P. juarezensis*' status to *P. quadrifolia*. We also found no evidence that *P. quadrifolia* is of hybrid origin from *P. juarezensis* x *P. monophylla* but is instead a genetically distinct species with natural needle number variation that has yet to be explained. Hybridization does occur in this complex, but mostly between *P. quadrifolia* and *P. californiarum*, and less commonly between *P. quadrifolia* and *P. monophylla*. Interestingly, some hybrid derivatives were detected between both single-needled taxa, *P. monophylla* and *P. californiarum*, a hybrid combination that has not yet been proposed. Hybrids have intermediate morphology when they have similar genetic contributions from both parental species; however, when one parent contributes more, hybrid derivatives resemble the parent with higher genetic contribution, resulting in cryptic introgression. Our detailed sampling across the distribution of this complex allowed us to describe the patterns of hybridization among these taxa, resolved an ancient taxonomic conflict, and provided insights on the challenges of exclusively using morphological traits when identifying these taxa with cryptic hybridization and variable morphology.

Keywords: Cryptic introgression; genomics; hybridization; needle number; pinyon pines; resin canals

Introduction

Interspecies hybridization is relatively common across plant taxa and is thought to be the cause of several major speciation events (Soltis and Soltis, 2009; Abbot *et al.*, 2013). Hybridization tends to occur when species lack strong reproductive isolating mechanisms (Bigelow, 1965), which normally act as barriers to reproduction and help in the process of speciation (Rieseberg and Willis, 2007). If species have loose genetic barriers (compatible zygotes) and at least partially overlap spatially and temporally, then hybrids are likely to form (Zhao *et al.*, 2014). Intermediate forms can be detected in recent hybridization events (Zavarin *el al.*, 1980, Delaporte *et al.*, 2001); however, when backcrossing occurs in a hybrid system, morphological traits are present either as a continuum between the two parental species (Holman *et al.*, 2003), or more representative of one of the parents resulting in cryptic introgression (Pfenninger *et al.*, 2002; Jasińska *et al.*, 2010; Neri *et al.*, 2017). The use of morphological traits to determine hybridization events can be useful when divergent features exist between parental species, but when morphological limits are not well defined in taxonomically challenging groups, it becomes difficult to distinguish intermediacy from interpopulation variation (Wei *et al.*, 2015). For example, in some systems, hybridization leads to a mosaic of forms or even to extreme or novel characters (Rieseberg *et al.*, 1993).

Hybridization in pines has been widely acknowledged (Critchfield, 1975, 1986; Willyard *et al.*, 2009; Menon *et al.*, 2018, 2020). The pinyon pine complex is an excellent system to study hybridization due to its lack of: strong reproductive isolating mechanisms (Lanner, 1974a), taxonomic congruence (Gernandt *et al.*, 2001, 2003), and conclusive genetic studies (Montes *et al.*, 2019). Most pine species lack interspecific incompatibility mechanisms (Critchfield, 1975) and have wind-mediated, long-distance pollen dispersals (Williams, 2010) that can facilitate gene flow between allopatric groups (Wright, 1952). Most notably, the Parry pinyon pine, *Pinus quadrifolia* Parl. ex Sudworth, has been proposed to have a hybrid origin based on morphological features and geography (Lanner, 1974a). *Pinus quadrifolia* commonly has four needles per fascicle, as its epithet suggests, and occurs from Riverside County, California to northern Baja California, Mexico. Based on intermediate morphology, Lanner (1974a) hypothesized its origin as the result of interspecific hybridization between two species with similar pollen dispersal times (Malusa, 1992; Farjon and Styles, 1997) and partially overlapping distributions (Fig. 1), *P.*

monophylla Torrey & Fremont and *P. juarezensis* Lanner. At the time of his discovery, needle number was the main morphological trait used in classifying pinyon pines, so he reasoned that individuals with one needle per fascicle were one species (*P. monophylla*), individuals with five needles were another (*P. juarezensis*), and the three- and four-needled individuals (*P. quadrifolia*) were intermediate hybrids of the former two.

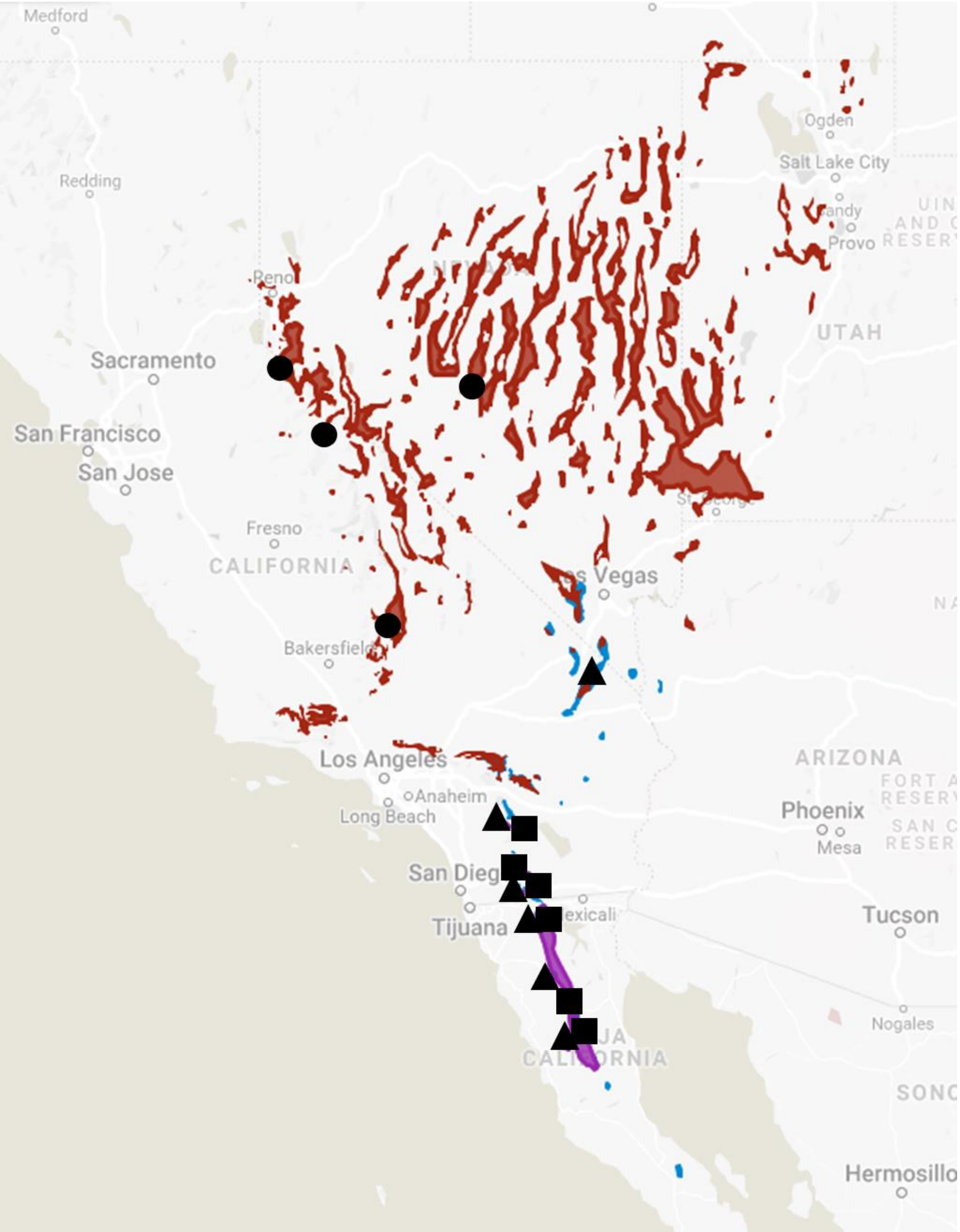


Figure 1. Putative distribution map of *P. monophylla* (red), *P. californiarum* (blue), and *P. quadrifolia+P. juarezensis* (purple) (USGS, 1999; Cole *et al.*, 2008). Sampling locations are represented by the overlaid shapes (circle = *P. monophylla*, triangle = *P. californiarum*, square = *P. quadrifolia+P. juarezensis*).

Several attempts have been made to resolve the hybrid nature *of P. quadrifolia,* however they have provided inconclusive results. The idea of hybridization was explored using additional morphological features other than needle number, but these studies focused on the potential hybridization between *P. monophylla* and *P. quadrifolia* without considering *P. juarezensis* (Zavarin *et al.*, 1980; Snajberk *et al.*, 1982). These studies also explored chemical profiles, such as mono and sesquiterpenoids that could be used for species identification between *P. quadrifolia and P. monophylla* (Zavarin *et al.*, 1980), however an intermediate chemical profile was detected in individuals in sympatry suggesting hybridization. Additionally, the only genetic study including the three putative taxa (*P. juarezensis*, *P. quadrifolia*, and *P. monophylla*) was done with ITS (Gernandt *et al.*, 2001), which has been proven to have paralogy in pines and other plants (Grimm and Denk, 2007; Xiao *et al.*, 2010; Flores-Rentería *et al.*, 2013, 2017). When using this nuclear marker, species were not monophyletic, so the authors then suggested a potential hybridization event between *P. monophylla* and *P. juarezensis* or incomplete lineage sorting as the cause. Therefore, the question of whether *P. quadrifolia* has a hybrid origin has remained unanswered for more than 40 years.

Since Lanner's publication (1974a) suggesting the hybrid origin of *P. quadrifolia* from the parental species *P. monophylla* and *P. juarezensis*, several taxonomic studies have been conducted, some of them suggest the splitting of *P. monophylla* into three different taxa based on morphology and niche variability: *P. monophylla*, *P. californiarum*, and *P. fallax* (Little, 1968; Bailey, 1987; Cole *et al.*, 2008). One of these suggested taxa, *P. californiarum* D.K. Bailey, has more resin canals and occupies drier habitats in higher elevations than *P. monophylla*. In a recent study, hybridization between *P. californiarum* and *P. quadrifolia* was explored using a Hyb-seq method, yet no gene flow was detected (Montes *et al.*, 2019). One reason could be the lack of adequate sample size due to the taxonomic focus of the paper which aimed to capture as many species as possible using only a few individuals. Montes *et al.* (2019) stated themselves that increased sampling would be required to draw accurate conclusions about admixture events in this system. In our study, we will be testing whether *P. juarezensis* hybridizes with *P. monophylla* or with the recently proposed taxa *P. californiarum.* In contrast, *P. juarezensis* has not been recognized as an independent species by some authors, but rather as a natural morphological variation of *P. quadrifolia*, able to produce three to five needles per fascicle. For example, Farjon and Styles (1997) challenged Lanner's conclusion, pointing to the extreme variation in needle number on a single tree and the year to year variation on trees reported by Lanner (1974a) himself. We will be testing the taxonomic validity of *P. juarezensis* as an independent species.

With advancements in next generation sequencing (NGS), the ability to detect reticulated evolution through hybridization has drastically improved (Pritchard *et al.*, 2000; Twyford and Ennos, 2012; Melville *et al.*, 2017). We are now able to revisit these hybridization studies with a finer scale, allowing us to examine the purported hybrid origin of *P. quadrifolia* and address the taxonomic standing of *P. juarezensis*. The goals of this study are (1) to determine the taxonomic validity of *P. juarezensis* as a species, (2) to assess the hybrid origin of *P. quadrifolia*, and (3) to test for hybridization and intermediate morphology among species in this complex. We have included samples from the taxa of interest using nuclear genomic data as well as morphological data to assess these questions.

Materials and Methods

Sampling

Samples were taken from 13 locations across the Southwestern US and Baja California (Table 1, Fig. 1), randomly sampling ten trees per putative species per site, at least 30 meters apart from each other. The distribution of this complex extends from southern Idaho, U.S. to southern Baja California, Mexico, with a majority of *P. monophylla* occurring in western Utah, Nevada and central California, *P. californiarum* occurring in southern Nevada, southern California, and down into Baja California, and *P. juarezensis* (five needle) and *P. quadrifolia* (three to four needle) occurring in southern California and Baja California. These taxa are characterized by inhabiting small disjunct areas in southern California and Baja California. Ten centimetres of branch tips were cut from each tree, representing two to three years of growth and averaging approximately 50 fascicles per tree. Most collections were made from late 2017 to early 2019, with the exception of a few collections made previously (2011-2013). We collected a larger number of individuals in San Jacinto because Lanner (1974a) proposed this area as one of the primary hybrid zones between *P. juarezensis* and *P. monophylla.*

Table 1. Sample locations with population coordinates and number of trees sampled per population. California (CA), Nevada (NV), Baja California (Baja CA).

Location	Latitude	Longitude	# Trees genotyped	Needle Type
Markleeville, CA	38.69894	-119.77082	10	*P. monophylla*
Manhattan, NV	38.54101	-117.05981	6	*P. monophylla*
Mono Lake, CA	37.92145	-119.06433	7	*P. monophylla*
Pine Mt, CA	36.02982	-118.15251	9	*P. monophylla*
Mojave Desert, CA	35.17397	-115.40707	13	*P. californiarum* Hybrids
San Jacinto, CA	33.55887	-116.60994	37	*P. quadrifolia* *P. californiarum* *P. juarezensis* Hybrids
Laguna Mts, CA	32.87554	-116.41017	9	*P. quadrifolia* *P. juarezensis*
Jacumba, CA	32.63317	-116.09395	15	*P. quadrifolia* *P. californiarum* *P. juarezensis* Hybrids
La Rumorosa, Baja CA	32.521544	-116.04122	21	*P. quadrifolia* *P. californiarum* *P. juarezensis* Hybrids
San Salvador, Baja CA	31.742415	-115.97815	8	*P. quadrifolia* *P. juarezensis*
Lazaro Cardenas, Baja CA	31.257595	-115.59940	11	*P. californiarum*
Northern San Pedro Martir, Baja CA	31.034611	-115.46477	18	*P. quadrifolia* *P. californiarum* *P. juarezensis*
Southern San Pedro Martir, Baja CA	31.02131	-115.51488	10	*P. quadrifolia* *P. juarezensis*

Genomic clustering analyses

In order to examine the taxonomic validity of *P. juarezensis* as a species, the hybrid origin of *P. quadrifolia*, and hybridization among species of this complex we first identified genetic clustering and admixed individuals. Nuclear DNA was extracted using Doyle and Doyle's (1987) CTAB 2% protocol, quantified, and sent to Diversity Arrays Technology (DArT), who produced a reduced library and sequenced on an Illumina Hiseq2500 system (DArTseq). DArTseq™ represents a combination of DArT complexity reduction methods and NGS platforms (Jaccoud *et al.*, 2001). Genome reduction is achieved by a combination of endonucleases that specifically target low-copy DNA areas, rather than repetitive DNA fragments (Wenzl *et al.*, 2004). This allows for detection of a high number of informative SNPs across the genome. The result is a genomic 'representation', comprising both constant and polymorphic fragments across individuals. NGS of these 'representations' reveals the sequence (approx. 70 bp) of an informative DNA fragment and each individual's state compared with all others, namely (i) homozygosity with reference allele, (ii) homozygosity with alternate allele, or (iii) heterozygosity, comprising both a reference and an alternate SNP allele. The technology was optimized for our taxa using combinations of enzymes (PstI/HpaII, PstI/SphI, SbfI/HpaII, SbfI/MseI) to select the most appropriate complexity reduction method, both in terms of the size of the representation and the fraction of a genome selected for assays. DArTseq™ has been successfully applied in genomic studies exploring species boundaries and hybridization in plants and animals (Cruz *et al.*, 2013; Melville *et al.*, 2017, Rutherford *et al.*, 2018). This technique enables genome-wide studies of non-model organisms, those for which there is limited genomic information. For our analysis we selected *P. lambertiana* as reference genome (GCA_001447015.2) due to its relatively close evolutionary relationship to the pinyon subsection *Cembroides*. A low-density DArTseq assay resulted in 18,518 SNPs. These SNPs have an index generated by reproducing the data independently, which is “the proportion of technical replicate assay pairs for which the marker score is consistent” (Gruber *et al.*, 2019). We filtered out all loci with reproducibility lower than 100%, missing data lower than 15%, all monomorphs, all loci departing from Hardy-Weinberg equilibrium, and all but one locus where there was more than one locus per sequence tag, resulting in a final data set of 1,868 loci. Input data were the metadata provided by DArTseq, saved as an xlsx file: ‘0’ (homozygosity with reference allele); ‘1’ (homozygosity with alternate allele); ‘2’ (heterozygote) and ‘-’, fragment missing in representation—double null (absence of fragment with SNP in genomic representation). The processed marker data were reformatted into appropriate file types for downstream analyses using the R program dartR (Gruber *et al.*, 2018).

The 1,868 loci were included in four complementary methods of genetic clustering: PCoA, DAPC, fastSTRUCTURE, and ADMIXTURE, which also allowed us to detect admixed individuals between different genetic clusters. We then evaluated whether *P. juarezensis* (five needled individuals) forms its own genetic cluster and whether *P. quadrifolia* (three and four needled individuals) is formed by admixed individuals from the parental species *P. monophylla* and *P. juarezensis* or whether there are other pairs of species hybridizing in this complex.

PCoA - A Principal Coordinates Analysis (PCoA), which considers differences in allele frequencies between individuals, was performed in dartR (gl.pcoa.plot) to examine genetic distance among populations (Gruber *et al.*, 2019).

DAPC - Population clustering was determined using a Discriminant Analysis of Principal Components (DAPC) (Jombart *et al.*, 2010) in R with the adegenet package (Jombart, 2008). DAPC uses a model-free k-means clustering algorithm on a PCA-transformed dataset, which maximizes variation between groups while reducing the number of variables and computation time needed to identify existing genetic clusters (Jombart and Collins, 2015). Differing cluster solutions are compared using Bayesian Information Criterion (BIC), with the lowest BIC score corresponding to the optimal cluster solution. The xvalDapc command, which uses a stratified random sampling to ensure one member of each population is represented in both a training set consisting of 90% of the data and a validation set consisting of the remaining 10% of the data, was used as a cross-validation method to determine the appropriate number of principal components to retain (Jombart and Collins, 2015).

fastSTRUCTURE - A Bayesian analysis of population clustering was performed in the software fastSTRUCTURE (Raj *et al.*, 2014) using the logistic prior and five cross-validations. Output files include the mean Q value for each individual, defining the mean probability to belong to any one of the populations K1 to Kx. Model complexity (K) was selected using the chooseK command built into fastSTRUCTURE. The resulting Q mean bar plots were visualized using the online application pophelper (Francis, 2017). Results were confirmed using ADMIXTURE (Alexander *et al.*, 2009) and a sub-structuring method, wherein samples with the same population identity (K) were reanalysed in fastSTRUCTURE using the logistic prior and five cross-validations to observe lower hierarchical structuring (Raj *et al.*, 2014).

Hybrid Generation Identification

NewHybrids – In order to determine an individual's hybrid category, e.g. early F_1 or advanced generation (F_2 and backcross), we used a Bayesian model-based clustering method in the software NewHybrids 1.0 (Anderson and Thompson, 2002). The program uses Markov chain Monte Carlo simulations to compute the posterior probability of an individual belonging to predefined ancestry categories, including pure, F_1, F_2, or backcrossed (Table 2). The program compares two parental genotypes at a time, so we created three data sets to represent each pairwise species cross by removing individuals with ancestry of the third genotype (e.g. for *P. monophylla* x *P. quadrifolia* analysis we removed pure *P. californiarum* individuals and individuals with *P. californiarum* ancestry). Runs were initiated at different random starting points using the Jeffrey's prior for both theta and pi with a burn-in of 10,000 and 100,000 sweeps (Couch *et al.*, 2016).

Table 2. Twelve genotype frequency categories input into NewHybrids analyses based on Chhatre *et al.*, 2018.

Genotype Category	Cross Type	Expected ancestry proportions			
		AA	Aa	aA	aa
Pure	Species 1	1	0	0	0
Pure	Species 2	0	0	0	1
F_1	Species 1 x Species 2	0	0.5	0.5	0
F_2	F_1 x F_1	0.25	0.25	0.25	0.25
F_1 Backcross 1	F_1 x Species 1	0.5	0.25	0.25	0
F_1 Backcross 2	F_1 x Species 2	0	0.25	0.25	0.5
F_2 Backcross 1	F_2 x Species 1	0.5	0.125	0.125	0.25
F_2 Backcross 2	F_2 x Species 2	0.25	0.125	0.125	0.5
1 Backcross x F_1 Backcross 1	Species 1 x (F_1 x Species 1)	0.75	0.125	0.125	0
2 Backcross x F_1 Backcross 2	Species 2 x (F_1 x Species 2)	0	0.125	0.125	0.75
1 Backcross x F_2 Backcross 1	Species 1 x (F_2 x Species 1)	0.625	0.125	0.125	0.125
2 Backcross x F_2 Backcross 2	Species 2 x (F_2 x Species 2)	0.125	0.125	0.125	0.625

Morphology

In order to determine what morphological features are consistent with the genetic clusters and to detect intermediate morphology associated with hybridization, we analysed a range of leaf morphological traits that have been used in most taxonomic studies of pinyon pines, including needle number, number of resin canals, and number of stomata rows (Little, 1968; Bailey, 1987; Lanner and Phillips, 1992; Malusa, 1992; Christensen *et al.*, 1995; Cole *et al.*, 2008; Flores-Rentería *et al.*, 2013). All sample shoots were kept in the freezer at -20°C and analysed individually at room temperature. All fascicles found in the 10 cm tip branches were visually examined for needle number. Most trees retain their needles, but some drop a few needles. To avoid any bias on needle number estimates, close examination of branch tips was done under a stereomicroscope to inspect fascicle scars, allowing us to detect dropped needles and accurately estimate the range of needle number per branch. Then we identified whether trees have the same number of needles across the branch (uniform number) or if they have a range of needle number (non-uniform). One fascicle was measured from branch tips with uniform needle per fascicle numbers. For branch tips with non-uniform needle per fascicle numbers, one fascicle of each needle number variant was measured. Clear nail polish was used when necessary to enhance the visualization of stomatal rows. A stomatal position index was created ranging from 1 to 5, where 1 means ventral position only, 2 is full ventral and partial dorsal, 3 is complete on both sides, 4 is full dorsal and partial ventral, and 5 is dorsal only. In addition to stomatal position, cross-sectional area and thickness of hypodermal layers were determined (Flores-Rentería *et al.*, 2013). Every needle sample was cross sectioned a few times with a razor blade in the midsection and rehydrated with water before measuring the cross-sectional area, dermal thickness, and number of dermal layers. Measurements for cross area sections were traced in micrometres and thickness of hypodermal layers was recorded as the average of four measurements around the needle. Variation in the number of resin canals in needles from different fascicles of the same branch was initially observed, thus two resin canal measurements were taken; however, the difference between these two measurements proved to be non-significant so only the first measurement was included in further analyses. All measurements were done using an ultra high-resolution Nikon SMZ25 stereoscopic microscope zoom 0.5-1.6X and NIS Elements software. We used Microsoft Excel to plot the means and standard errors of needle number by resin canals and stomatal rows by resin canals per population with and without genetic hybrids, following Cole *et al*. (2008). This method allows us to examine differences among species while showing the intraspecific variation present among populations. It also allows us to examine the morphology of hybrids and compare them to general species characteristics.

Results

Genetic clustering analyses

The PCoA analysis (Fig. 2) shows individuals separating into three groups with PCoA axis one representing 9.6% of variation in genetic distance and PCoA axis two representing 7.9%. While those two axes represent the two most informative axes, the subsequent axes incrementally represent less variation, with PCoA axis three representing 1.5%, four representing 1.3%, and so on. Each group represents individuals that have the morphology of *P. monophylla, P. quadrifolia* / *P. juarezensis,* and *P. californiarum*, respectively. Individuals with allele frequencies in-between *P. monophylla* and *P. quadrifolia*, *P. monophylla* and *P. californiarum*, and *P. quadrifolia* and *P. californiarum* can be seen falling out of the three main groups, suggesting possible admixture among the three taxa. It is important to note that a PCoA does not present statistical clusters, but instead functions as a tool to visualize differences in allele frequency and therefore must be interpreted with caution. Individuals on the PCoA graph may group together due to similarities in allele frequencies, but a PCoA does not provide a group assessment as in K-means clustering methods like DAPC and fastSTRUCTURE.

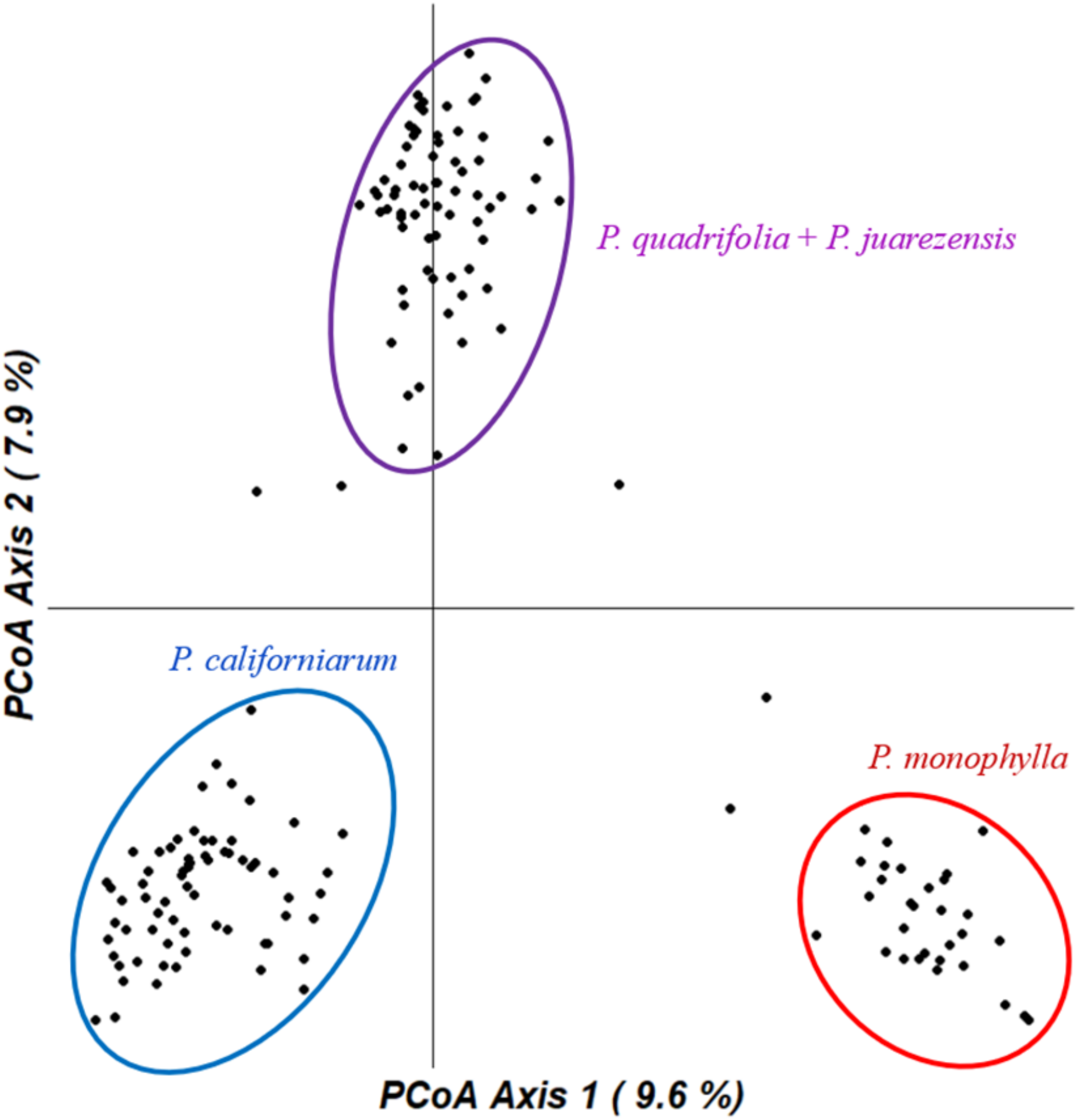


Figure 2. Principal Coordinates Analysis showing the variation in genetic distance with three main groups circled by genetic identity (red = *P. monophylla*, blue = *P. californiarum*, purple = *P. quadrifolia*+*P. juarezensis*). Each black dot represents an individual, dots outside of circles represent potential hybrids and hybrid derivatives.

DAPC analysis (Fig. 3) of the SNP dataset resulted in a value of K=3 by K-means clustering, retaining twenty principal components and two discriminant functions. All sample populations of *P. monophylla* clustered together as one group (colored in red). All sample populations of *P. quadrifolia* and purported *P. juarezensis* clustered together as one group (colored in purple). All sample populations of *P. californiarum* (colored in blue) clustered together as one group.

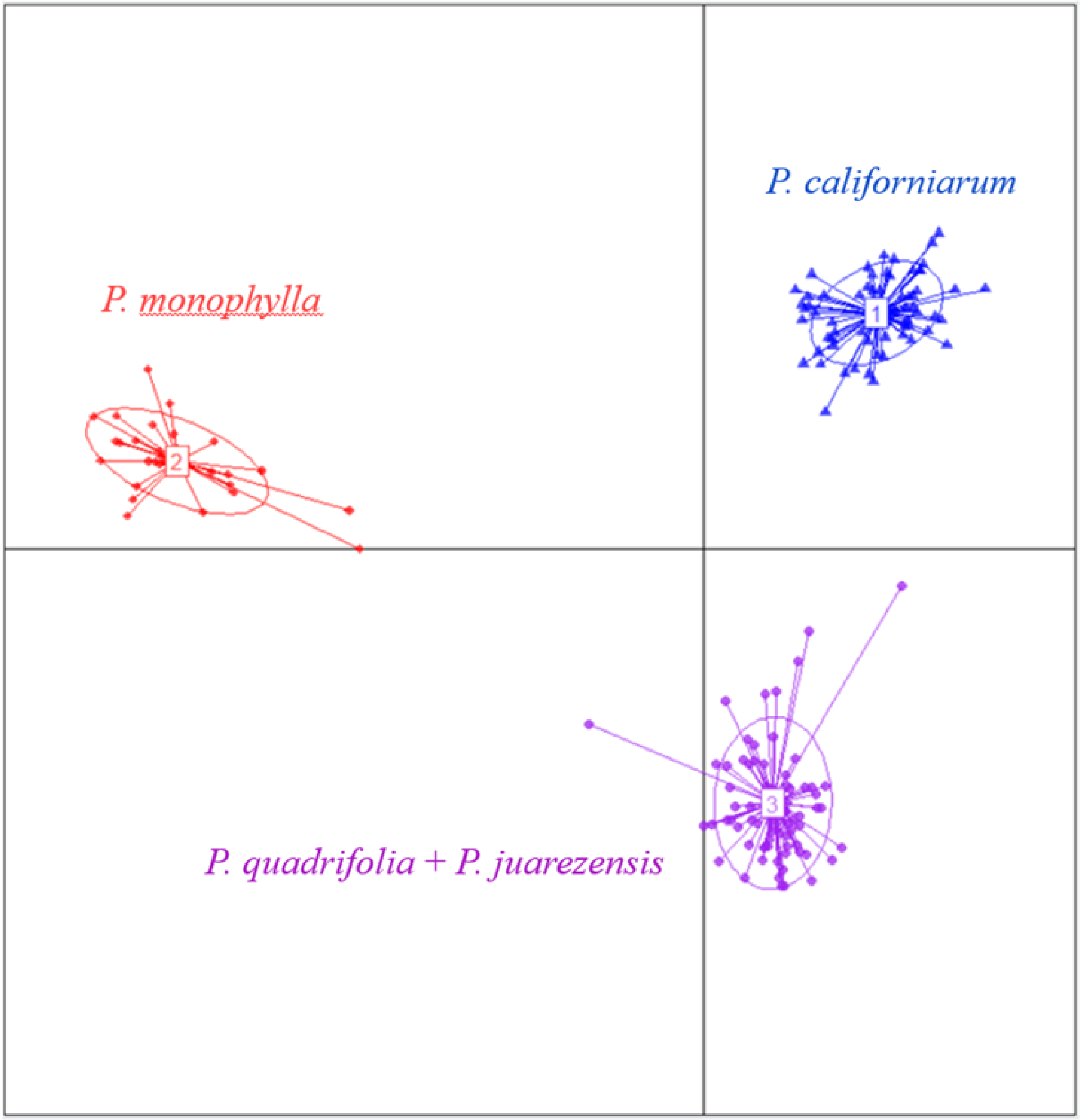


Figure 3. Discriminant Analysis of Principal Components showing three distinct genetic clusters colored by genetic identity (red = *P. monophylla*, blue = *P. californiarum*, purple = *P. quadrifolia*+*P. juarezensis*). Each dot represents an individual connected to the genetic cluster's centroid by a line, inertia ellipses indicate their assignment to one of the three genetic clusters inferred by DAPC.

Our analyses of fastSTRUCTURE show three distinct genetic clusters using the chooseK method and visualized in pophelper (Fig. 4; Suppl. Fig. 1), supporting the DAPC and PCoA results. Again, *P. monophylla* forms one cluster (colored in red), *P. quadrifolia* and purported *P. juarezensis* another cluster (colored in purple), and *P. californiarum* the third cluster (colored in blue). These results show no genetic distinction between individuals with morphological characteristics of *P. quadrifolia* (three to four needle) and *P. juarezensis* (five needle), this was further confirmed by the substructuring analysis of the cluster containing only the putative *P. juarezensis* and *P. quadrifolia* (Suppl. Fig. 2). Moreover, this analysis did not support the hypothesized hybrid origin of *P. quadrifolia*. However, these results do show extensive admixture among the three taxa with several examples of *P. monophylla* x *P. quadrifolia* (as in San Jacinto), *P. monophylla* x *P. californiarum* (as in Mojave), and *P. quadrifolia* x *P. californiarum* (as in Rumorosa, Mojave, San Jacinto, and Jacumba) admixture events. These admixture events are evident in the fastSTRUCTURE plot (Fig. 4) where individuals have more than one genetic cluster assigned to them. ADMIXTURE analyses supported three genetic clusters with the lowest cross-validation error at K=3 (Suppl. Fig. 3). Sub-structuring using fastSTRUCTURE's logistic prior and ten cross-validations found no lower hierarchical structuring below K=3.

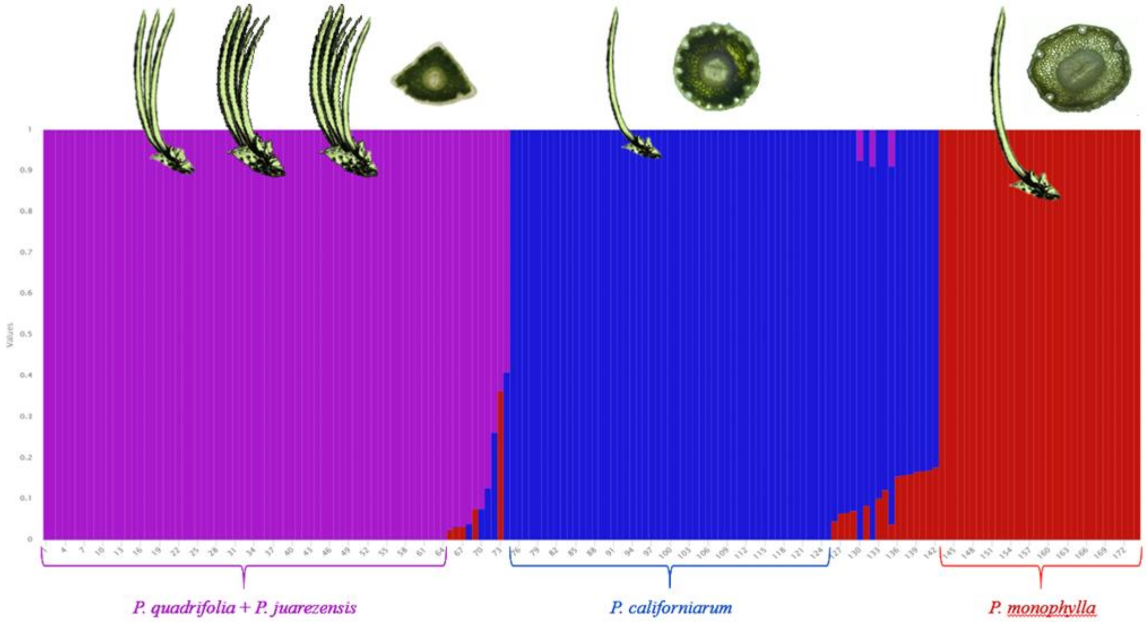


Figure 4. fastSTRUCTURE plot showing three genetic clusters (K=3) colored by genetic identity (red = *P. monophylla*, blue = *P. californiarum*, purple = *P. quadrifolia*+*P. juarezensis*). Each line on the x-axis represents an individual and the probability of those individuals belonging to a certain genetic cluster is represented by the proportion of color on the y-axis. Lines with multiple colors represent individuals with admixture from multiple genetic clusters. Cartoons showing the number of needles per fascicle and cross-sectional pictures showing the resin canals are overlaid on their corresponding genetic groups.

Hybrid Generation Identification

Interestingly, the NewHybrids analyses (Table 3) show that most individuals initially identified as having hybrid ancestry in fastSTRUCTURE end up grouping as advanced generation backcrosses, suggesting F_1 hybrids are able to reproduce further. Several individuals had multiple category assignments with probability scores lower than 0.98, suggesting uncertainty in assignment (Anderson and Thompson, 2002). These pure and mixed assignment individuals are possibly backcrosses of more advanced generation than our genotype categories can detect. Individuals that appeared near a Q value of 0.4 on the fastSTRUCTURE plot were at least third or fourth generation hybrid derivatives, with the most recent hybrid ancestry generation of F_2 x *P. quadrifolia* and no F_1 or F_2 individuals detected. A majority of the individuals were identified as fourth generation backcrosses, having a genotype class of Backcross x (F_2 x Backcross).

Table 3. Levels of ancestry determined by NewHybrids for the apparent hybrid derivatives indicated in fastSTRUCTURE analyses.

NewHybrids assignment	Contributing species	# indiv.
pure *P. quadrifolia*	*P. quadrifolia*	61
pure *P. monophylla*	*P. monophylla*	32
pure *P. californiarum*	*P. californiarum*	53
P. quadrifolia x (F_2 x *P. quadrifolia*)	*P. monophylla, P. quadrifolia*	7
P. quadrifolia x (F_2 x *P. quadrifolia*)	*P. californiarum, P. quadrifolia*	4
F_2 x *P. quadrifolia*	*P. monophylla, P. quadrifolia*	1
F_2 x *P. quadrifolia*	*P. californiarum, P. quadrifolia*	1
P. californiarum x (F_2 x *P. californiarum*)	*P. monophylla, P. californiarum*	14

Morphological analysis

In order to properly assign morphological traits to species, characteristics were assigned to the genetic groups resulting from the genetic structuring analyses, particularly fastSTRUCTURE and NewHybrids. Individuals assigned exclusively to one genetic cluster were determined to be pure, while individuals of mixed assignment were run through NewHybrids and distinguished as hybrids (F_1) and hybrid derivatives (F_2, backcrosses, or advanced generation backcrosses). All morphological results are listed in Table 4 and distinguishing traits are plotted in Fig. 5. Morphological results also supported three groups with *P. monophylla* and *P. californiarum* having one needle per fascicle on average but separated by the larger number of resin canals in *P. californiarum.* In contrast, *P. quadrifolia+P. juarezensis* formed one group, having the highest number of needles but lower number of stomata rows, and stomata just on the ventral surface. Intermediate forms were detected when hybrids and hybrid derivatives were plotted (Fig. 5B and 5C), however, in some cases some individuals more closely resembled one of the parental species.

Table 4. Morphological characteristics based on genetic identity. Note that *P. juarezensis* is treated as synonym of *P. quadrifolia* because no genetic differences were found between them. Letters in the hybrid cross rows represent the first initial of the species name (M = *P. monophylla*, Q = *P. quadrifolia*, C = *P. californiarum*).

Genetic identity	Needle per fascicle		Resin canals		Stomatal rows	Stomatal position	Area (mm^2)	Number of dermal layers	Thickness of dermal layers (µm)
	Mean	Range	Mean	Range					
P. monophylla	1.06 ± 0.24	1-2	2.82 ± 1.08	1-6	24.48 ± 4.09	complete	$1.9 \times 10^3 \pm 7.0 \times 10^2$	2.36 ± 0.65	59.08 ± 13.86
P. quadrifolia	4.19 ± 0.77	2-5	1.91 ± 0.78	0-7	7.69 ± 1.80	ventral only	$7.9 \times 10^2 \pm 2.5 \times 10^3$	2.59 ± 0.55	36.76 ± 37.57
P. californiarum	1.11 ± 0.31	1-2	12.30 ± 2.96	6-23	19.81 ± 4.02	complete	$1.1 \times 10^3 \pm 2.1 \times 10^3$	2.49 ± 0.50	37.43 ± 25.25
M × Q	3.71 ± 1.03	2-5	2.04 ± 0.26	2-3	8.71 ± 1.98	ventral only or complete	$5.6 \times 10^2 \pm 3.6 \times 10^2$	2.14± 0.64	40.37 ± 13.90
Q × C	2.83 ± 1.40	1-5	4.22 ± 4.60	1-17	11.58 ± 3.90	ventral only or complete	$5.9 \times 10^2 \pm 4.0 \times 10^2$	2.66 ± 0.47	38.49 ± 18.42
C × M	1.13 ± 0.33	1-2	8.94 ± 2.65	1-12	16.92 ± 2.56	complete	$2.0 \times 10^3 \pm 2.4 \times 10^3$	2.97 ± 0.62	51.38 ± 12.86
M × Q × C	1.50 ± 0.50	1-2	8.50 ± 0.50	8-9	14.00 ± 0.00	complete	$8.8 \times 10^2 \pm 9.5 \times 10^1$	2.00 ± 0.00	51.38 ± 0.84

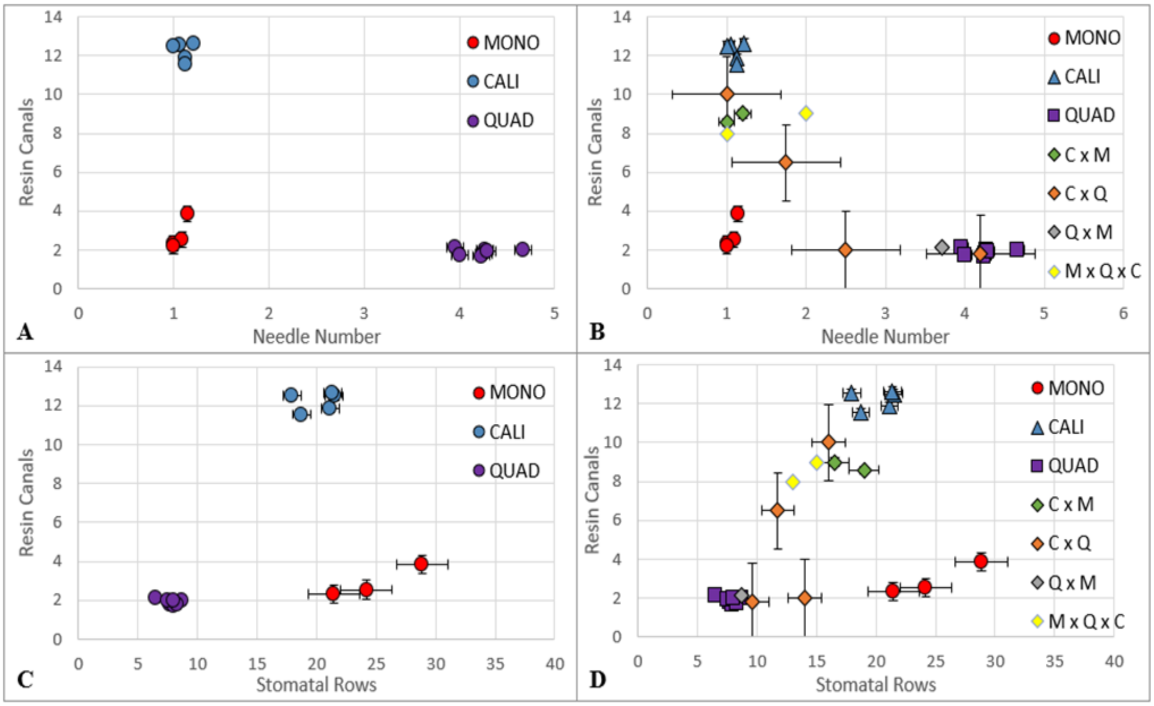


Figure 5. Morphological analyses showing in panels A & B) needle number by resin canals and C & D) stomatal rows by resin canals, with A & C) showing pure species and B & D) showing pure species and hybrid derivatives. Each dot represents the mean of a population with standard error bars, each color corresponds to a species (red circle = *P. monophylla*, blue triangle = *P. californiarum*, purple square = *P. quadrifolia*, diamonds = hybrid derivatives).

Discussion

Genetic clustering analyses show three groups: *P. monophylla, P. californiarum*, and *P. quadrifolia*. Five-needled individuals are not genetically distinguishable from three- and four-needled ones, contrary to Lanner's original hypothesis. Individuals with hybrid ancestry were detected, however the hypothesized hybrid origin of *P. quadrifolia* from the species tested was not supported. Morphological analyses showed a clear distinction between *P. californiarum* and *P. monophylla* in the number of resin canals, while both *P. californiarum* and *P. monophylla* were distinguishable from *P. quadrifolia* in the number of needles per fascicle. *Pinus quadrifolia* showed variation in the number of needles per fascicle throughout its range, with most individuals having a variable number of needles on the same branch.

Lanner's hypotheses on the existence of *P. juarezensis* and the hybrid origin *P. quadrifolia* proved to be incorrect. While *P. quadrifolia* does have variation in its needle number, this variation is not explained by interspecific hybridization. *Pinus juarezensis* is not a valid taxon, but rather a synonym of *P. quadrifolia,* as five-needled individuals genetically cluster with the three- and four-needled ones. Hybridization was detected between *P. quadrifolia* and *P. monophylla*, as well as *P. quadrifolia* and *P. californiarum*. Interestingly, a new admixture combination was detected between *P. californiarum* and *P. monophylla*. The morphological traits of recent hybrid derivatives appeared intermediate between the parental species, yet as apparent backcrossing occurred, the advanced generation hybrid derivatives appeared morphologically indistinguishable from their main genetically contributing parent.

Genetic clustering analyses do not support P. juarezensis *as a valid taxon*

Previous studies have either failed to include *P. juarezensis* in their analyses or found inconclusive results regarding its taxonomic validity (Zavarin *et al.*, 1980; Snajberk *et al.*, 1982; Gernandt *et al.*, 2001; Montes *et al.*, 2019). In our study, we included several populations with the five-needled pinyon individuals, including some from Lanner's original discovery (Lanner, 1974a). Genetic analyses, including PCoA, DAPC, fastSTRUCTURE, and ADMIXTURE, showed three distinct genetic groups, with no difference between *P. quadrifolia* and the purported *P. juarezensis*. This may initially appear to align with Lanner's hypothesis that the five-needled taxon would eventually be introgressed out of existence (Lanner, 1974a); however, purely five-needled trees were sampled and did not have a distinct genetic structure from the three- and four-needled samples identified as *P. quadrifolia*, implying they are all one species. With pine trees' long generation time and slow mutation rate (Buschiazzo *et al.*, 2012, De La Torre *et al.*, 2017) it is not plausible that a species could have disappeared from the gene pool since Lanner's observation (1974a) as a result of genomic swamping. It is more likely that *P. quadrifolia* can express three to five needles per fascicle. Additionally, purely five-needled trees were so rare compared to the mixed three-, four-, and five-needled individuals that they seem to be more of an exception, rather than the standard for a species. Lanner himself reported such rarity in his initial study, with only three individuals having 100% five needle coverage for one year but then expressing three or four needles in subsequent years (Lanner 1974a). With all this evidence considered, we reduce *P. juarezensis* to a synonym of *P. quadrifolia* supporting Farjon and Styles' (1997) hypothesis.

Pinus quadrifolia *is not of* P. monophylla *x* P. juarezensis *hybrid origin*

The unique needle number variation (i.e. three to five needles per fascicle) present in *P. quadrifolia* populations has led many to question its origin. Hybridization events between other pinyon species, such as *P. monophylla* and *P. edulis* for example, produce individuals with mixed needle number (Lanner, 1974b; Lanner and Phillips, 1992; Cole *et al.*, 2008). Previous research explored the possibility that the same mechanisms could be producing such variation in *P. quadrifolia*. For the first time, our study used NGS methods to test for the hybrid origin of *P. quadrifolia* suggested by the intermediate morphology between five-needle and one-needle pinyon pines. Genetic analyses showed no evidence that *P. quadrifolia* was of *P. monophylla* x *P. juarezensis* hybrid origin, but instead has a unique genetic identity. FastSTRUCTURE results showed *P. quadrifolia* as a distinct genetic group, with most individuals absent of genetic admixture. Individuals with a combination of three, four, and five needles were most common, yet showed no genetic distinction from the individuals with only four or only five needles. We therefore reject Lanner's declassification of the Parry pinyon to the hybrid epithet *P.* × *quadrifolia* and support its species status as *P. quadrifolia*. The morphological variation originally seen by Lanner and others is likely not due to hybridization events but may be explained by intraspecies variation or environmentally induced plasticity. In fact, in several tree stands where *P. quadrifolia* and *P. californiarum* grow together, the needle variation is not increased compared to isolated pure *P. quadrifolia* stands but instead this needle number range occurs throughout its distribution. This variation could be environmentally related, yet *P. californiarum* shares a large portion of its distribution with *P. quadrifolia* and does not seem to have nearly as much variation; however, the mode of monophylly inheritance is unknown (Gabilo and Mogensen, 1973).

Tridirectional hybridization: Detection of hybridization among the valid taxa P. quadrifolia, P. californiarum, and P. monophylla

Our data supports the existence of three taxa: *P. quadrifolia, P. californiarum*, and *P. monophylla.* Lanner (1974a) treated the one-needled individuals from Southern California and Baja California (San Jacinto Mountains, Rumorosa, San Pedro Martir) as *P.*

monophylla. We found that these populations are actually *P. californiarum* based on the genetic data. Furthermore, the morphology of these populations matches the morphology of other *P. californiarum* populations, with a high number of resin canals. We found no evidence of *P. monophylla* in Baja California. While *P. quadrifolia* is not of *P. monophylla* x *P. juarezensis* hybrid origin, it hybridizes with *P. californiarum* and *P. monophylla*. Previous studies hypothesized hybridization occurred between *P. quadrifolia* and *P. monophylla* (Zavarin *et al.*, 1980; Snajberk *et al.*, 1982; Kral, 1993); however, we also discovered genetic admixture between *P. quadrifolia* and *P. californiarum*. These two species share a significant amount of their range and even have stands where they grow a few meters apart. It appears that there are no strong genetic/postzygotic barriers preventing hybrids of these species from forming and proliferating within natural populations. However, more extensive pre- and postzygotic studies examining shifted phenology and outbreeding depression are needed to understand why some populations have more hybrid derivatives than others. A few instances of *P. quadrifolia* hybridizing with *P. monophylla* were seen in San Jacinto, and while these species' ranges do not overlap geographically, their nearest distributions are only 30 miles apart, a feasible distance for both pollen and seeds dispersed by corvids to travel (Wells, 1983b; Williams, 2010). This southern direction of gene flow would correspond to the Santa Ana winds that occur through October to May (Muhs *et al.*, 2007), well within the pollen dispersal times of *P. monophylla* and *P. quadrifolia* around May (Malusa, 1992; Farjon and Styles, 1997). Given the lower frequency of *P. monophylla* genes present in *P. quadrifolia*, these individuals may be relics from the early Holocene when *P. monophylla* had a more restricted southernly range, a pattern of distribution that has been detected in some populations of *P. monophylla* and its close relative *Pinus edulis* (Madsen and Rhode, 1990; Duran *et al.*, 2012). This contraction in distribution has been considered the source of hybridization in the latter (Duran *et al.*, 2012).

Surprisingly, a majority of the introgression detected in this study was between two species not previously thought to hybridize, *P. monophylla* and *P. californiarum*. Originally considered synonymous, these two species' potential for hybridization has long been overlooked, possibly due to their identical needle number. They share a significant portion of their range in southern Nevada and southern California but differ quite drastically in their number of resin canals. Curiously, in the site where hybrid derivatives are present (Mojave National Park), only one *P. californiarum* individual was detected, indicating how readily these two species hybridize. This could additionally imply heterosis, with increased fertility of the F_1 generation allowing for frequent backcrossing, or selection potentially favouring hybrid phenotypes, promoting the proliferation of individuals with hybrid ancestry. Future studies on the fitness of *P. monophylla* x *P. californiarum* crosses must be done to explore these hypotheses.

Morphological Traits in Pure Species vs Hybrids

Our study delimited species by genetic clusters first, without any *a priori* bias towards morphological traits. We then generalized the average characteristics of each genetic group to assign morphologic qualities to each species. In doing so, our analyses were not influenced by strict morphological cut-offs and we were able to more accurately identify individuals with hybrid ancestry and intraspecific variation compared to other studies in pinyon pines. It is important to note that while some distinguishing morphological patterns were found, these should be used as suggestions for species identification, rather than strict species guidelines/boundaries.

Interestingly, the morphology of all three hybrid types largely depends on the species contributing more of its genotype. For example, in *P. quadrifolia x P. californiarum* hybrid derivatives, individuals with a majority of *P. californiarum* genotype tend to have one to two needles per fascicle with over ten resin canals; whereas individuals with a majority of *P. quadrifolia* genotype tend to have two to four needles per fascicle and two to four resin canals. This implies needle morphology is largely genetically controlled and may be less influenced by environment than previously thought. The only exceptions to this were in recent hybrid derivatives, whose traits were more intermediate between their parents. For example, one *P. monophylla* x *P. quadrifolia* hybrid derivative found in San Jacinto had exclusively two needles per fascicle, 12 stomatal rows, and two resin canals, resembling *P. edulis*. It is possible that as these individuals of hybrid ancestry further backcross, the intermediate traits could be recombined or introgressed out to more closely represent one of the parental species. This unfortunately makes the identification of individuals with advanced generation hybrid ancestry quite difficult without using genetic methods because these hybrid derivatives are often indistinguishable from their main genetically contributing parent, concordant with other hybrid systems (Rieseberg *et al*., 1993).

This study provides an insight into the morphological and genetic outcomes of species undergoing gene flow. We plan to further explore these outcomes and patterns of hybridization in pinyon pines, especially in our newly discovered admixture combination of *P. monophylla* and *P. californiarum*. While our study addresses long-standing taxonomic issues within a small group of pinyon pines, the NGS methods and genetics first approach we used can be applied to a variety of taxa with controversial standings and cryptic introgression.

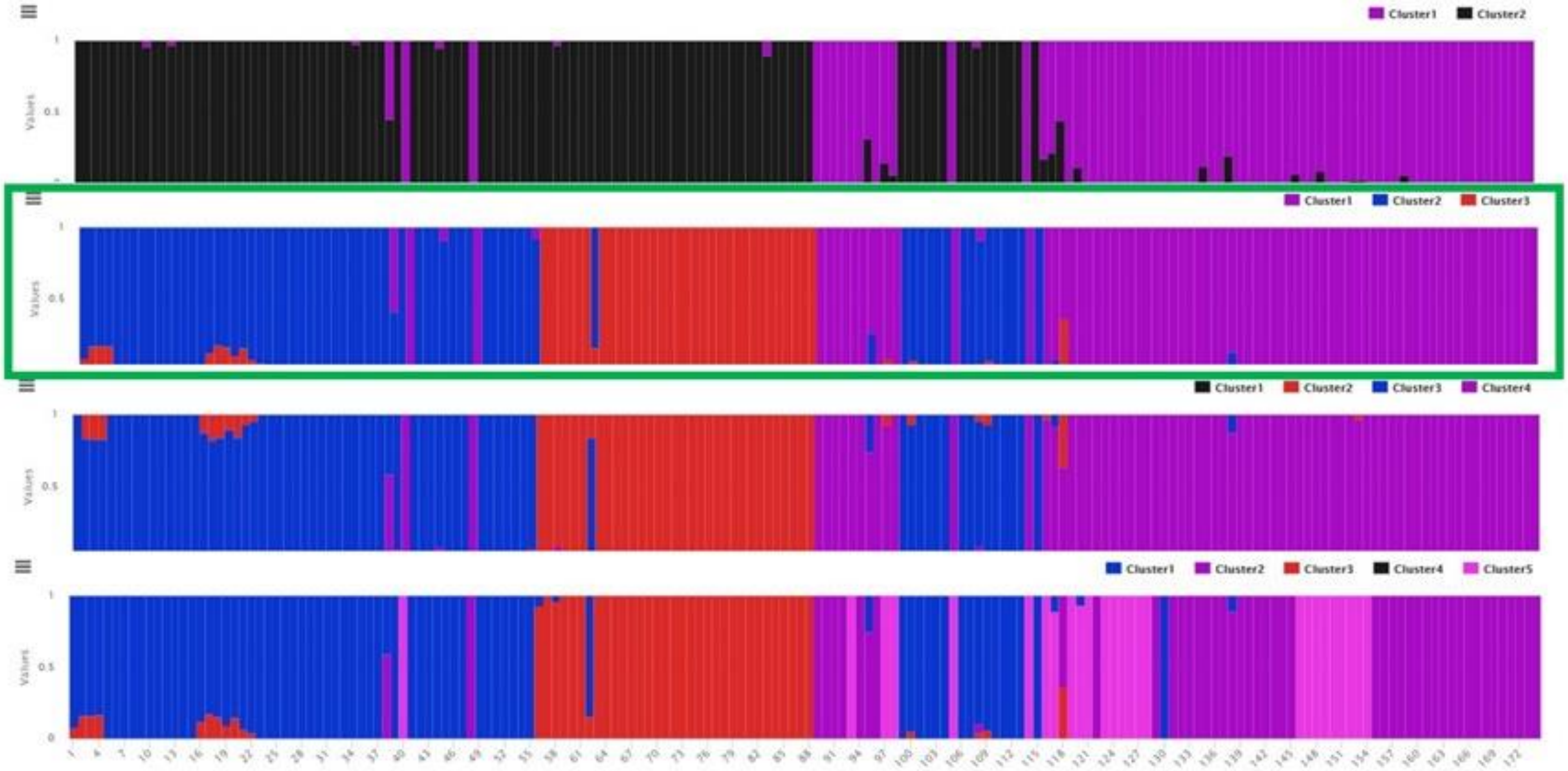
Funding: This work was supported by the Hispanic-Serving Institutions Education Grants (HSI) Program [grant number 2018-38422-28614/project accession no. 1016839] from the USDA National Institute of Food and Agriculture, by the National Science Foundation [grant number DEB-0816675], and by the University Grants Program of San Diego State University.

Acknowledgements: We thank the Forest Service, Mojave National Park, San Pedro Martir National Park, Drs. Sula Vanderplank, Jose Delgadillo, Rhett Michelson, Amy Whipple, Nuri Benet-Pierce, Rhett Michelson, Jamie Lamit, Hillary Cooper, Jean-Guillaume Lonjaret, Alexandra McElwee-Adame, Amy Orduño-Baez, Kyle Gunther, Mario Galvan, Stephanie Stragier, and Brian Myers for facilitating plant collections, Katya Geissler for her help in leaf morphology, The Binational Species Group, and the Flores-Rentería lab for feedback to the manuscript.

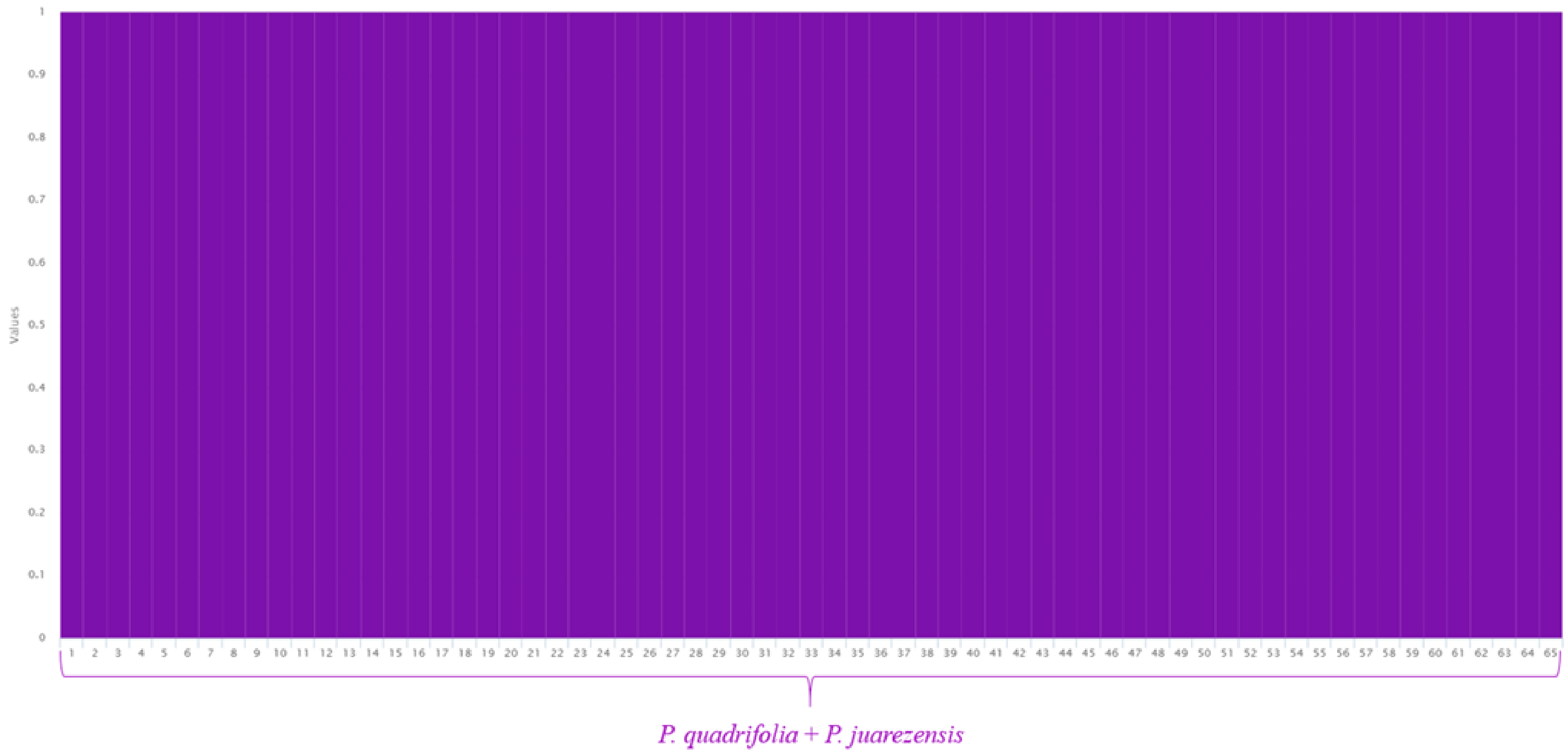
Author Contributions: R.B. and L.F-R. made study design, collected plant material, conducted lab work, analysed and interpreted data and wrote paper; S.H. and A.H.

collected morphological data and helped write the method section. All authors provided feedback to the manuscript.

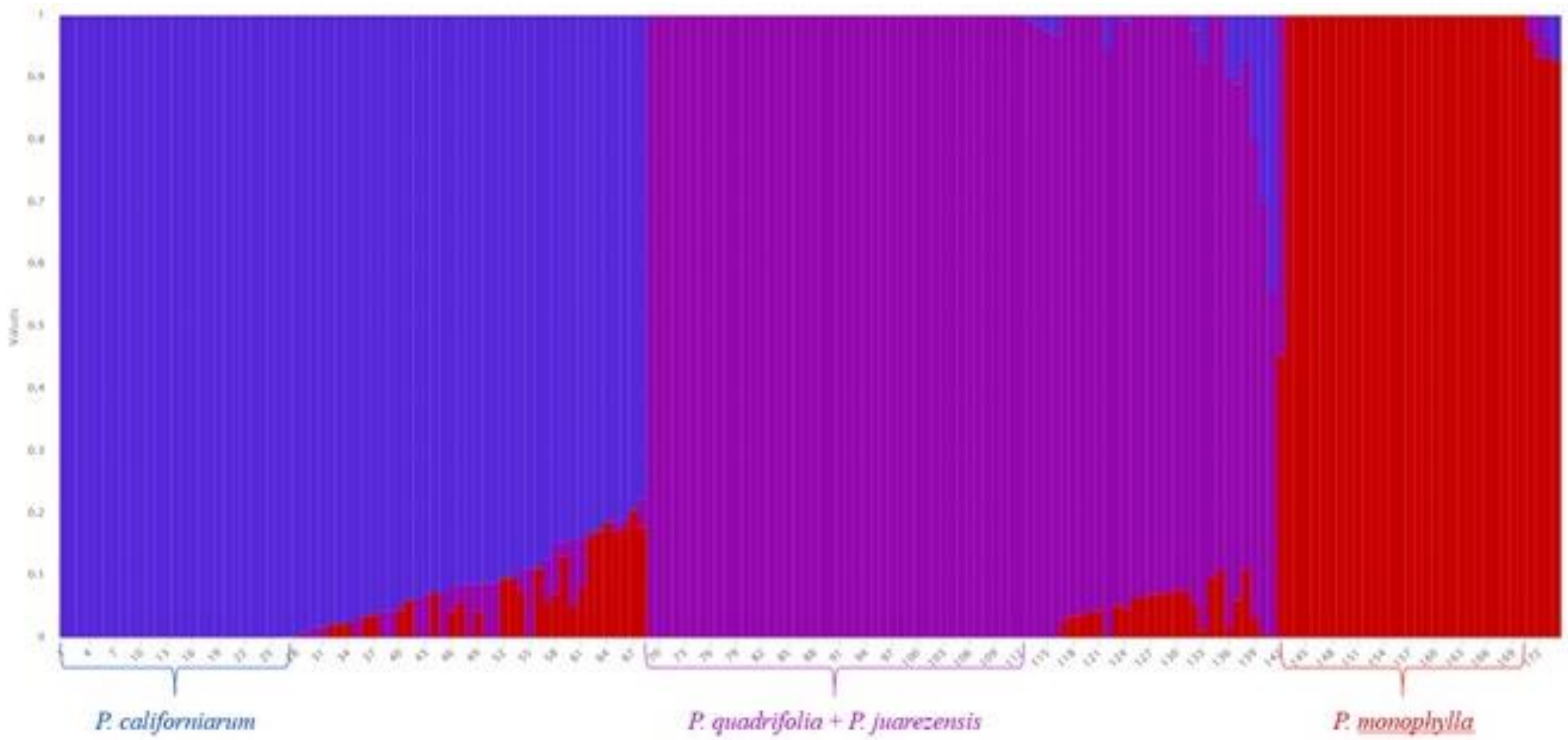
Supplementary Information: Three supplementary figures are included. The first shows the fastSTRUCTURE bar plots of a growing number of genetic clusters (K=2-5), with the software determined best K (K=3) framed. The second visualizes the substructuring of *P. quadrifolia* and *P. juarezensis* samples in fastSTRUCTURE, showing one genetic cluster (K=1). The third visualizes the ADMIXTURE analysis of all samples, showing three genetic clusters (K=3) corresponding to *P. monophylla*, *P. californiarum*, and *P. quadrifolia*+*P. juarezensis*.



Supplementary Figure 1. fastSTRUCTURE plots showing two to five genetic clusters (K= 2-5) with the best K (K=3) enclosed by a green square. Plots are colored by genetic identity (K=2: black = *P. californiarum* + *P. monophylla*, purple = *P. quadrifolia*+*P. juarezensis*; K=3: blue = *P. californiarum*, red = *P. monophylla*, purple= *P. quadrifolia*; K=4: similar to K=3 plus a cluster in black that is not distinguishable; K=5, as K=4 but it further splits *P. quadrifolia*= light purple; this secondary split of *P. quadrifolia* corresponds to geographical distribution, not morphology, with individuals in light purple from the northern distribution and dark purple from the southern. Each line on the x-axis represents an individual and the probability of those individuals belonging to a certain genetic cluster is represented by the proportion of color on the y-axis. Lines with multiple colors represent individuals with admixture from multiple genetic clusters.



Supplementary Figure 2. Hierarchical fastSTRUCTURE analysis of only the cluster containing the putative *P. juarezensis* and *P. quadrifolia* showing one genetic cluster (K=1) colored by genetic identity (purple = *P. quadrifolia*+*P.juarezensis*). Each line on the x-axis represents an individual and the probability of those individuals belonging to a certain genetic cluster is represented by the proportion of color on the y-axis.



Supplementary Figure 3. ADMIXTURE analysis showing three genetic clusters (K=3) colored by genetic identity (red = *P. monophylla*, blue = *P*. *californiarum*, purple = *P*. *quadrifolia*+*P*. *juarezensis*). Each line on the x-axis represents an individual and the probability of those individuals belonging to a certain genetic cluster is represented by the proportion of color on the y-axis. Lines with multiple colors represent individuals with admixture from multiple genetic clusters.

References

Abbott, R., Albach, D., Ansell, S., Arntzen, J.W., Baird, S.J., Bierne, N., Boughman, J., Brelsford, A., Buerkle, C.A., Buggs, R. and Butlin, R.K., 2013. Hybridization and speciation. *Journal of evolutionary biology*, *26*(2), pp.229-246.

Alexander, D.H., Novembre, J. and Lange, K., 2009. Fast model-based estimation of ancestry in unrelated individuals. *Genome research*, *19*(9), pp.1655-1664.

Anderson, E.C. and Thompson, E., 2002. A model-based method for identifying species hybrids using multilocus genetic data. *Genetics*, *160*(3), pp.1217-1229.

Bailey, .K., 1987. A study of *Pinus* subsection *Cembroides*. I: The single-needle pinyons of the Californias and the Great Basin. *Notes from the Royal Botanic Garden, Edinburgh*, *44*(2), pp.275-310.

Bigelow, R.S., 1965. Hybrid zones and reproductive isolation. *Evolution*, *19*(4), pp.449-458.

Buschiazzo, E., Ritland, C., Bohlmann, J. and Ritland, K., 2012. Slow but not low: genomic comparisons reveal slower evolutionary rate and higher dN/dS in conifers compared to angiosperms. *BMC Evolutionary Biology*, *12*(1), pp.1-15.

Chhatre, V.E., Evans, L.M., DiFazio, S.P. and Keller, S.R., 2018. Adaptive introgression and maintenance of a trispecies hybrid complex in range-edge populations of *Populus*. *Molecular ecology*, *27*(23), pp.4820-4838.

Christensen, K.M., Whitham, T.G. and Keim, P., 1995. Herbivory and tree mortality across a pinyon pine hybrid zone. *Oecologia*, *101*(1), pp.29-36.

Cole, K.L., Fisher, J., Arundel, S.T., Cannella, J. and Swift, S., 2008. Geographical and climatic limits of needle types of one-and two-needled pinyon pines. *Journal of Biogeography*, *35*(2), pp.257-269.

Couch, A.J., Unmack, P.J., Dyer, F.J. and Lintermans, M., 2016. Who's your mama? Riverine hybridisation of threatened freshwater Trout Cod and Murray Cod. *PeerJ*, *4*, p.e2593.

Critchfield, W.B., 1975. Interspecific hybridization in *Pinus*: a summary review. In*: Symp. on Interspecific and Interprovenance Hybridization in Forest Trees.[Ed] DP Fowler and CY Yeatman. Proc. 14th Meeting, Canad. Tree Improv. Assoc., Part II: p. 99-105.*Critchfield WB. 1986. Critchfield, W.B., 1986. Hybridization and classification of the white pines (*Pinus* section *Strobus*). *Taxon*, *35*(4), pp.647-656.

Delaporte, K.L., Conran, J.G. and Sedgley, M., 2001. Interspecific hybridization within *Eucalyptus* (myrtaceae): subgenus *symphyomyrtus*, sections *Bisectae* and *Adnataria*. *International Journal of Plant Sciences*, *162*(6), pp.1317-1326.

Cruz, V.M.V., Kilian, A. and Dierig, D.A., 2013. Development of DArT marker platforms and genetic diversity assessment of the US collection of the new oilseed crop lesquerella and related species. *PloS one*, *8*(5), p.e64062.

De La Torre, A.R., Li, Z., Van de Peer, Y. and Ingvarsson, P.K., 2017. Contrasting rates of molecular evolution and patterns of selection among gymnosperms and flowering plants. *Molecular Biology and Evolution*, *34*(6), pp.1363-1377.

Doyle, J.J., Doyle J.L, 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Focus*, *12*, pp.13-15.

Duran, K.L., Pardo, A. and Mitton, J.B., 2012. From middens to molecules: phylogeography of the piñon pine, *Pinus edulis*. *Journal of biogeography*, *39*(8), pp.1536-1544.

Farjon A., Styles B.T., 1997. *Flora Neotropica. Pinus (*Pinaceae*).* New York: Organization for Flora Neotropica, New York Botanical Garden.

Flores-Rentería, L., Wegier, A., Del Vecchyo, D.O., Ortíz-Medrano, A., Piñero, D., Whipple, A.V., Molina-Freaner, F. and Domínguez, C.A., 2013. Genetic, morphological, geographical and ecological approaches reveal phylogenetic relationships in complex groups, an example of recently diverged pinyon pine species (Subsection *Cembroides*). *Molecular phylogenetics and evolution*, *69*(3), pp.940-949.

Flores-Rentería, L., Rymer, P.D. and Riegler, M., 2017. Unpacking boxes: integration of molecular, morphological and ecological approaches reveals extensive patterns of reticulate evolution in box eucalypts. *Molecular Phylogenetics and Evolution*, *108*, pp.70-87.

Francis, R.M., 2017. pophelper: an R package and web app to analyse and visualize population structure. *Molecular ecology resources*, *17*(1), pp.27-32.

Gabilo, E.M. and Mogensen, H.L., 1973. Foliar initiation and the fate of the dwarf-shoot apex in *Pinus monophylla*. *American Journal of Botany*, *60*(7), pp.671-677.

Gernandt, D.S., Liston, A. and Piñero, D., 2001. Variation in the nrDNA ITS of *Pinus* subsection *Cembroides*: implications for molecular systematic studies of pine species complexes. *Molecular Phylogenetics and Evolution*, *21*(3), pp.449-467.

Gernandt, D.S., Liston, A. and Piñero, D., 2003. Phylogenetics of *Pinus* subsections *Cembroides* and *Nelsoniae* inferred from cpDNA sequences. *Systematic Botany*, *28*(4), pp.657-673.Grimm Grimm, G.W. and Denk, T., 2008. ITS evolution in *Platanus* (Platanaceae): homoeologues, pseudogenes and ancient hybridization. *Annals of Botany*, *101*(3), pp.403-419.

Gruber, B., Unmack, P.J., Berry, O.F. and Georges, A., 2018. dartr: An r package to facilitate analysis of SNP data generated from reduced representation genome sequencing. *Molecular Ecology Resources*, *18*(3), pp.691-699.

Gruber, B., Unmack, P., Berry, O., Georges, A., 2019. Introduction to dartR. https://rdrr.io/cran/dartR/f/inst/doc/IntroTutorial_dartR.pdf. 29 Nov 2019.

Holman, J.E., Hughes, J.M. and Fensham, R.J., 2003. A morphological cline in *Eucalyptus*: a genetic perspective. *Molecular Ecology*, *12*(11), pp.3013-3025.

Jaccoud, D., Peng, K., Feinstein, D. and Kilian, A., 2001. Diversity arrays: a solid state technology for sequence information independent genotyping. *Nucleic acids research*, *29*(4), pp.e25-e25.

Jasińska, A.K., Wachowiak, W., Muchewicz, E., Boratyńska, K., Montserrat, J.M., Boratyński. A., 2010. Cryptic hybrids between *Pinus uncinata* and *P. sylvestris*. *Botanical Journal of the Linnean Society* 163: pp.473-485.

Jombart, T., 2008. adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*, *24*(11), pp.1403-1405.

Jombart, T., Devillard, S. and Balloux, F., 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC genetics*, *11*(1), pp.1-15.

Jombart, T., Collins, C., 2015. *A tutorial for discriminant analysis of principal components (DAPC) using adegenet 2.0.* http://adegenet.r-forge.r-project.org/files/tutorial-dapc.pdf. 29 Nov 2019.

Kral R., 1993. *Pinus*. In: Flora of North America Editorial Committee, eds. *Flora of North America North of Mexico* New York: Oxford University Press, pp.373-398.

Lanner, R.M., 1974a. A new pine from Baja California and the hybrid origin of *Pinus quadrifolia*. *The Southwestern Naturalist*, pp.75-95.

Lanner, R.M., 1974b. Natural hybridization between *Pinus edulis* and *Pinus monophylla* in the American southwest. *Silvae Genetica, 23,* pp.108–116.

Lanner, R.M. and Phillips III, A.M., 1992. Natural hybridization and introgression of pinyon pines in northwestern Arizona. *International Journal of Plant Sciences*, *153*(2), pp.250-257.

Little, E.L., 1968. Two new pinyon varieties from Arizona. *Phytologia*, *17*, pp.329-342.

Madsen, D.B. and Rhode, D., 1990. Early Holocene pinyon (*Pinus monophylla*) in the northeastern Great Basin. *Quaternary Research*, *33*(1), pp.94-101.

Malusa, J.. 1992. Phylogeny and biogeography of the pinyon pines (*Pinus* subsect. *Cembroides*). *Systematic Botany, 17*, pp.42-66.

Melville, J., Haines, M.L., Boysen, K., Hodkinson, L., Kilian, A., Smith Date, K.L., Potvin, D.A. and Parris, K.M., 2017. Identifying hybridization and admixture using SNPs: application of the DArTseq platform in phylogeographic research on vertebrates. *Royal Society open science*, *4*(7), p.161061.

Menon, M., Bagley, J.C., Friedline, C.J., Whipple, A.V., Schoettle, A.W., Leal-Sàenz, A., Wehenkel, C., Molina-Freaner, F., Flores-Rentería, L., Gonzalez-Elizondo, M.S. and Sniezko, R.A., 2018. The role of hybridization during ecological divergence of southwestern white pine (*Pinus strobiformis*) and limber pine (*P. flexilis*). *Molecular Ecology*, *27*(5), pp.1245-1260.

Menon, M., Landguth, E., Leal-Saenz, A., Bagley, J.C., Schoettle, A.W., Wehenkel, C., Flores-Renteria, L., Cushman, S.A., Waring, K.M. and Eckert, A.J., 2020. Tracing the footprints of a moving hybrid zone under a demographic history of speciation with gene flow. *Evolutionary applications*, *13*(1), pp.195-209.

Montes, J.R., Peláez, P., Willyard, A., Moreno-Letelier, A., Piñero, D. and Gernandt, D.S., 2019. Phylogenetics of *Pinus* subsection *Cembroides* Engelm.(Pinaceae) inferred from low-copy nuclear gene sequences. *Systematic Botany*, *44*(3), pp.501-518.

Muhs, D.R., Budahn, J., Reheis, M., Beann, J., Skipp, G. and Fisher, E., 2007. Airborne dust transport to the eastern Pacific Ocean off southern California: Evidence from San Clemente Island. *Journal of Geophysical Research: Atmospheres*, *112*(D13).

Neri, J., Wendt, T. and Palma-Silva, C., 2018. Natural hybridization and genetic and morphological variation between two epiphytic bromeliads. *AoB Plants*, *10*(1), p.plx061.

Pfenninger, M., Reinhardt, F. and Streit, B., 2002. Evidence for cryptic hybridization between different evolutionary lineages of the invasive clam genus *Corbicula* (Veneroida, Bivalvia). *Journal of Evolutionary Biology*, *15*(5), pp.818-829.

Pritchard, J.K., Stephens, M. and Donnelly, P., 2000. Inference of population structure using multilocus genotype data. *Genetics*, *155*(2), pp.945-959.

Raj, A., Stephens, M. and Pritchard, J.K., 2014. fastSTRUCTURE: variational inference of population structure in large SNP data sets. *Genetics*, *197*(2), pp.573-589.

Rieseberg, L.H., Ellstrand, N.C. and Arnold, M., 1993. What can molecular and morphological markers tell us about plant hybridization?. *Critical reviews in plant sciences*, *12*(3), pp.213-241.

Rieseberg, L.H. and Willis, J.H., 2007. Plant speciation. *Science*, *317*(5840), pp.910-914.

Rutherford, S., Rossetto, M., Bragg, J.G., McPherson, H., Benson, D., Bonser, S.P. and Wilson, P.G., 2018. Speciation in the presence of gene flow: population genomics of closely related and diverging *Eucalyptus* species. *Heredity*, *121*(2), pp.126-141.

Snajberk, K., Zavarin, E. and Debry, R., 1982. Terpenoid and morphological variability of *Pinus quadrifolia* and the natural hybridization with *Pinus monophylla* in the San Jacinto Mountains in California. *Biochemical Systematics and Ecology*, *10*(2), pp.121-132.

Soltis, P.S. and Soltis, D.E., 2009. The role of hybridization in plant speciation. *Annual review of plant biology*, *60*, pp.561-588.

Twyford, A.D. and Ennos, R.A., 2012. Next-generation hybridization and introgression. *Heredity*, *108*(3), pp.179-189.

USGS. 1999. *Digital representation of "Atlas of United States Trees" by E.L. Little, Jr.* https://archive.usgs.gov/archive/sites/www.usgs.gov/science/cite-view.php-cite=22.html. 29 Nov 2019.

Wei, L., Li, Y.F., Zhang, H. and Liao, W.J., 2015. Variation in morphological traits in a recent hybrid zone between closely related *Quercus liaotungensis* and *Q. mongolica* (Fagaceae). *Journal of Plant Ecology*, *8*(2), pp.224-229.

Wells, P.V., 1983. Paleobiogeography of montane islands in the Great Basin since the last glaciopluvial. *Ecological Monographs*, *53*(4), pp.341-382.

Wenzl, P., Carling, J., Kudrna, D., Jaccoud, D., Huttner, E., Kleinhofs, A. and Kilian, A., 2004. Diversity Arrays Technology (DArT) for whole-genome profiling of barley. *Proceedings of the National Academy of Sciences*, *101*(26), pp.9915-9920.

Williams, C.G., 2010. Long-distance pine pollen still germinates after meso-scale dispersal. *American Journal of Botany*, *97*(5), pp.846-855.

Wright, J.W., 1952. Pollen dispersion of some forest trees. *Station Paper NE-46. Upper Darby, PA: US Department of Agriculture, Forest Service, Northeastern Forest Experiment Station. 42 p.*, *46*.

Willyard, A., Cronn, R. and Liston, A., 2009. Reticulate evolution and incomplete lineage sorting among the ponderosa pines. *Molecular Phylogenetics and Evolution*, *52*(2), pp.498-511.

Xiao, L.Q., Möller, M. and Zhu, H., 2010. High nrDNA ITS polymorphism in the ancient extant seed plant Cycas: incomplete concerted evolution and the origin of pseudogenes. *Molecular Phylogenetics and Evolution*, *55*(1), pp.168-177.

Zavarin, E., Snajberk, K. and Debry, R., 1980. Terpenoid and morphological variability of *Pinus quadrifolia* and its natural hybridization with *Pinus monophylla* in northern Baja California and adjoining United States. *Biochemical systematics and ecology*, *8*(3), pp.225-235.

Zhao, W., Meng, J., Wang, B., Zhang, L., Xu, Y., Zeng, Q.Y., Li, Y., Mao, J.F. and Wang, X.R., 2014. Weak crossability barrier but strong juvenile selection supports ecological speciation of the hybrid pine *Pinus densata* on the Tibetan plateau. *Evolution*, *68*(11), pp.3120-3133.

CHAPTER III

Sequential homoploid hybridization may have facilitated ecological transitions in the Southwestern pinyon pine syngameon

Ryan Buck[1], Diego Ortega-Del Vecchyo[2], Catherine Gehring[3], Rhett Michelson[4], Dulce Flores-Rentería[5], Barbara Klein[6], Amy V. Whipple[3], Lluvia Flores-Rentería[1]

[1]Department of Biology, San Diego State University, San Diego, CA, USA

[2]Laboratorio Internacional de Investigación sobre el Genoma Humano, Universidad Nacional Autónoma de México, Santiago de Querétaro, Querétaro 76230, México

[3]Department of Biological Sciences and Center for Adaptive Western Landscapes, Northern Arizona University, Flagstaff, AZ, USA

[4]Dept. of Biological Sciences, College of Southern Nevada, Nevada, USA

[5]CONACYT-CINVESTAV Unidad Saltillo, Grupo de Sustentabilidad de los Recursos Naturales y Energía, Av. Industria Metalúrgica 1062, Parque Industrial Ramos Arizpe, 25900, Ramos Arizpe, Coahuila, Mexico

[6]Diné College, School of Science, Technology, Engineering and Mathematics, Tsaile, AZ 86556, USA

Significance

While hybridization between two species has been widely studied, hybridization among three or more species, known as a syngameon, remains largely unexplored. These syngameonic interactions may facilitate evolutionary events such as niche diversification, speciation, and adaptive introgression. Here, we explore the genetic structure and evolutionary history of an established pinyon pine syngameon. We demonstrate that participation in the syngameon may have resulted in the creation and accelerated speciation of a drought tolerant hybrid species, which then hybridized with one of its parental species to create a second drought tolerant hybrid lineage. These traits have the potential to introgress across the syngameon network and protect other participating species from a changing world.

Summary

- Multispecies interbreeding networks, or syngameons, have been increasingly reported in natural systems. However, the formation, structure, and maintenance of syngameons have received little attention. Through gene flow, syngameons can increase genetic diversity, facilitate the colonization of new environments, and contribute to hybrid speciation.
- In this study, we evaluated the history, patterns, and consequences of hybridization in a pinyon pine syngameon using morphological and genomic data to assess genetic structure, demographic history, and geographic and climatic data to determine niche differentiation.
- We demonstrated that *Pinus edulis*, a dominant species in the Southwestern US and a barometer of climate change, is a core participant in the syngameon and involved in the formation of two drought-adapted hybrid lineages including the parapatric and taxonomically controversial *fallax*-type. We found that species remain morphologically and genetically distinct

at range cores, maintaining species boundaries while undergoing extensive gene flow in areas of sympatry at range peripheries.

- Our study shows that sequential hybridization may have caused relatively rapid speciation and facilitated the colonization of different niches resulting in the rapid formation of two new lineages. Participation in the syngameon may allow adaptive traits to be introgressed across species barriers and provide the changes needed to survive future climate scenarios.

Introduction

Natural hybridization has played an important role in the evolution of plants, birds, fish, fungi, insects, and mammals, including humans (Mallet 2005; Abbott *et al.*, 2013; Taylor and Larson, 2019). Although most studies on hybridization assess gene flow between two species in sympatry, hybridization can occur among multiple species simultaneously, with well documented examples in cichlids, oaks, heliconias, sunflowers, corals, and pinyon pines among others (Grant 1981; Rieseberg 1991b; Schliewen and Klee, 2004; Mallet *et al.*, 2007; Ladner 2012; Buck *et al.*, 2020). Multispecies interbreeding networks were defined as syngameons almost a century ago (Lotsy, 1925). Despite the long recognition of syngameons' existence, our understanding of how multiple species can collectively exchange genes and the resulting evolutionary consequences is limited (Buck and Flores-Rentería, in review). It is not understood how participating in a syngameon has affected the evolutionary history of species and how it may affect their future evolutionary trajectories, but some evidence shows syngameons may be responsible for rapid radiations, niche diversification, and even island colonization (Seehausen 2004; Caujapé-Castells and Bramwell, 2011; Liu *et al.*, 2017).

One hypothesis proposes that speciation events could be more common in a syngameon due to the increased amount of hybrid pairs (Seehausen 2004; Liu *et al.*, 2017). If a hybrid speciation event occurred within a syngameon, would reproductive barriers form at all, and if they did, would they remain and exclude the new hybrid species from joining the syngameon or would they quickly dissipate to enable syngameonic participation? While reproductive barriers aren't necessarily a requirement for hybrid speciation (Mallet 2007; Mavárez and Linares, 2008; Cannon 2021), understanding the process of speciation within a syngameon can give insight into the processes of speciation under gene flow and in sympatry. It is also not clear how species are able to maintain their identities in the face of extensive gene flow yet simultaneously act as a collective gene pool (Cannon and Petit, 2020). Spatial and temporal patterns of gene flow, as well as its direction and magnitude, constitute the structure of a syngameon (Ladner 2012; Boecklen 2017), but it is not known how these patterns affect the stability of the entire hybrid network. For example, intermittent waves of gene flow could facilitate the periodic introduction of genetic variation while providing enough time for traits to sort, allowing lineages to remain distinct (Cannon and Lerdau, 2015). Alternatively, constant interspecific gene flow could supply adaptive traits when needed through a sustained pool of genetic variation (Givnish 2010; Racimo *et al.*, 2015; Dannemann and Racimo, 2018; Menon *et al.*, 2021; Storz and Signore, 2021).

Pinyon pines (subsection *Cembroides*) are prone to hybridization due to their lack of genetic isolating mechanisms (Critchfield 1975), partially overlapping distributions (Mansfield-Jones 1967; Lanner 1971; USGS 1999; Cole *et al.*, 2003), and long-distance, wind-mediated pollination systems with overlapping pollen dispersal times (Malusa 1992; Farjon and Styles, 1997; Williams 2010). This is further supported by the intermediate morphology observed in areas of sympatry (Lanner 1974; des Lauriers and Ikeda, 1986). Therefore, pinyon pines are an excellent system to explore these multispecies hybridization dynamics and outcomes due to the known presence of a syngameon with at least three hybridizing species in the southwestern US that have varying patterns of hybridization (Buck *et al.*, 2020) and unique ecological niches (Cole *et al.*, 2008). Buck *et al.* (2020) was the first conclusive genetic study to demonstrate syngameonic hybridization in the Southwestern pinyon pines, examining gene flow among *Pinus monophylla*, *P. quadrifolia*, and *P. californiarum*. This paper includes two additional taxa with overlapping distributions (Fig. 1) that may also participate in this pinyon pine syngameon: *P. edulis* and the taxonomically controversial *fallax*-type (Lanner 1971; Lanner and Hutchison, 1972; Lanner 1974b). Understanding whether these taxa form part of the syngameonic network is critical because *P. edulis* is a dominant species in the Southwestern US and a barometer of climate change (Gitlin *et al*., 2006), experiencing massive rates of mortality after recent drought events (Ogle *et al.*, 2000; Mueller *et al.*, 2005; Mueller *et al.*, 2019). *Pinus edulis* is a foundation species, supporting thousands of other species above and below ground, including its avian and rodent seed dispersers (i.e. the scrub, pinyon, and Steller's jays, Clark's nutcracker, chipmunks, squirrels, deer mice, pinyon mice, Great Basin pocket mice, and Panamint kangaroo rats (Gottfried, 1987; Hollander and Vander Wall, 2004)), so its mass mortality has a large effect on the ecosystem it supports (Gehring *et al*., 2014; Paulin *et al*., 1999; Patterson *et al*., 2018; Stone *et al*., 2018). Compared to the other taxa in this syngameon, *P. edulis* inhabits higher elevation areas with monsoon-like precipitation, whereas *fallax*-type inhabits areas with monsoon-like precipitation but drier by 12mm and hotter by 2-4°C, showing that *P. edulis* and *fallax*-type have different niches (Cole *et al*., 2008). If *P. edulis* does participate in this syngameon, natural hybridization could facilitate the exchange of genes from more arid adapted taxa into *P. edulis*, potentially allowing it to withstand the changing climate and prevent further die-off. Looking at interspecific gene flow, examining demographic history, and establishing species barriers in this complex is therefore crucial to understanding the structure of a syngameon and the future of these species and the species they support. Also, syngameon networks might play a role in maintaining biodiversity in the face of climate change (Cannon and Petit, 2020).

While hybridization between *fallax*-type and *P. edulis* has been proposed (Lanner and Phillips, 1992; LaHood 1995), no genetic studies have found evidence of admixture between the two taxa (Duran *et al*., 2012; Montes *et al*., 2019). *Fallax*-type typically has one needle per fascicle with two resin canals (Little 1968; Cole *et al*., 2008), whereas *P. edulis* has two needles and two resin canals. *Fallax*-type's mixture of traits might suggest a hybrid origin between the single-needled *P. monophylla* and the two-resin-canaled *P. edulis* (Businksý 2008). However, while intermediate morphology can indicate hybridization (Mavárez *et al*., 2006; Curtu *et al*., 2007), individuals with hybrid ancestry are often indistinguishable from their parental species after a few generations of backcrossing (Rieseberg *et al.*, 1993; Buck *et al.*, 2020), so it is unclear why admixed individuals would have this consistent and distinguishable *fallax*-type morphology (Bailey 1987; Cole *et al.*, 2008). Further, while *fallax*-type itself overlaps in distribution with *P. edulis*, it rarely occurs in the areas where *P. edulis* and *P. monophylla* overlap. The unique niche occupied by *fallax*-type (Cole *et al.*, 2008) along with its distinct morphological traits suggest that it may either be a valid taxon or an incipient species, regardless of hybrid origins. However, if *fallax*-type does have hybrid origin and is also its own species, it would be a rare case of hybrid speciation (Rieseberg 1997; Goulet *et al.*, 2017). Alternatively, *fallax*-type might currently be undergoing speciation, providing an extraordinary glimpse into ongoing incipient speciation processes, especially in the face of extensive syngameonic gene flow.

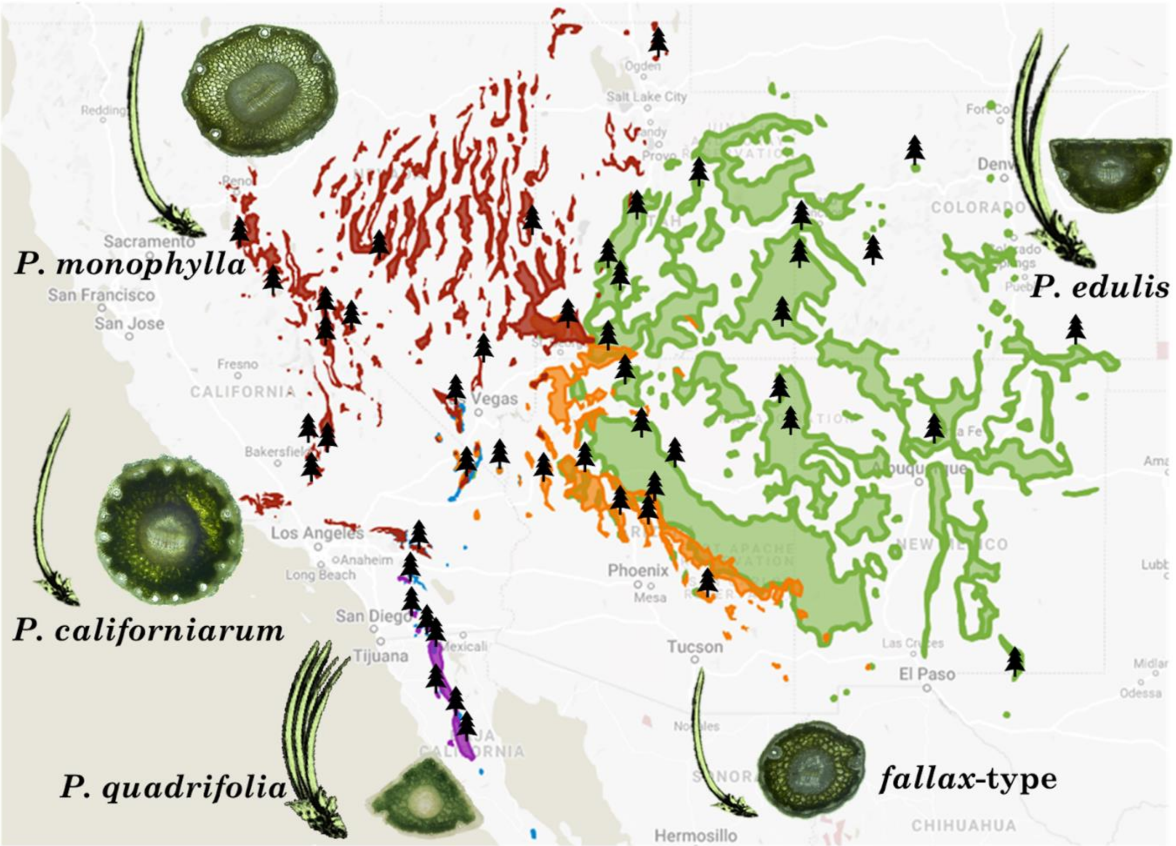


Figure 1. Distribution map colored by taxa identified in Cole *et al.* (2008): *P. monophylla* (red), *P. edulis* (green), *P. quadrifolia* (purple), *P. californiarum* (blue), *fallax*-type (orange). Superimposed images display the number of needles per fascicle and leaf cross section, showing its typical shape and the number of resin canals (ducts in the periphery of the leaf). Sampling points are represented as black tree symbols (based on USGS 1999; Cole *et al.*, 2003).

In this study, we generated morphological and genomic data from 46 populations of the five taxa hypothesized to participate in the pinyon syngameon present in Southwestern North America (Fig. 1). We used these data to 1) determine whether *P. edulis* and *fallax*-type participate in the known pinyon pine syngameon, 2) to elucidate whether any syngameon members (i.e. *fallax*-type) are a result of hybrid speciation, and 3) assess the structure of the syngameon and explore its evolutionary history.

Materials and Methods

Sampling

A total of 1,131 branch samples were taken from 46 locations across the Southwestern US and Baja California (Table S1), randomly sampling ten trees per putative species per site, at least 30 meters apart from each other. The distribution of this complex extends from southern Idaho, U.S. to southern Baja California, Mexico and from the eastern tip of Oklahoma to California, with a majority of *P. edulis* occurring in New Mexico, Colorado, Utah, and Arizona, *fallax*-type occurring in Arizona and the southern regions of Utah, Nevada, and California, *P. monophylla* occurring in western Utah, Nevada, and central California, *P. californiarum* occurring in southern Nevada, southern California, and south into Baja California, and *P. quadrifolia* occurring in southern California and Baja California (Fig. 1). Ten centimeters of branch tips were cut from each tree, representing two to three years of growth and averaging approximately 50 fascicles per tree. Collections were made from 2011 to 2013 and 2017 to 2019.

Nuclear DNA Sequencing

Nuclear DNA was extracted using Doyle and Doyle's (1987) CTAB 2% protocol, quantified, and sent to Diversity Arrays Technology (DArT), who use genome reduction methods and next-generation sequencing platforms similar to ddRAD-seq but to target low-copy fragments (Jaccoud *et al.*, 2001). This method produces a higher number of informative SNPs across the genome that are then compared to a reference genome (Buck *et al.*, 2020). We selected *P. lambertiana* as the reference genome (GCA_001447015.2) because of its relatively close evolutionary relationship to the pinyon subsection *Cembroides*. A low-density DArTseq assay of 353 individuals (Table S1) subset from across the distribution (Fig 2b) resulted in 24,109 SNPs, coded as “0” (homozygous to reference allele), “1” (homozygous to alternative allele, “2” (heterozygous), and “-” (fragment missing). Appropriate file types needed for downstream analyses were reformatted from the processed marker data using the R program dartR (Gruber *et al.*, 2018).

Genomic clustering analyses

From the original raw SNP dataset, SNPs were filtered to remove all loci with more than 10% missing data, all monomorphs, all loci with a reproducibility quality score lower than 1, all loci departing from Hardy-Weinberg equilibrium, and all but one locus where there was more than one locus per sequence tag, resulting in a final cleaned data set of 1,891 loci and 353 individuals. We analyzed patterns of population structure among individuals using a Principal Coordinates Analysis (PCoA) with dartR using the gl.pcoa.plot command and 95% confidence intervals (Gruber *et al.*, 2019). Additionally, we determined population clusters in the Bayesian software fastSTRUCTURE (Raj *et al.*, 2014) using the logistic prior and ten cross-validations. The chooseK command was used to select model complexity (K). The resulting Q mean bar plots were visualized using the online application pophelper (Francis, 2017).

Genetic Differentiation Analyses

Fst - As hybridization can bias fixation index results (Nei 1986), we separated the pure individuals from the admixed individuals identified in our fastStructure analyses (Q score < 1) and used the stamppFst command in the R package *adegenet* (Jombart 2008) to calculate the F statistic, applying 1,000 bootstraps and a 95% confidence interval.

Heterozygosity - In order to estimate the genetic diversity of each taxon, we calculated the heterozygosity of each taxon, using the same dataset as in Fst, and implemented the gl.report.heterozygosity command in the *adegenet* package in R.

Hybrid Generation Identification

NewHybrids – We applied a Bayesian model-based clustering method to determine each individual's hybrid category in the software NewHybrids 1.0 (Anderson and Thompson, 2002), which uses a Markov chain Monte Carlo method to compute the posterior probability of an individual belonging to predefined ancestry categories (Table S2). The program only compares two parental genotypes at a time, so we created six data sets to represent each pairwise species cross by removing individuals with ancestry of the third and fourth species (e.g. for *P. monophylla* x *P. quadrifolia* analysis we removed pure *P. californiarum* and *P. edulis* individuals and individuals with *P. californiarum* or *P. edulis* ancestry). We initiated runs at different random starting points with a burn-in of 10,000 and 100,000 sweeps, using the Jeffrey's prior for both pi and theta (Couch *et al.*, 2016).

Demographic Parameter Estimation

fastsimcoal – We tested eleven demographic scenarios (Fig. S1) in fastsimcoal version 2.6 (Excoffier and Foll, 2011; Excoffier *et al.*, 2013). To increase the SNP count required for the analysis, we created a modified SNP dataset, keeping monomorphs, multiple SNPs per locus, and SNPs with less than 20% missing data, but removing individuals with more than 40% missing SNP data. This resulted in a dataset containing 307 individuals and 16,427 SNPs. All admixed individuals were grouped with their main genetic contributor identified in fastStructure (Q > 0.5). We used easySFS (https://github.com/isaacovercast/easySFS) to generate a folded multidimensional site frequency spectrum, downsampling our sample sizes to 30 individuals per deme due to the missing data in our SNP dataset and the computational limitations of fastsimcoal2. 457,303 monomorphic sites in the locus surrounding the SNPs were enumerated and added to the final site frequency spectrum input into fastsimcoal2. Each demographic model was replicated 100 times using 40 expectation-conditional maximization (ECM) cycles, a log precision of 18, and 200,000 simulations to calculate the likelihood. A mutation rate of 8.04×10^{-8} per site per generation as an average of the reported nuclear mutation rate for *Pinus* (Willyard *et al.*, 2007) and a generation time range of 40-80 years was used as these trees begin producing female cones around 40 years old and become sexually mature around 80 (Ronco 1990; Cobb *et al.*, 2002). We used Akaike's information criterion (AIC) to find the most probable model given the observed data. Confidence intervals on parameter estimates for the best-supported model were generated via block bootstrapping by randomly sampling 100 blocks of the SNP dataset and creating 100 new SFS per model in easySFS. These 100 SFS were then run 50 times each in fastsimcoal2 using the same parameters and conditions as the original runs for both of the best models. The estimates from the best runs (highest likelihood) out of the replicated 50 were then used as confidence intervals, removing the lowest and highest five estimates for each parameter. These analyses took 417 days of run-time using 356 processors on two computing clusters: UCR's High-Performing Computing Center (HPCC) and UNAM's Laboratorio Nacional de Visualización Científica Avanzada.

Chloroplast DNA analysis

Microsatellites and Multiplex Design - A total of 18 chloroplast microsatellite markers developed for subsection *Cembroides* (Flores-Rentería and Whipple, 2011) were considered in this study due to the lack of variation among chloroplasts. Amplification was done from DNA of *P. edulis, P. monophylla*, *P. quadrifolia*, *P. californiarum* and *fallax-type* (Table S1). Four of 14 markers were excluded due to homoplasy (compound microsatellite 102213), lack of variation (61350, 108909) or other causes (108297). Additional markers (26081, 72502, 15147, 48509 and 79293) were developed following Flores-Rentería and Whipple's (2011) protocol, which reduces stutter in mononucleotide repeats. Multiplex design, PCR conditions and subsequent genotyping analyses were performed as described in Flores-Rentería *et al.* (2013).

PCA - In order to visualize genetic variation among chlorotypes, we ran a Principal Components Analysis in R using the adegenet package. The resulting PCA was plotted using the ggplot2 package, coloring individuals based on their nuclear genetic clustering identified in fastStructure.

Niche Divergence

Maxent - To model current habitat suitability of the taxa in this study, environmental niche models were generated using MaxEnt v3.4.4 (Phillips *et al.*, 2006). The original 30 arc-second variables used to build the suitability models were 19 bioclimatic variables and an elevation raster from WorldClim2 (Fick *et al.*, 2017), and 36 categorical soil classes, 8 slope classes, and 5 slope aspect classes from the Harmonized World Soil Database (Fischer *et al.*, 2008). All rasters were cropped to latitudes 23°N and 45°N and longitudes 99°W and 125°W to encompass the study area. Taxa occurrence points included our collection sites, sample sites from Zavarin *et al.* (1990), as well as citizen science and herbaria data from the Consortium of California Herbaria (CCH2), Baja Flora, Global Biodiversity Information Facility (GBIF), and la Red Mundial de Información sobre Biodiversidad (REMIB). Occurrence points were hand filtered to remove any erroneous points and correct any misidentified taxa. This filtering resulted in a final observation count of 1449 for *P. monophylla*, 1353 for *P. edulis*, 1343 for *P. californiarum*, 745 for *P. quadrifolia*, and 172 for *fallax*-type. Independent models were created for each species by removing correlated variables ($|r| > 0.70$) (Merow *et al.*, 2013) and those with 0% contribution to the model, and using 10 bootstrapped replicates, jackknifing, a random test percentage of 30 (Hernández-Quiroz *et al.*, 2018), a maximum of 5000 iterations, a 1.5 regularization multiplier (Merow *et al.*, 2013), and a 10 percentile training presence.

ENMTools - To determine if niche models of taxa are divergent, niche models were compared in ENMTools 1.4 (Warren *et al.*, 2008, 2010) using the overlap test and either the identity test for species with overlapping distributions (*P. edulis* and *P.* x *fallax*, *P. quadrifolia* and *P. californiarum*) or the background test for allopatric species (all other species pairs) (Warren *et al.*, 2008). In the overlap test, a D value for each pairwise comparison was produced using the average rasters created in the Maxent runs above. For the identity tests, pairwise comparisons using 100 replicates and all input occurrence points generated a distribution of D values that were then compared to the D value from the overlap test. For the background tests, pairwise comparisons using 100 replicates and a background raster with a 30km buffer around occurrence points generated a distribution of D values that were then compared to the D value from the overlap test.

Morphology

In order to determine what morphological features are consistent with the genetic clusters and to detect intermediate morphology associated with hybridization, we analyzed a range of leaf morphological traits typical to pine analyses (Little 1968; Bailey 1987; Lanner and Phillips, 1992; Malusa 1992; Christensen *et al.*, 1995; Cole *et al.*, 2008; Flores-Rentería *et al.*, 2013) following the protocol used in Buck *et al.* (2020). We used an ultra high-resolution Nikon SMZ25 stereoscopic microscope zoom 0.5-1.6X and NIS Elements software for all measurements. We also ran a PCA analysis to visualize morphological differences among the taxa and estimate the contribution of each variable to the separation, normalizing the data using the R package *bestNormalize*.

Results

Genetic Structure

In order to determine whether *P. edulis* and *fallax*-type participate in the syngameon, we performed nuclear genetic clustering and admixture analyses on 353 individuals with 1,891 SNPs. Our fastSTRUCTURE analyses (Fig. 2a, Fig. S2) visualized in Pophelper (Francis 2017) and superimposed onto the distribution map to help discern the structure of the syngameon, show four distinct genetic clusters: *Pinus edulis*, *P. quadrifolia*, *P. monophylla*, and *P. californiarum*. Extensive admixture among all species is observed, with most admixed individuals appearing near areas of sympatry, while pure individuals exist at the range cores (Fig. 2b). Notably, the individuals identified as *fallax*-type show up as hybrid descendants of *P. edulis* and *P. californiarum*.

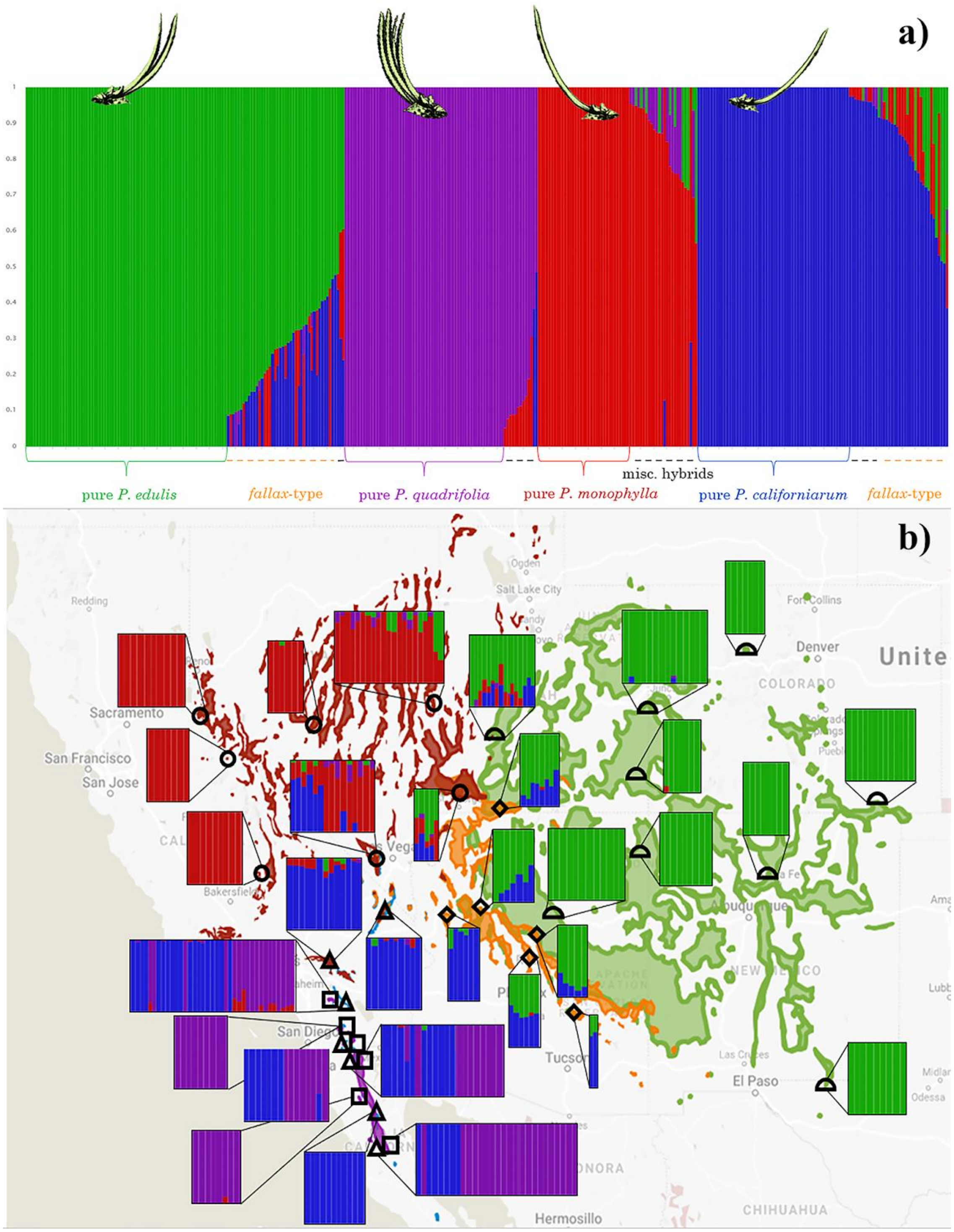
a)
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
pure P. edulis
fallax-type
pure P. quadrifolia
pure P. monophylla
misc. hybrids
pure P. californiarum
fallax-type
b)
Ogden
Salt Lake City
Redding
Fort Collins
Denver
United
COLORADO
Sacramento
San Francisco
San Jose
CAL
Bakersfield
Las Vegas
Albuquerque
NEW MEXICO
APACHE
Phoenix
Tucson
San Diego
Las Cruces
El Paso
Lubbock
Midland
Odessa
SONORA
CHIHUAHUA
Hermosillo

Figure 2. fastStructure plot showing four genetic clusters with extensive admixture among all species. A) Each vertical line represents an individual, with the color of the line illustrating the probability of that individual belonging to a certain genetic cluster. Lines with a single color represent “pure” individuals belonging to only one genetic cluster, while lines with multiple colors represent admixed individuals belonging to multiple genetic clusters. *Fallax*-type individuals are generally composed of only green and blue admixed bars. Superimposed figures represent the typical needle number associated with each genetic cluster. B) fastStructure plots on the distribution map by population showing admixture in areas of sympatry. Clusters are colored by genetic identity and populations are represented by different morphological symbols: *P. edulis* (green, half circles), *P. quadrifolia* (purple, squares), *P. monophylla* (red, circles), and *P. californiarum* (blue, triangles). Individuals identified as *fallax*-type (diamonds) appear as admixed individuals of *P. edulis* and *P. californiarum*.

Our Principal Coordinates Analysis (PCoA) (Fig. 3a) showed three groups of individuals composed of *P. monophylla*, *P. quadrifolia*, and a group of *P. edulis*, *P. californiarum*, and *fallax*-type overlapping. The circles surrounding each group are meant to aid in identifying the genetic background of its members. PCoA axis one accounted for 13.2% of the variation and matches the west to east distribution of the taxa, while axis two accounted for 7.9% of the variation and matches the south to north distribution of the taxa.

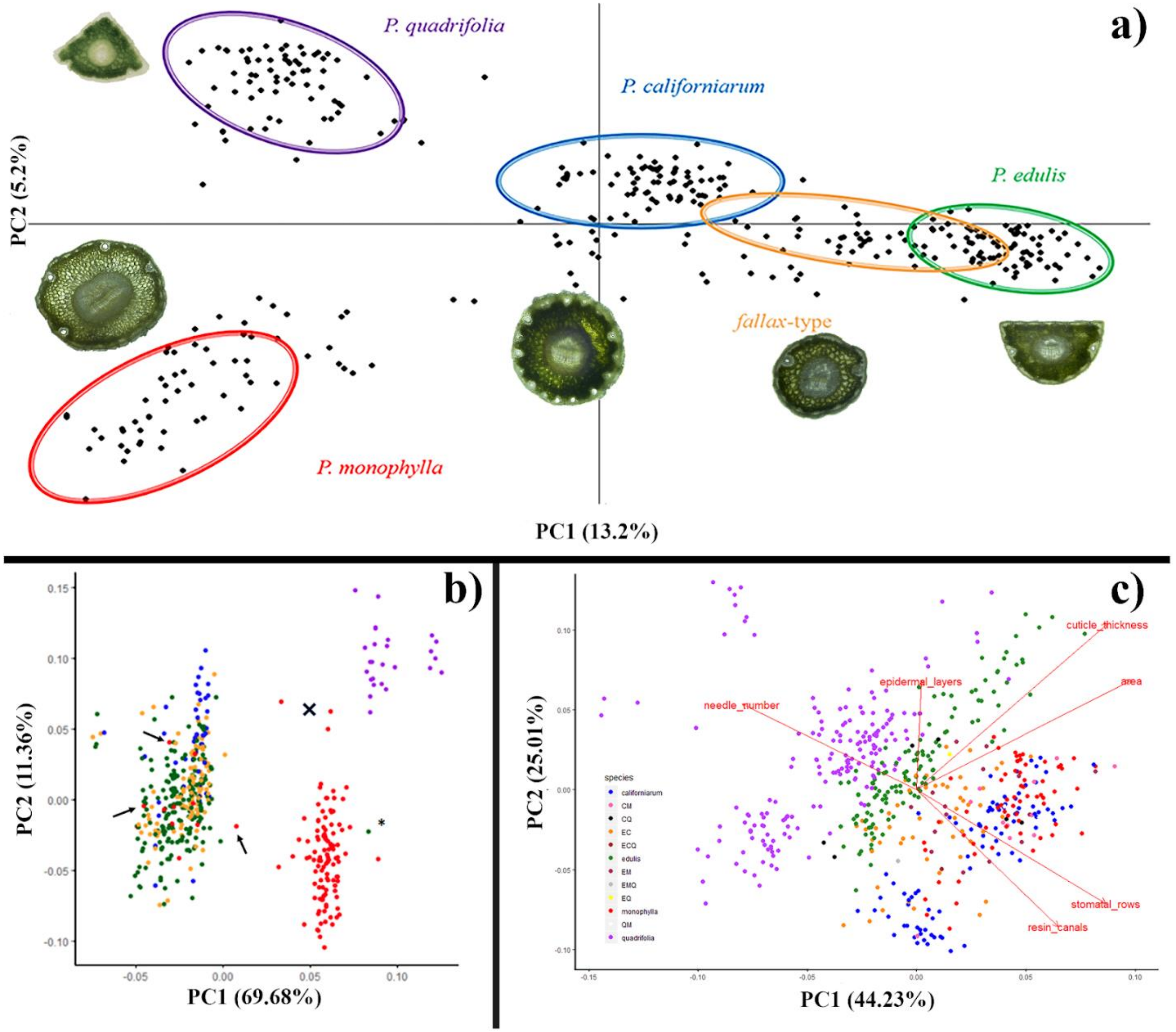
a)
P. quadrifolia
P. californiarum
P. edulis
fallax-type
P. monophylla
PC2 (5.2%)
PC1 (13.2%)
b)
PC2 (11.36%)
PC1 (69.68%)
0.15
0.10
0.05
0.00
-0.05
-0.10
-0.05
0.00
0.05
0.10
×
*
c)
PC2 (25.01%)
PC1 (44.23%)
0.10
0.05
0.00
-0.05
-0.10
-0.15
-0.10
-0.05
0.00
0.05
0.10
cuticle_thickness
area
epidermal_layers
needle_number
stomatal_rows
resin_canals
species
californiarum
CM
CQ
EC
ECQ
edulis
EM
EMQ
EQ
monophylla
QM
quadrifolia

Figure 3. Nuclear, chloroplast, and morphology shows patterns of hybridization and chloroplast capture on *P. edulis* (green), *P. californiarum* (blue), *fallax*-type (orange), *P. monophylla* (red), and *P. quadrifolia* (purple) with individuals represented as dots. A) Principal coordinates analysis using 1891 unlinked nuclear SNPs showing the variation in genetic distances among five main groups circled with 95% confidence intervals by color. Dots outside of circles corresponding to some of the individuals with hybrid ancestry. Individuals forming *fallax*-type emerge between *P. edulis* and *P. californiarum*. Images represent the typical leaf morphology and relative size for each genetic group. B) Principal component analysis using 18 chloroplast markers, suggesting *P. edulis*, *P. californiarum*, and *fallax*-type share a similar chloroplast, while *P. monophylla* and *P. quadrifolia* are distinct. Dots are colored by their main nuclear identity. One individual that had the nuclear identity of *P. edulis* had the chloroplast of *P. monophylla* (asterisk). In contrast, some trees that had the nuclear identity of *P. monophylla* had the chloroplast of *P. edulis/fallax*-type/*P. californiarum* (arrows) and *P. quadrifolia* (X mark), suggesting chloroplast capture events. C) Principal Component Analysis of the leaf morphology among species colored by nuclear genetic identity. Hybrid individuals are represented as *P. californiarum* x *P. monophylla* (CM), *P. californiarum* x *P. quadrifolia* (CQ), *P. edulis* x *P. californiarum* (EC), *P. edulis* x *P. californiarum* x *P. quadrifolia* (ECQ), *P. edulis* x *P. monophylla* (EM), *P. edulis* x *P. monophylla x P. quadrifolia* (EMQ), *P. edulis* x *P. quadrifolia* (EQ), *P. quadrifolia* x *P. monophylla* (QM).

Genetic differentiation

Our Fst results (Fig. 4b) indicated low differentiation between *fallax*-type and *P. californiarum* and *fallax*-type and *P. edulis*. Hybridization is expected to increase genetic diversity, therefore, we estimated heterozygosity as a measure of genetic diversity among taxa. The heterozygosity of each taxa were as follows: *P. quadrifolia* 0.0242, *P. edulis* 0.0269, *fallax*-type 0.0350, admixed individuals 0.0401, *P. monophylla* 0.0405, and *P. californiarum* 0.0479.

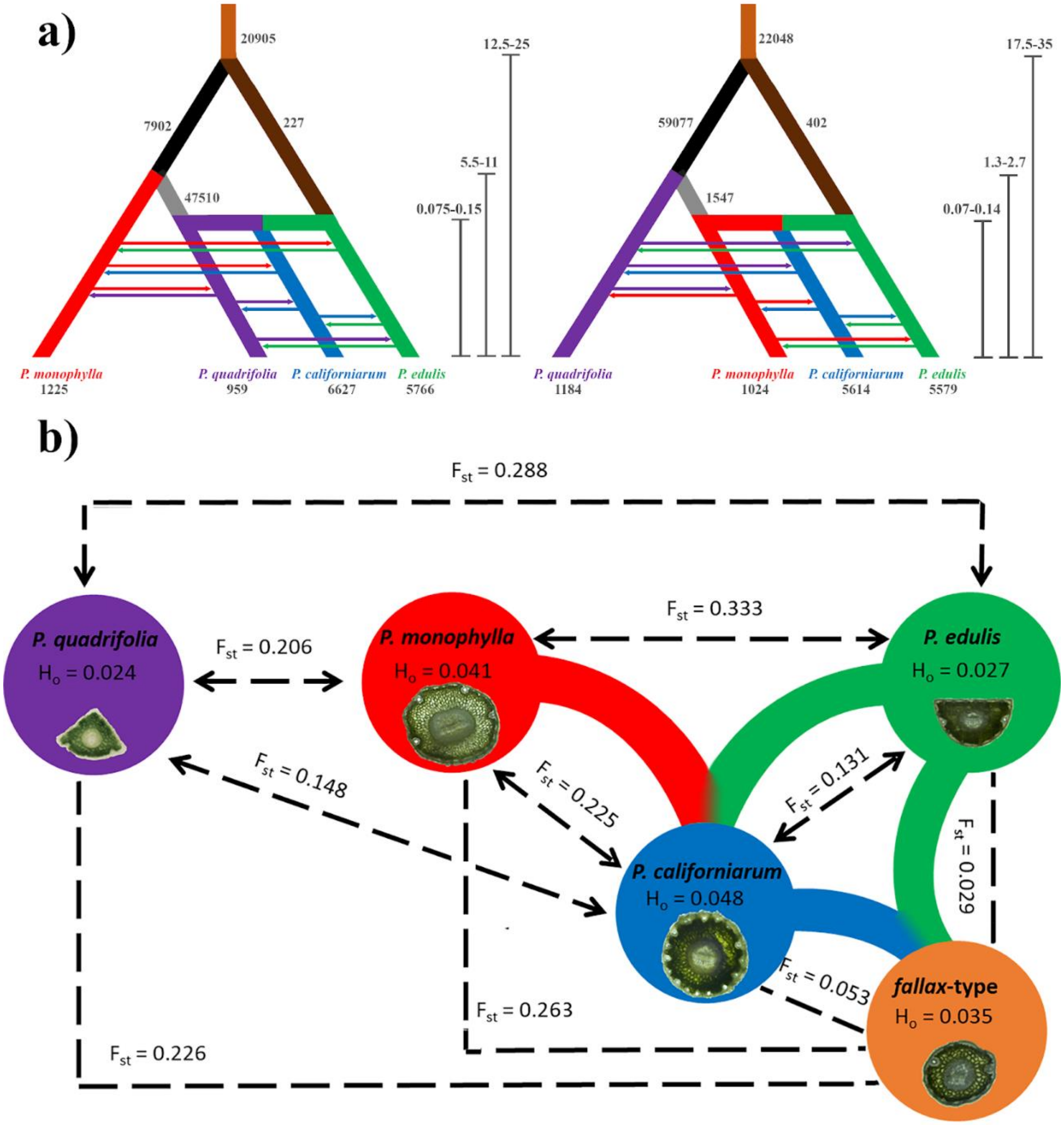


Figure 4. Hybridization has played a role in the speciation process in pinyon pines. A) The two most likely fastsimcoal scenarios determined by AIC calculation showing the hybrid origin of *P. californiarum* from *P. edulis* and either *P. monophylla* or *P. quadrifolia*. Estimated effective population sizes are represented by the numbers on the sides of the models, while estimated divergence times are reported in bars to the right of each model. B) Sequential hybridization diagram showing the hybrid origins of *P. californiarum* from *P. monophylla* and *P. edulis* and of *fallax*-type from *P. californiarum* and *P. edulis*. Dashed lines with arrows represent gene flow connections, with population differentiation (Fst) estimates shown above all dashed lines. Heterozygosity estimates are shown within each taxon's circle, along with an image of a needle cross-section typical for the taxon.

Chloroplast capture

As the chloroplast is inherited paternally in most pines (Neale and Sederoff, 1989; Kormutak *et al.*, 2018), if two species hybridize and subsequently backcross, one parental species can acquire the chloroplast of the other parental species, a process known as chloroplast capture. The comparison of nuclear and chloroplast data can aid in the understanding of parental contribution and the detection of chloroplast capture events and cryptic hybrids. In the chloroplast PCA analysis (Fig. 3b), we colored individuals based on their main nuclear genetic clustering in fastStructure and found three groups. *Pinus monophylla* and *P. quadrifolia* had distinct chlorotypes, while *P. edulis*, *P. californiarum*, and *fallax*-type share a similar chlorotype. One admixed individual with the nuclear identity of *P. edulis* grouped with the *P. monophylla* chlorotype (Fig. 3b - asterisk). Similarly, several admixed individuals with the main nuclear identity of *P. monophylla* grouped with the *P. edulis-P. californiarum-fallax-type* chlorotype (Fig. 3b - arrows) and the *P. quadrifolia* chlorotype (Fig. 3b - X). All of these individuals originate from the contact zones where *P. edulis* and *P. monophylla* overlap in distribution, suggesting chloroplast capture events.

Demography

In order to understand the temporal structure of this syngameon and estimate the direction and magnitude of gene flow we compared 11 demographic scenarios (Fig. S1) using fastsimcoal2.6 (Excoffier and Foll, 2011; Excoffier *et al.*, 2013). The two scenarios that best explain our data (Fig. 4a) revealed a recent hybrid speciation event, with either *P.*

quadrifolia or *P. monophylla* hybridizing with *P. edulis* 72,000-150,000 years ago to create *P. californiarum*; however the AIC scores were too close to determine which is the best model (Table S3). Extensive gene flow among all species was detected in the present time (Table S4), corroborating the syngameonic behavior found in the fastStructure results.

Hybrid Generation

We assessed whether the detected hybrid populations including *fallax-type* are recent and recurring F_1s or are stable lineages with advanced generation admixed individuals using the software NewHybrids with nuclear data (Anderson and Thompson, 2002). The hybrid generations determined by NewHybrids (Table S5) showed most of the admixed individuals identified in fastStructure as advanced generation backcrosses, with 20 individuals identified as fourth generation or further. 100 individuals had lower confidence in their generation assignment, potentially due to the pairwise limitations of NewHybrids or that they are beyond the generation categories tested.

Niche Divergence

To explore if the hybrid origins of *P. californiarum* and *fallax*-type could have been driven by differing niches, habitat suitability models were generated using Maxent v3.4.4 and compared in ENMTools v1.3. All models showed an AUC score greater than 0.9, suggesting they are useful models (Elith 2000; Phillips and Dudik, 2008). The suitability models indicated the niche of *P. californiarum* was divergent from *P. edulis* and *P. quadrifolia*, with potential divergence from *P.* x *fallax* and *P. monophylla*. *Pinus* *monophylla* had a divergent niche from *P. edulis*, with potential divergence from *P. quadrifolia*. Notably, *P.* x *fallax* showed a niche divergent from *P. edulis* and *P. monophylla*, with a potential divergence from *P.* x *fallax* and *P. quadrifolia* (Fig. S3).

Morphology

In order to determine if genetic clusters had identifiable morphological characteristics, we examined several needle traits. Our morphological analyses (Fig. 3c, Fig. S4) revealed that species can typically be identified by using the number of needles per fascicle, stomatal rows, and resin canals together. Individuals from pure *P. monophylla* descent tend to have one needle per fascicle with two to four resin canals and 20-31 stomatal rows. Pure *P. edulis* individuals usually have two needles per fascicle with one to two resin canals and 9-12 stomatal rows. Individuals from pure *P. quadrifolia* typically have four to five needles per fascicle (but individuals with three needles have been observed) with two resin canals and six to nine stomatal rows. Pure *P. californiarum* individuals generally have one needle per fascicle with 8-14 resin canals and 15-23 stomatal rows. *Fallax*-type has a varying morphology falling between *Pinus californiarum* and *P. edulis* (Fig. S4a, S4a), with 0-8 resin canals and 12-17 stomatal rows, depending on the contribution of each parent. Admixed individuals tend to have the morphology of the parent that contributes more to the genome, especially ones identified as advanced generation backcrosses (Fig. S4b, S4d).

Discussion

Our results showed that hybridization is widespread and has played an important role in the pinyon pine complex. This is the first study documenting sequential hybridization events in natural systems, in which a novel independent lineage resulting from gene flow hybridized with one of its parents to create a second novel lineage (Fig. 4b).

Syngameonic hybridization restricted to range edges

In order to understand the dynamics and structures of syngameons, it is crucial to determine what species are participating in them and how those species are connected spatially and through gene flow. Our genetic and morphologic data suggest *P. edulis, P. monophylla, P. quadrifolia,* and *P. californiarum* are distinct species that all hybridize among each other. *Pinus edulis* participates in this syngameon, having extensive admixture with the three described species, especially with *P. monophylla* and *P. californiarum*. Admixed individuals in this syngameon are generally restricted to areas of sympatry at the edges of the participating species' ranges, suggesting that these patterns are not produced by incomplete lineage sorting, supporting the range-edge formation hypothesis (Crohn and Suarez-Gonzalez, 2018), and giving insight into the structure of this syngameon. The lack of introgression further into the species' cores could explain how they can participate in the syngameon yet maintain their species' identities. Usually, a narrow hybrid zone is a result of postzygotic isolating mechanisms (Barton and Hewitt, 1985); however, our NewHybrids analyses showed that many individuals appear to be a result of several generations of admixture, implying that F_1s and F_2s are able to form, are viable, and are fertile enough to mate with each other and backcross with the parental species. Further, these pines lack genetic isolating mechanisms (Critchfield 1975) and have overlapping pollen dispersal times (Malusa 1992; Farjon and Styles, 1997). This may mean that there are no discernable intrinsic reproductive isolating barriers present among these species of pinyon pines and that extrinsic factors such as climate are more important in isolating species (Cole *et al.*, 2008; Menon *et al.*, 2018).

These patterns of restricted hybridization could be a result of strong selective pressures against the parental species in the edges of their ranges where hybrids may have a selective advantage (Pfennig *et al.*, 2016), resulting in the replacement and local extinction of pure parental genotypes i.e., genetic swamping (Todesco *et al.*, 2016). This genetic swamping could explain the patterns seen in a few isolated desert populations at the furthest range edges of *P. monophylla*, *P. californiarum,* and *fallax*-type that consist of mainly admixed individuals and few, if any, parental individuals and have significantly hotter and drier conditions. At the same time, there could be strong extrinsic factors selecting against unfit hybrids towards the core of the parental range, resulting in the local extinction of hybrids and the rarer parental lineage i.e.,demographic swamping (Todesco *et al.*, 2016), preventing further introgression and interspecific migration into range cores.

Hybrid origins of P. californiarum

Pinus edulis and *P. californiarum* shared a similar chlorotype, hinting at a possible shared ancestry or chloroplast capture event. This led to our exploration of demographic scenarios in which *P. californiarum* was of hybrid origin as well. Our two most likely models supported the chloroplast data, showing that *P. californiarum* was indeed a hybrid species from *P. edulis* and either *P. monophylla* or *P. quadrifolia*. The alternative model testing the origin of *P. californiarum* from a *P. monophylla* x *P. quadrifolia* hybridization event was not well supported (Fig. S1, model #22). The high rate of gene flow from *P. quadrifolia*/*P. monophylla* into *P. edulis* detected by the nuclear data coupled with the shared chlorotype data suggests that *P. edulis* was mainly a pollen donor while the hybrid descendants may have backcrossed multiple times with *P. quadrifolia*/*P. monophylla* to produce *P. californiarum*. The timing of this hybrid speciation event (72,000 – 150,000 years ago) is also intriguing because that is a relatively short time (around 1,800 generations using a generation time of 80 years) for pines to speciate (Menon *et al.*, 2018; Jin *et al.*, 2021), suggesting that hybridization may have accelerated the speciation process (Mallet 2007; Abbott *et al.*, 2013). The estimated ancient distribution of these pines is restricted to 50,000 years ago by the radiocarbon dating limit and presence of packrat middens (Cole *et al.*, 2013), so we cannot be sure which species shared distributions during the creation of *P. californiarum*. While *P. monophylla* and *P. edulis* are thought to have begun overlapping in distribution at some time during the last Full Glacial 14,700 to 23,400 years ago (Cole *et al.*, 2013), it is possible they were in contact at an earlier stage before migrating north. Unfortunately, not much is known about ancient range of the more Mexico-distributed *P. quadrifolia* beyond 18,000 years ago (Wells *et al.*, 2000), however *P. edulis*' distribution was considerably farther south than it currently is (Duran *et al.*, 2012; Cole *et al.*, 2008, 2013), therefore, it is possible that their ranges once overlapped or were at least close enough for pollen and seed dispersal (Wells 1983; Williams 2010). Future studies modeling climate beyond the reach of packrat midden data are needed to approximate ancient distributions and potential instances of sympatry. Hybridization with *P. monophylla* could explain the origin of its rare single-needle morphology or increased number of resin canals as a transgressive trait common in hybrid systems (Rieseberg *et al.*, 1999; Dittrich-Reed and Fitzpatrick, 2013). Further, a hybrid origin of *P. californiarum* could explain why it has the highest heterozygosity of any taxa, yet has the most restricted range. This inflated heterozygosity in turn could explain the large population sizes estimated for *P. californiarum*. Lastly, *P. californiarum* and *P. quadrifolia* often grow on the same mountain ranges often only miles apart in Southern California and Baja California, co-occurring in at least two locations (La Rumorosa and San Pedro Martir, Lanner 1974a). This neighboring distribution could be a remnant of shared ancestry or a sign of homoploid speciation without allopatry, as demonstrated in other pines (Menon *et al.*, 2018).

Hybrid origins of fallax-*type*

Individuals of *fallax*-type, originally identified as having one needle per fascicle and two resin canals (Little 1968), appear to be of hybrid descent, composed of both *P. edulis* and *P. californiarum*; therefore, we will hereby refer to it as *P.* x *fallax*. As *P.* x *fallax* is of hybrid origin, we expect to see its chlorotype grouping with either *P. edulis* or *P. californiarum* depending on the pollen donor, however, as *P. californiarum* likely inherited its chloroplast from *P. edulis*, the three taxa shared similar chlorotypes, while *P.*

quadrifolia and *P. monophylla* remained distinct. *Pinus edulis* and *P. californiarum* have never been hypothesized to hybridize, probably due to the lack of both current overlapping distribution and samples with intermediate morphology, characteristics most authors have used to identify hybrids in this complex (Cole 1956; Little 1968; Lanner 1974a,b; Lanner and Phillips, 1992). As the non-intermediate *P.* x *fallax* morphotype has been found in middens from over 48,000 years ago (Cole *et al.*, 2013), its stability likely also contributed to the lack of hybrid origin hypotheses. While F_1 hybrids of *P. monophylla* x *P. edulis* and *P. monophylla* x *P. quadrifolia* have consistent morphotypes (Critchfield in Lanner 1974b, Buck *et al.*, 2020), further backcrossing reverts these later generation crosses to a parental phenotype (Buck *et al.*, 2020). This pattern of cryptic hybridization is also exemplified here, in which admixed individuals of later generations appear to have indistinguishable/similar morphology from their parents (Fig. 3). However, *P.* x *fallax* populations are mostly composed of advanced generation admixed individuals, yet they are still a constant and distinguishable lineage even though the parental species, *P. edulis* and *P. californiarum*, have no currently overlapping distributions. The consistently forming morphology coupled with the uniquely hotter and drier niche that *P.* x *fallax* occupies (Bailey 1987; Cole *et al.*, 2008, Fig S3) point to an incipient species with a possible selective advantage allowing individuals of hybrid ancestry from *P. edulis* and *P. californiarum* to inhabit a new arid lower elevation environment. This process is similar to other hybridizing pine species in which the hybrid zones occur in a unique niche space that is hotter and drier than either parental species' habitats (Menon *et al.*, 2018). *Pinus* x *fallax*'s ability to occupy this niche may be a result of the introgression of drought-adapted traits from *P. californiarum*. If so, these traits may also be introduced into *P. edulis* or the other species *P. californiarum* is hybridizing with and could potentially help them survive under future climate scenarios. This is especially important as *P. edulis* is a particularly sensitive indicator of climate change (Gitlin *et al.*, 2006) whose distribution has been susceptible to changes in climate for the last 40,000 years (Betancourt 1987; Betancourt *et al.*, 1990).

Sequential Hybridization

The implications of a hybrid speciation event in a syngameon gives insight into the inner workings of a network undergoing frequent gene flow and emphasizes the potential creative force of hybridization. *Pinus californiarum* inhabits a unique niche that is both drier and hotter than that of *P. monophylla* and *P. edulis* (Cole *et al.*, 2008; Ortiz-Medrano *et al.*, 2016). This suggests that the colonization of a new niche, specifically driven by drought stress tolerance, could have promoted the hybrid speciation event, similar to the process demonstrated in *Helianthus* (Rieseberg *et al.*, 2007) and other pines (Menon *et al.*, 2021). This type of hybrid speciation may be happening again with *P. californiarum* and *P. edulis* creating the morphologically stable *P.* x *fallax*. Further, pines have been described as diploid (Ledig 1998; Grotkopp *et al.*, 2004; Ohri 2021), meaning that they could be examples of rare homoploid hybrid speciation (Goulet *et al*., 2017). For some authors, hybrid speciation is hard to prove, especially in systems with ongoing gene flow, because complete reproductive isolation is not guaranteed. However, many agree (Mallet 2007; Mavárez and Linares, 2008; Cannon 2021) that complete reproductive isolation is not necessarily required for hybrid speciation as long as there are several lines of evidence, including genetic, morphological, geographic, and ecological separation, all of which are presented here or have been extensively documented for this group (Cole *et al.*, 2008, 2013; Ortiz-Medrano *et al.*, 2016). This is the first study that documents sequential hybridization events resulting in the formation of novel independent lineages that occupy a different ecological niche.

In the pinyon syngameon, species remain morphologically and genetically distinct at range cores and are able to maintain species boundaries while undergoing extensive gene flow in areas of sympatry at range peripheries. Additionally, the syngameon seems to be structured with several "hub species" that contribute more genetic information than they receive and are connected to every other species through gene flow. Our study shows that hybridization has resulted in the formation of two new lineages in a relatively short time; thus, in this system, sequential hybridization may have accelerated the speciation process by providing a new combination of genes that allowed the colonization of different niches (Fig. S3). For example, *P. californiarum* which likely resulted from hybridization between *P. monophylla* and *P. edulis*, retained one needle per fascicle but with fewer stomata rows. Both traits may be useful in colonizing drier and hotter environments, as drier environments co-occur with a decrease in needle number (Ortiz-Medrano *et al.*, 2016), and fewer stomata have been associated with drought tolerance (Knauf and Bilan, 1974; but see Guérin *et al.*, 2018). *Pinus californiarum* and *P. edulis* hybridize to create *P.* x *fallax* individuals and adaptive introgression from *P. californiarum* may allow *P.* x *fallax* to colonize hotter and drier habitats. The shared chlorotypes of *P. edulis*, *P. californiarum*, and *P.* x *fallax* led us to explore the hybrid origin of *P. californiarum*. The most likely models showed *P. californiarum* resulting from a hybrid speciation event between *P. edulis* and either *P. monophylla* or *P. quadrifolia*. Hybrid speciation could explain *P. californiarum*'s ability to occupy a unique niche, its transgressive resin canal number, its shared chlorotype with *P. edulis*, and its inflated heterozygosity despite having the most restricted range. The participation of *P. edulis* in this syngameon and the formation of two drought-adapted lineages begs larger questions about the stability and future of the pinyon syngameon. If one participating member goes extinct or develops reproductive barriers, will the structure remain intact or will the whole network collapse? If the network collapses, will the newly isolated species survive on their own, or was their dependence on the syngameon critical to their existence? Participation in the syngameon may allow adaptive traits to be introgressed across species barriers and provide the changes needed to survive under future climate scenarios.

Funding: This work was supported by the Hispanic-Serving Institutions Education Grants (HSI) Program [grant no. 2018-38422-28614/project accession no. 1016839] from the USDA National Institute of Food and Agriculture, by the National Science Foundation [grant no. DEB-0816675], by PAPIIT-UNAM IA200620, UC MEXUS-CONACYT fellowship CN-19-29, and by the University Grants Program of San Diego State University.

Permits: USDA Permits for San Bernardino National Forest #2019-RB-SDSU and Cleveland National Forest #003571; California State Park Permit for Anza-Borrego

#CDD-2018-008-ABDSP; USNPS Permits for Mojave #MOJA-2018-SCI-0036 and Joshua Tree #JOTR-2018-SCI-0026; Mexico Permits #SPGA/DGVS/009790/18 and #SGPA/DGGFS/712/1631/18

Acknowledgements: We thank the Forest Service, the Navajo Nation, Joshua Tree National Park, Mojave National Preserve, San Pedro Martir National Park, Drs. Sula Vanderplank, Jose Delgadillo, Nuri Benet-Pierce, Jamie Lamit, Hillary Cooper, Lela Andrews, Adair Patterson, Arizona State Forestry Crew, Kim Haase, Todd Wojtowicz, Lisa Prato, Collin Hafey, Craig Allen, David Charlet, Amy Van Gundy, Miranda Redmond, Tegan May, Jean-Guillaume Lonjaret, Alexandra McElwee-Adame, Amy Orduño-Baez, Kyle Gunther, Mario Galvan, Stephanie Stragier, and Brian Myers for facilitating plant collections; Katya Geissler, Alice Hossfeld, Sandra Hyssat for their help measuring leaf morphology; The Binational Species Group and the Flores-Rentería lab for feedback on the manuscript. We thank UNAM LAVIS and UCR HPCC for providing access to their computer clusters.

Author contribution statement: L.F-R., R.B. and A.W. designed the study; R.B., L.F-R., R.M., and B.K. collected plant material; L.F-R., R.B., R.M. and D.F.-R. conducted lab work; L.F-R., D.O.D.V. and R.B. analyzed and interpreted data. R.B. and L.F-R. wrote the paper and made figures; L.F-R., C.G., and A.W. secured funding; L.F-R., C.G., and A.W. provided access to facilities. All authors provided feedback on the manuscript.

Data Availability: The data that support the findings of this study are openly available in Cyverse at https://data.cyverse.org/dav-non/iplant/home/rcbuck/Fallax_paper_dataset.csv

Supplementary Information:

Supplementary Table 1. Sample locations ordered by longitude (west to east) with population coordinates and number of trees sampled per population and per analysis type. California (CA), Nevada (NV), Arizona (AZ), Colorado (CO), New Mexico (NM), Texas (TX), Baja California (Baja CA).

Location	Latitude	Longitude	Needle type	# Trees nuclear genotyped	#Trees chlorotyped	# Trees morphotyped
Markleeville, CA	38.6989	-119.7708	*P. monophylla*	10	0	11
Mono Lake, CA	37.9215	-119.0643	*P. monophylla*	7	10	23
Lake Isabella, CA	35.5669	-118.4986	*P. monophylla*	9	0	26
Tehachapi, CA	35.1947	-118.3354	*P. monophylla*	0	0	15
Black Mountain, CA	37.2831	-118.1545	*P. monophylla*	0	0	15
Dyer, NV	37.7154	-118.1166	*P. monophylla*	0	0	6
Walker Pass, CA	35.6648	-118.0369	*P. monophylla*	0	7	7
Lida, NV	37.4399	-117.4878	*P. monophylla*	0	0	15
Manhattan, NV	38.5410	-117.0598	*P. monophylla*	6	22	32
Pinyon Pines, CA	33.5589	-116.6099	*P. quadrifolia* *P. californiarum* *Hybrids*	18 19	29 0	29 18 1
Mount Laguna, CA	32.8755	-116.4102	*P. quadrifolia*	9	0	11
Joshua Tree National Park, CA	34.0156	-116.1677	*P. californiarum* *Hybrids*	10	34	29
Jacumba Hot Springs, CA	32.6332	-116.0939	*P. californiarum* *P. quadrifolia*	7 7	0	10 10

La Rumorosa, Baja CA	32.5215	-116.0412	*P. quadrifolia* *P. californiarum* *Hybrids*	9 11	0	17 18
Puntas Coloradas, Baja CA	31.7424	-115.9782	*P. quadrifolia*	8	0	15
San Pedro Martir, Baja CA	30.9709	-115.5997	*P. quadrifolia* *P. californiarum*	21 7	0	39 9
Albérchigos, Baja CA	31.2576	-115.5994	*P. californiarum*	11	0	15
Mountain Springs, NV	35.9925	-115.5188	*P. monophylla* *P. californiarum* *hybrids*	15	58	49
Mojave Desert, CA	35.1739	-115.4071	*P. californiarum* *Hybrids*	13	12	22
Sheep Range, NV	36.8542	-115.1063	*P. monophylla*	0	23	12
Palm Gardens, NV	35.2637	-114.7520	*P. californiarum*	0	24	14
Baker, NV	38.9220	-114.1259	*P. monophylla*	21	22	32
Pinion Pines, AZ	35.1244	-113.9139	*P. californiarum*	6	22	11
Pine Valley, UT	37.4162	-113.6026	*fallax-type*	5	23	12
Seligman, AZ	35.2921	-113.0885	*fallax-type*	6	15	15
Kanab, UT	37.0705	-112.7019	*fallax-type*	8	12	12
Beaver, UT	38.2769	-112.6411	*P. edulis*	16	7	35
Panguitch, UT	38.0037	-112.5067	*P. edulis*	0	11	4
Williamson, AZ	34.6294	-112.4972	*fallax-type*	0	2	0
Jacob Lake, AZ	36.7157	-112.2171	*P. edulis*	0	11	11
Scipio, UT	39.1858	-112.1010	*P. edulis*	0	0	20

Wingfield, AZ	34.4722	-111.8373	*fallax-type*	6	22	31
Sedona, AZ	34.9146	-111.7254	*fallax-type*	5	12	20
Paradise, UT	41.5169	-111.7178	*P. monophylla*	0	8	4
Grand Canyon Village, AZ	35.8746	-111.6829	*fallax-type*	0	6	3
Flagstaff, AZ	35.3983	-111.4234	*P. edulis*	11	8	15
Castle Gate, UT	39.7197	-110.8667	*P. edulis*	0	2	2
Globe, AZ	33.3329	-110.8421	*fallax-type*	2	28	26
Blanding, UT	37.7378	-109.4067	*P. edulis*	7	10	10
Tsaile, NM	36.2998	-109.1799	*P. edulis*	9	25	14
Navajo, NM	35.9046	-109.0142	*P. edulis*	0	27	17
Glade Park, CO	38.9439	-108.9134	*P. edulis*	10	0	40
Cedar Point, CO	38.6194	-107.5907	*P. edulis*	4	0	6
McCoy, CO	39.9203	-106.7364	*P. edulis*	7	29	14
Bandelier, NM	35.8003	-106.2822	*P. edulis*	8	26	40
Pine Springs, TX	31.8887	-104.7873	*P. edulis*	9	9	10
Kim, CO	37.3524	-103.6998	*P. edulis*	9	0	30

Supplementary Table 2. Twelve genotype frequency categories input into NewHybrids analyses (based on Chhatre *et al*., 2018)

Genotype category	Cross type	Expected ancestry proportions			
		AA	Aa	aA	aa
Pure	Species 1	1	0	0	0
Pure	Species 2	0	0	0	1
F_1	Species 1 x Species 2	0	0.5	0.5	0
F_2	F_1 x F_1	0.25	0.25	0.25	0.25
F_1 Backcross 1	F_1 x Species 1	0.5	0.25	0.25	0
F_1 Backcross 2	F_1 x Species 2	0	0.25	0.25	0.5
F_2 Backcross 1	F_2 x Species 1	0.5	0.125	0.125	0.25
F_2 Backcross 2	F_2 x Species 2	0.25	0.125	0.125	0.5
1 Backcross x F_1 Backcross 1	Species 1 x (F_1 x Species 1)	0.75	0.125	0.125	0
2 Backcross x F_1 Backcross 2	Species 2 x (F_1 x Species 2)	0	0.125	0.125	0.75
1 Backcross x F_2 Backcross 1	Species 1 x (F_2 x Species 1)	0.625	0.125	0.125	0.125
2 Backcross x F_2 Backcross 2	Species 2 x (F_2 x Species 2)	0.125	0.125	0.125	0.625

Supplementary Table 3. AIC scores of the 11 models

Model	MaxEstLhood	Number of parameters	AIC score
20	-10500.43	24	21049
21	-10501.20	24	21050
22	-10502.71	24	21053
29	-10508.64	28	21073
192	-10525.69	22	21095
19	-10525.70	22	21095
16	-10525.76	22	21096
30	-10525.66	26	21103
31	-10525.69	26	21103
32	-10525.80	26	21104
18	-10525.45	28	21107

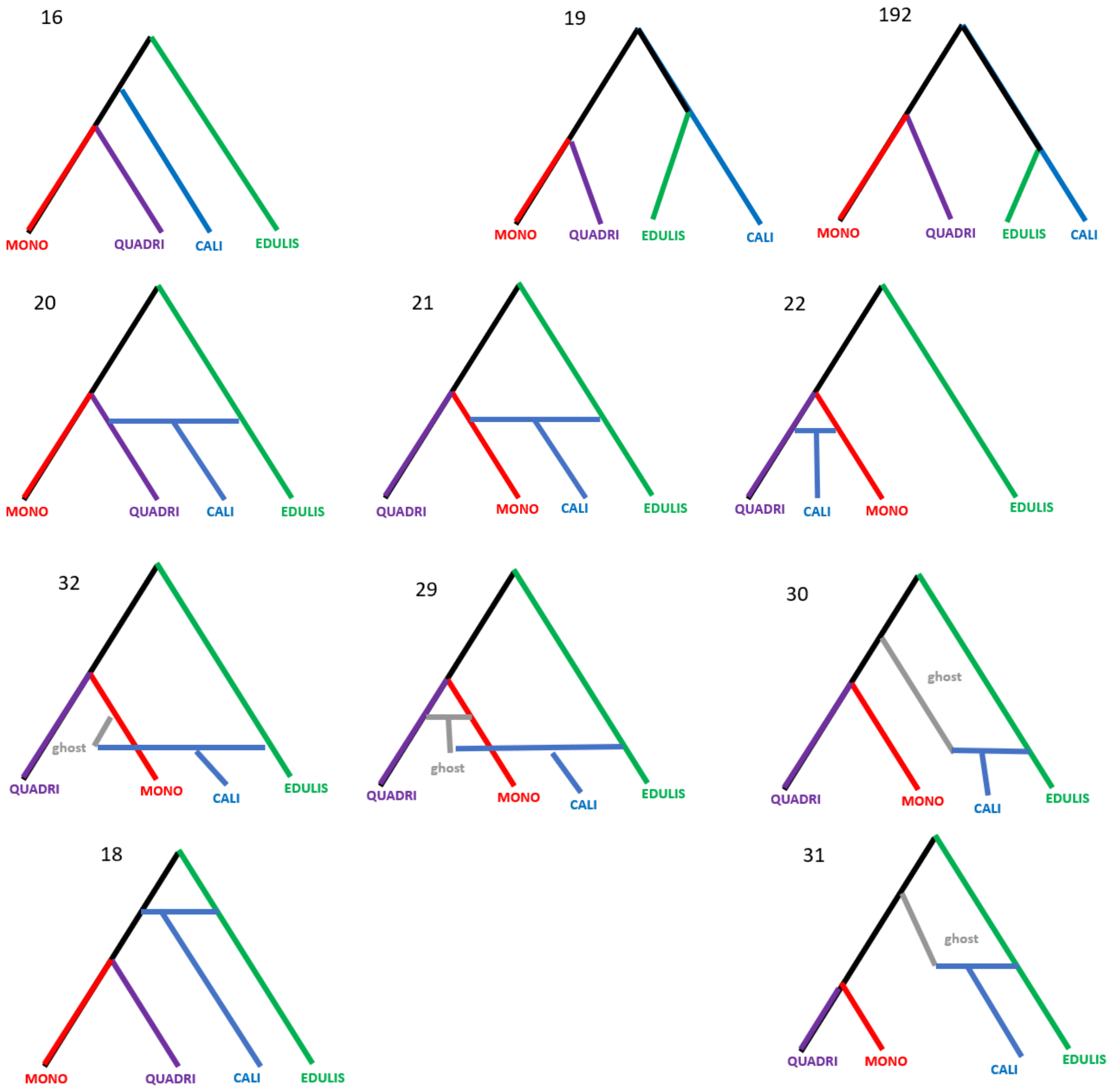
Supplementary Table 4. fastsimcoal2 results

Demographic parameter	Model 20 estimates	Model 21 estimates
N_e *P. monophylla*	1125	1024
N_e *P. edulis*	5766	5579
N_e *P. californiarum*	6627	5614
N_e *P. quadrifolia*	959	1184
Ne Ancestor of *P. californiarum- P. monophylla*	N/A	1547
N_e Ancestor of *P. californiarum- P. quadrifolia*	47510	N/A
N_e Ancestor of *P. californiarum- P. edulis*	227	402
N_e Ancestor of *P. monophylla- P. quadrifolia*	7902	59077
N_e *P. edulis- P. monophylla*	20905	22048
Time of *P. californiarum* hybrid origin	1871	1788
Time of *P. monophylla-P. quadrifolia* coalescence	139101	337796
Time of *P. monophylla-P. edulis* coalescence	315232	439076
Rate of *P. californiarum* creation	1.31x10-5	2.17x10-6
Gene flow from *P. edulis* into *P. monophylla*	0.48 migrants/generation	0.34 migrants/generation
Gene flow from *P. californiarum* into *P. monophylla*	3.29 migrants/generation	3.17 migrants/generation

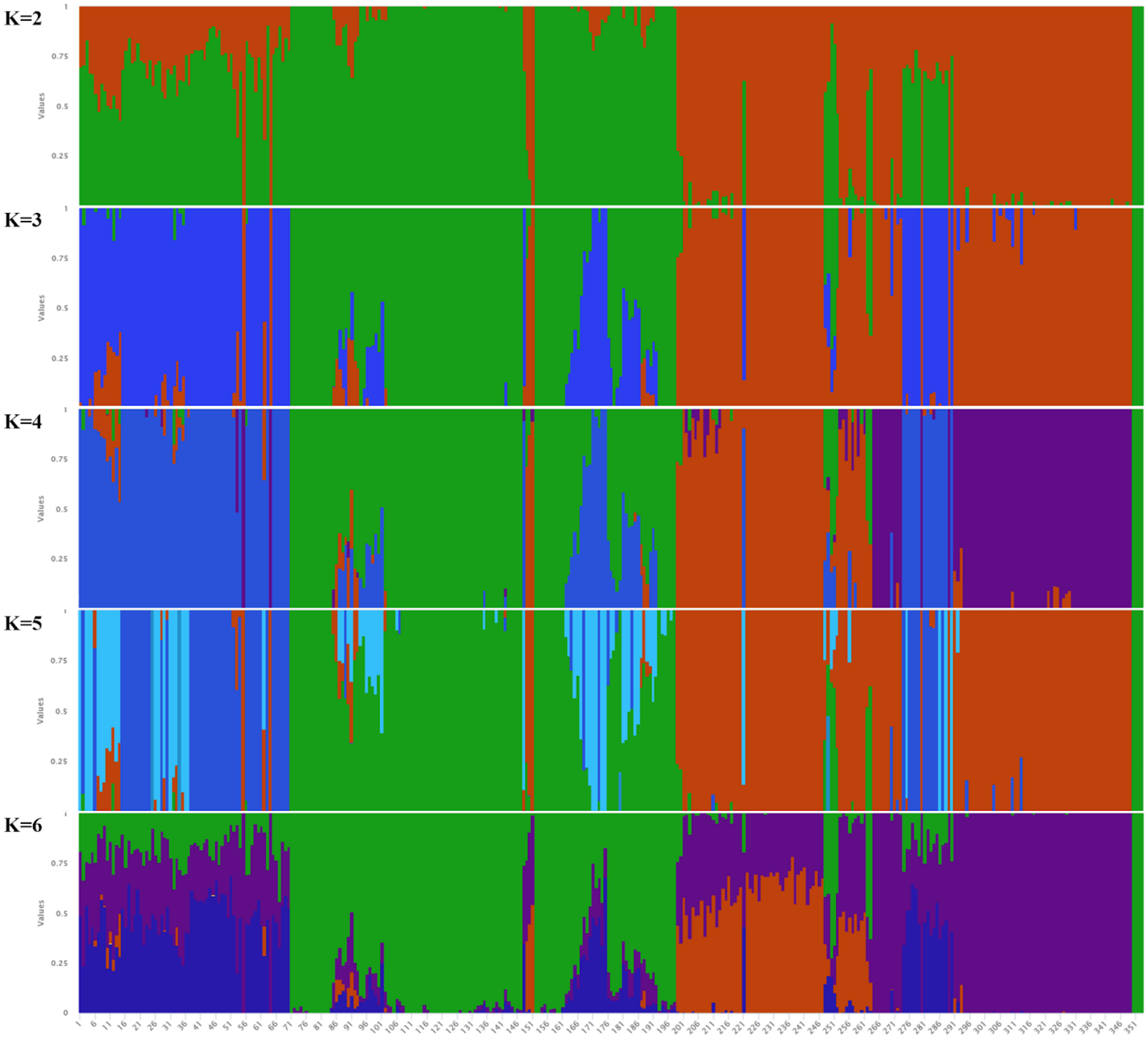
Gene flow from *P. quadrifolia* into *P. monophylla*	3.07 migrants/generation	3.54 migrants/generation
Gene flow from *P. monophylla* into *P. edulis*	0.01 migrants/generation	0.02 migrants/generation
Gene flow from *P. californiarum* into *P. edulis*	5.17 migrants/generation	5.86 migrants/generation
Gene flow from *P. quadrifolia* into *P. edulis*	0.02 migrants/generation	0.01 migrants/generation
Gene flow from *P. monophylla* into *P. californiarum*	0.01 migrants/generation	0.01 migrants/generation
Gene flow from *P. edulis* into *P. californiarum*	4.16 migrants/generation	3.91 migrants/generation
Gene flow from *P. quadrifolia* into *P. californiarum*	0.02 migrants/generation	0.01 migrants/generation
Gene flow from *P. monophylla* into *P. quadrifolia*	0.71 migrants/generation	0.16 migrants/generation
Gene flow from *P edulis* into *P. quadrifolia*	0.02 migrants/generation	0.06 migrants/generation
Gene flow from *P. californiarum* into *P. quadrifolia*	5.26 migrants/generation	6.88 migrants/generation

Supplementary Table 5. Levels of ancestry determined by NewHybrids for the hybrid derivatives identified in fastStructure analyses.

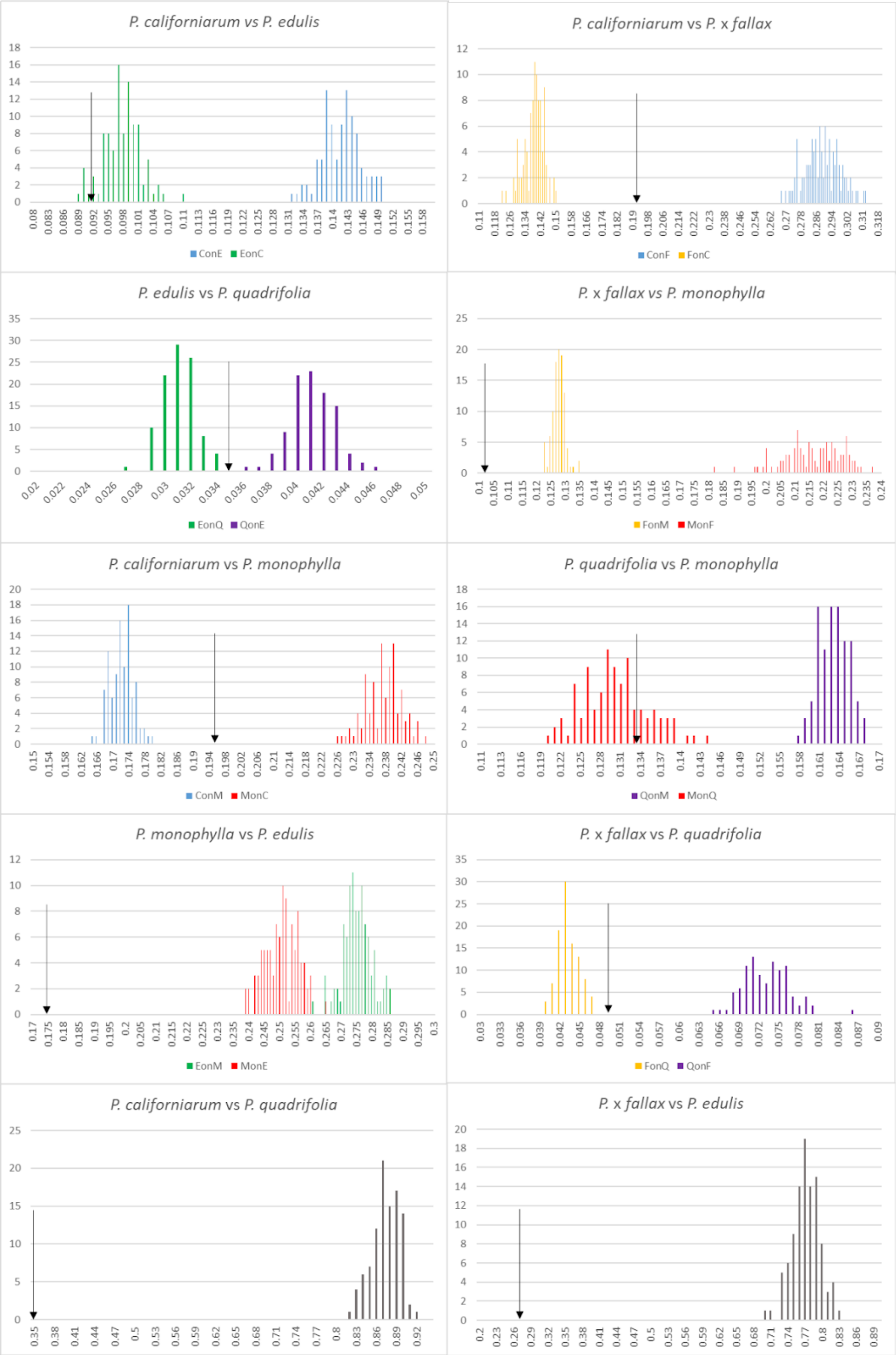
NewHybrids assignment	Contributing species	# indiv.
pure *P. californiarum*	*P. californiarum*	64
pure *P. edulis*	*P. edulis*	76
pure *P. monophylla*	*P. monophylla*	35
pure *P. quadrifolia*	*P. quadrifolia*	61
P. monophylla x (F_2 x *P. monophylla*)	*P. monophylla, P. quadrifolia*	6
P. quadrifolia x (F_2 x *P. quadrifolia*)	*P. monophylla, P. quadrifolia*	4
P. californiarum x (F_2 x *P. californiarum*)	*P. monophylla, P. californiarum*	7
P. edulis x (F_2 x *P. edulis*)	*P. monophylla, P. edulis*	2
P. monophylla (F_2 x *P. monophylla*)	*P. monophylla, P. edulis*	1
F_2 x *P. edulis*	*P. edulis, P. californiarum*	2



Supporting Information Figure 1. The topology of the 11 fastsimcoal scenarios tested. While not represented here, all models included estimates of divergence times, gene flow until the first historical event, and effective population sizes.



Supporting Information Figure 2. fastStructure results sorted by individual, showing K=2-6 with the best K=4.


P. californiarum vs *P. edulis*

P. californiarum vs *P. x fallax*

P. edulis vs *P. quadrifolia*

P. x fallax vs *P. monophylla*

P. californiarum vs *P. monophylla*

P. quadrifolia vs *P. monophylla*

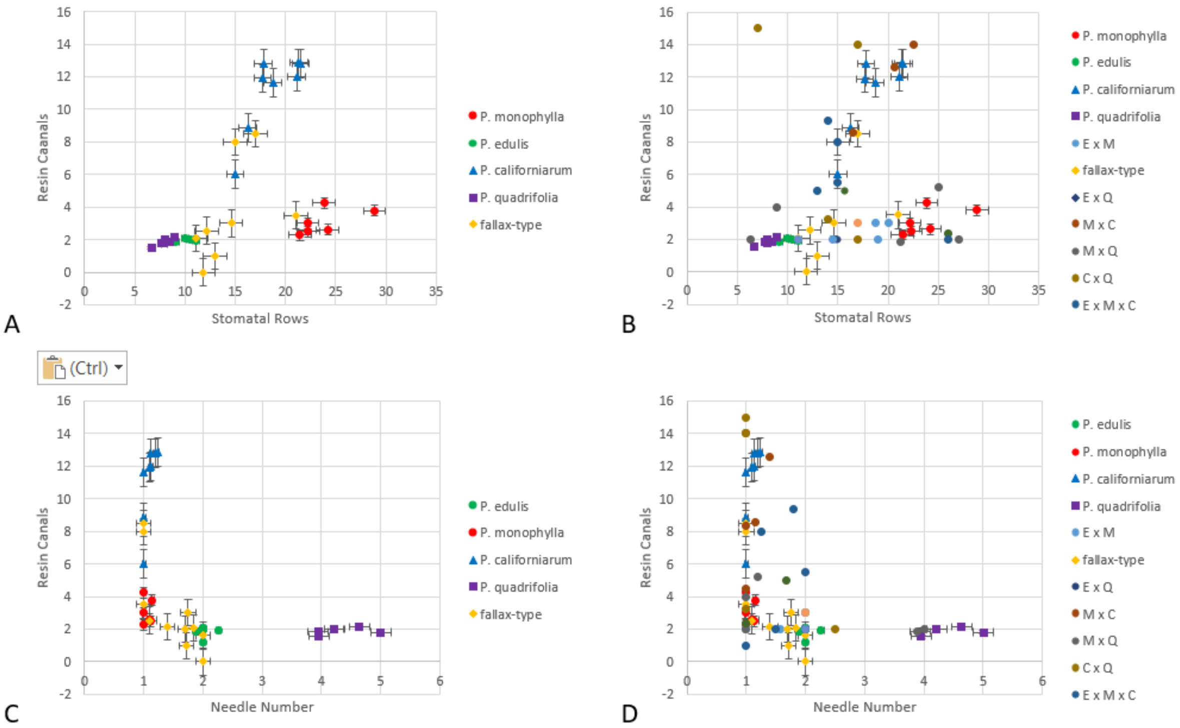
P. monophylla vs *P. edulis*

P. x fallax vs *P. quadrifolia*

P. californiarum vs *P. quadrifolia*

P. x fallax vs *P. edulis*

Supporting Information Figure 3. Niche divergence tests (identity and background), showing the D value distributions of the 100 pseudoreplicates with the overlap test D indicated by a red arrow. If the overlap test produces a smaller D value than the pseudoreplicate distribution, then niche divergence is inferred.



Supporting Information Figure 4. Scatter plots of morphology with and without admixed individuals. Morphological variation of stomatal rows by resin canals without (A) and with genetic hybrids (B). Morphological variation of needle number by resin canals without (C) and with genetic hybrids (D). Each figure with associated error bars represents a population of *P. monophylla* (red), *P. edulis* (green), *P. californiarum* (blue), *P. quadrifolia* (purple) and *fallax*-type (yellow). Hybrid individuals are represented as *P. californiarum* x *P. monophylla* (CM), *P. californiarum* x *P. quadrifolia* (CQ), *P. edulis* x *P. californiarum* (EC), *P. edulis* x *P. californiarum* x *P. quadrifolia* (ECQ), *P. edulis* x *P. monophylla* (EM), *P. edulis* x *P. monophylla x P. quadrifolia* (EMQ), *P. edulis* x *P. quadrifolia* (EQ), *P. quadrifolia* x *P. monophylla* (QM).

References

Abbott, R., Albach, D., Ansell, S., Arntzen, J.W., Baird, S.J., Bierne, N., Boughman, J., Brelsford, A., Buerkle, C.A., Buggs, R. and Butlin, R.K., 2013. Hybridization and speciation. *Journal of evolutionary biology*, *26*(2), pp.229-246.

Boecklen, W.J., 2017. Topology of syngameons. *Ecology and evolution*, *7*(24), pp.10486-10491.

Buck R, Flores-Rentería L. 2022. The Syngameon Enigma. *Plants*, *11*(7), p. 895.

Buck R, Hyasat S, Hossfeld A, Flores-Rentería L. 2020. Patterns of hybridization and cryptic introgression among one-and four-needled pinyon pines. *Annals of botany*, *126*(3), pp.401-11.

Businský, R., 2008. Výzkumný ústav SILVA TAROUCY pro krajinu a okrasné zahradnictví, vvi.

Cannon, C.H., 2021. Is speciation an inexorable march to reproductive isolation?. *Molecular Ecology, 30,* pp.4349-4352.

Cannon, C.H. and Lerdau, M., 2015. Variable mating behaviors and the maintenance of tropical biodiversity. *Frontiers in Genetics, 6,* p.183.

Cannon, C.H. and Petit, R.J., 2020. The oak syngameon: more than the sum of its parts. *New Phytologist*, *226*(4), pp.978-983.

Chhatre, V.E., Evans, L.M., DiFazio, S.P. and Keller, S.R., 2018. Adaptive introgression and maintenance of a trispecies hybrid complex in range-edge populations of *Populus*. *Molecular ecology*, *27*(23), pp.4820-4838.

Cobb, N.S., Trotter, R.T. and Whitham, T.G., 2002. Long-term sexual allocation in herbivore resistant and susceptible pinyon pine (*Pinus edulis*). *Oecologia*, *130*(1), pp.78-87.

Cole KL, Fisher J, Arundel ST, Cannella J, Swift S. 2008. Geographical and climatic limits of needle types of one-and two-needled pinyon pines. *Journal of Biogeography*, *35,* pp.257-269.

Couch, A.J., Unmack, P.J., Dyer, F.J. and Lintermans, M., 2016. Who’s your mama? Riverine hybridisation of threatened freshwater Trout Cod and Murray Cod. *PeerJ*, *4*, p.e2593.

Critchfield WB. 1975. Interspecific hybridization in Pinus: a summary review. In: Fowler DP, Yeatman CY eds. Symp. on Interspecific and Interprovenance Hybridization in Forest Trees. Proc. 14th Meeting, Canad. Tree Improv. Assoc., Part II, 99-105.

Dannemann, M. and Racimo, F., 2018. Something old, something borrowed: admixture and adaptation in human evolution. *Current opinion in genetics & development*, *53*, pp.1-8.

Dittrich-Reed, D.R. and Fitzpatrick, B.M., 2013. Transgressive hybrids as hopeful monsters. *Evolutionary biology*, *40*(2), pp.310-315.

Duran KL, Pardo A, Mitton JB. 2012. From middens to molecules: phylogeography of the piñon pine, *Pinus edulis*. Journal of biogeography *39*, pp. 1536-1544.

Elith, J., 2000. Quantitative methods for modeling species habitat: comparative performance and an application to Australian plants. In *Quantitative methods for conservation biology* (pp. 39-58). Springer, New York, NY.

Excoffier, L. and Foll, M., 2011. Fastsimcoal: a continuous-time coalescent simulator of genomic diversity under arbitrarily complex evolutionary scenarios. *Bioinformatics*, *27*(9), pp.1332-1334.

Excoffier, L., Dupanloup, I., Huerta-Sánchez, E., Sousa, V.C. and Foll, M., 2013. Robust demographic inference from genomic and SNP data. *PLoS genetics*, *9*(10), p.e1003905.

Farjon A, Styles BT. 1997. Flora Neotropica. *Pinus* (Pinaceae). New York: Organization for Flora Neotropica, New York Botanical Garden.

Fick, S.E. and Hijmans, R.J., 2017. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International journal of climatology*, *37*(12), pp.4302-4315.

Fischer, G., Nachtergaele, F., Prieler, S., Van Velthuizen, H.T., Verelst, L. and Wiberg, D., 2008. Global agro-ecological zones assessment for agriculture (GAEZ 2008). *IIASA, Laxenburg, Austria and FAO, Rome, Italy*, *10*.

Flores-Rentería, L., Wegier, A., Del Vecchyo, D.O., Ortíz-Medrano, A., Piñero, D., Whipple, A.V., Molina-Freaner, F. and Domínguez, C.A., 2013. Genetic, morphological, geographical and ecological approaches reveal phylogenetic relationships in complex groups, an example of recently diverged pinyon pine species (Subsection *Cembroides*). Molecular phylogenetics and evolution, 69(3), pp.940-949.

Flores-Rentería L, Whipple AV. 2011. A new approach to improve the scoring of mononucleotide microsatellite loci. *American journal of botany, 98*(3):e51-3.

Gehring C, Flores-Rentería D, Sthultz CM, Leonard TM, Flores-Rentería L, Whipple AV, Whitham TG. 2014. Plant genetics and interspecific competitive interactions determine ectomycorrhizal fungal community responses to climate change. *Molecular ecology, 23*(6), pp.1379-91.

Givnish, T.J., 2010. Ecology of plant speciation. *Taxon*, *59*(5), pp.1326-1366.

Gottfried, G.J., 1987. Regeneration of pinyon. In Proceedings: pinyon–juniper conference. US Department of Agriculture, Forest Service, Intermountain Research Station, Reno, Nevada, USA (pp. 249-254).

Grant, V. 1981. 18. The Syngameon. Plant Speciation. New York Chichester, West Sussex: Columbia University Press, pp. 234-241.

Grotkopp E, Rejmánek M, Sanderson MJ, Rost TL. 2004. Evolution of genome size in pines (*Pinus*) and its life-history correlates: supertree analyses. *Evolution*, *58*(8, pp.1705-29.

Hernández-Quiroz, N.S., Badano, E.I., Barragán-Torres, F., Flores, J. and Pinedo-Álvarez, C., 2018. Habitat suitability models to make conservation decisions based on areas of high species richness and endemism. *Biodiversity and Conservation*, *27*(12), pp.3185-3200.

Hollander, J.L. and Vander Wall, S.B., 2004. Effectiveness of six species of rodents as dispersers of singleleaf pinon pine (*Pinus monophylla*). *Oecologia*, *138*(1), pp.57-65.

Holmgren, C.A., Betancourt, J.L. and Rylander, K.A., 2011. Vegetation history along the eastern, desert escarpment of the Sierra San Pedro Mártir, Baja California, Mexico. *Quaternary Research*, *75*(3), pp.647-657.

Jombart, T., 2008. adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*, *24*(11), pp.1403-1405.

Kormutak, A., Galgoci, M., Sukenikova, D., Bolecek, P., Libantova, J. and Gőmőry, D., 2018. Maternal inheritance of chloroplast DNA in *Pinus mugo* Turra: A case study of *Pinus mugo* x *Pinus sylvestris* crossing. *Plant Systematics and Evolution*, *304*(1), pp.71-76.

Ladner, J.T., 2012. *Hidden diversity in corals and their endosymbionts*. Doctoral dissertation, Stanford University.

Lanner, R.M., 1974a. A new pine from Baja California and the hybrid origin of *Pinus quadrifolia*. *The Southwestern Naturalist*, pp.75-95.

Lanner RM. 1974b. Natural hybridization between *Pinus edulis* and *Pinus monophylla* in the American southwest. *Silvae Genetica 23,* pp. 108– 116.

Lanner RM, Phillips III AM. 1992. Natural hybridization and introgression of pinyon pines in northwestern Arizona. *International Journal of Plant Sciences, 153,* pp.250-257.

Linan, A.G., Lowry, P.P., Miller, A.J., Schatz, G.E., Sevathian, J.C. and Edwards, C.E., 2020. RAD‐sequencing reveals patterns of diversification and hybridization, and the accumulation of reproductive isolation in a clade of partially sympatric, tropical island trees. *Molecular Ecology, 30*(18), pp.4520-4537.

Little EL. 1968. Two new pinyon varieties from Arizona. *Phytologia, 17,* pp. 329-342.

Ledig FT (1998) Genetic variation in pines. In: Richardson DM (ed) Ecology and biogeography of *Pinus*. Cambridge University Press, Cambridge, pp.251–280

Mallet, J., 2007. Hybrid speciation. *Nature*, *446*(7133), pp.279-283.

Malusa J. 1992. Phylogeny and biogeography of the pinyon pines (*Pinus* subsect. *Cembroides*). *Systematic Botany, 17,* pp.42-66.

Mavárez, Jesús, and Mauricio Linares. 2008. "Homoploid hybrid speciation in animals." *Molecular Ecology 17*, pp.4181-4185.

Menon M, Bagley JC, Friedline CJ, Whipple AV, Schoettle AW, Leal‐Sàenz A, Wehenkel C, Molina‐Freaner F, Flores‐Rentería L, Gonzalez‐Elizondo MS, Sniezko RA. 2018. The role of hybridization during ecological divergence of southwestern white pine (*Pinus strobiformis*) and limber pine (*P. flexilis*). *Molecular Ecology,* 27(5), pp.1245-60.

Menon, M., Bagley, J.C., Page, G.F., Whipple, A.V., Schoettle, A.W., Still, C.J., Wehenkel, C., Waring, K.M., Flores-Rentería, L., Cushman, S.A. and Eckert, A.J., 2021. Adaptive evolution in a conifer hybrid zone is driven by a mosaic of recently introgressed and background genetic variants. *Communications biology, 4*(1), pp.1-14.

Merow, C., Smith, M.J. and Silander Jr, J.A., 2013. A practical guide to MaxEnt for modeling species' distributions: what it does, and why inputs and settings matter. *Ecography, 36*(10), pp.1058-1069.

Mueller, R.C., Scudder, C.M., Whitham, T.G. and Gehring, C.A., 2019. Legacy effects of tree mortality mediated by ectomycorrhizal fungal communities. *New Phytologist*, *224*(1), pp.155-165.

Montes JR, Peláez P, Willyard A, Moreno-Letelier A, Piñero D, Gernandt DS. 2019. Phylogenetics of *Pinus* Subsection *Cembroides* Engelm.(Pinaceae) Inferred from Low-Copy Nuclear Gene Sequences. *Systematic Botany, 44*, pp.501-518.

Mueller, R.C., Scudder, C.M., Porter, M.E., Talbot Trotter III, R., Gehring, C.A. and Whitham, T.G., 2005. Differential tree mortality in response to severe drought: evidence for long-term vegetation shifts. *Journal of Ecology, 93*(6), pp.1085-1093.

Neale, D.B. and Sederoff, R.R., 1989. Paternal inheritance of chloroplast DNA and maternal inheritance of mitochondrial DNA in loblolly pine. *Theoretical and Applied Genetics*, *77*(2), pp.212-216.

Nei, M., 1986. Definition and estimation of fixation indices. *Evolution*, *40*(3), pp.643-645.

Ohri, D., 2021. Polyploidy in Gymnosperms-A Reappraisal. *Silvae Genetica*, *70*(1), pp.22-38.

Patterson A, Flores-Rentería L, Whipple A, Whitham T, Gehring C. 2019. Common garden experiments disentangle plant genetic and environmental contributions to ectomycorrhizal fungal community structure. *New Phytologist*, *221*(1), pp.493-502.

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

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

Phillips, S.J., Dudik, A., and Schapire, R.E. [Internet] Maxent software for modeling species niches and distributions (Version 3.4.4). Available from url: http://biodiversityinformatics.amnh.org/open_source/maxent/. Accessed on 2021-11-03.

Racimo, F., Sankararaman, S., Nielsen, R. and Huerta-Sánchez, E., 2015. Evidence for archaic adaptive introgression in humans. *Nature Reviews Genetics*, *16*(6), pp.359-371.

Rieseberg, L.H., 1991. Hybridization in rare plants: insights from case studies in *Cercocarpus* and *Helianthus*. *Genetics and conservation of rare plants*, pp.171-181.

Rieseberg, L.H., 1997. Hybrid origins of plant species. *Annual review of Ecology and Systematics*, *28*(1), pp.359-389.

Rieseberg, L.H., Archer, M.A. and Wayne, R.K., 1999. Transgressive segregation, adaptation and speciation. *Heredity*, *83*(4), pp.363-372.

Rieseberg LH, Ellstrand NC, Arnold M. 1993. What can molecular and morphological markers tell us about plant hybridization? *Critical reviews in plant sciences, 12,* pp.213-241.

Ronco, Frank P., Jr. 1990. *Pinus edulis* Engelm. pinyon. In: Burns, Russell M.; Honkala, Barbara H., technical coordinators. Silvics of North America. Volume 1. Conifers. Agric. Handb. 654. Washington, DC: U.S. Department of Agriculture, Forest Service: 327-337.

Schliewen, U.K. and Klee, B., 2004. Reticulate sympatric speciation in Cameroonian crater lake cichlids. *Frontiers in zoology*, *1*(1), pp.1-12.

Soltis PS, Soltis DE. 2009. The role of hybridization in plant speciation. *Annual review of plant biology, 60,* pp.561-588.

Storz, Jay F., and Anthony V. Signore. 2021. "Introgressive Hybridization and Hypoxia Adaptation in High-Altitude Vertebrates." *Frontiers in Genetics, 12,* 696484.

Taylor SA, Larson EL. 2019. Insights from genomes into the evolutionary importance and prevalence of hybridization in nature. *Nature ecology & evolution*, *2,* pp.170-7.

Todesco, M., Pascual, M.A., Owens, G.L., Ostevik, K.L., Moyers, B.T., Hübner, S., Heredia, S.M., Hahn, M.A., Caseys, C., Bock, D.G. and Rieseberg, L.H., 2016. Hybridization and extinction. *Evolutionary applications*, *9*(7), pp.892-908.

USGS. 1999. *Digital representation of 'Atlas of United States Trees'.* https://archive.usgs.gov/archive/sites/www.usgs.gov/science/cite-view.php-cite=22.html. 29 November 2019.

Warren, D.L., Glor, R.E. and Turelli, M., 2008. Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. *Evolution: International Journal of Organic Evolution, 62*(11), pp.2868-2883.

Warren, D.L., Glor, R.E. and Turelli, M., 2010. ENMTools: a toolbox for comparative studies of environmental niche models. *Ecography, 33*(3), pp.607-611.

Wells PV. 1983. Paleobiogeography of montane islands in the Great Basin since the last glaciopluvial. *Ecological Monographs, 53,* pp.341-382.

Wells, P.V., 2000. Pleistocene macrofossil records of four-needled pinyon or juniper encinal in the Northern Vizcaino desert, Baja California del Norte. *Madroño,* pp.189-194.

Williams CG. 2010. Long-distance pine pollen still germinates after meso-scale dispersal. *American Journal of Botany, 97,* pp. 846-855.

Willyard, A., Syring, J., Gernandt, D.S., Liston, A. and Cronn, R., 2007. Fossil calibration of molecular divergence infers a moderate mutation rate and recent radiations for *Pinus*. *Molecular Biology and Evolution*, *24*(1), pp.90-101.

Zavarin, E., Snajberk, K. and Cool, L., 1990. Chemical differentiation in relation to the morphology of the single-needle pinyons. *Biochemical Systematics and Ecology*, *18*(2-3), pp.125-137.

CHAPTER IV

Predicting the Future: Habitat Suitability of the Southwestern Pinyon Pines Under Climate Change

Ryan Buck and Lluvia Flores-Rentería

Abstract

Pinyon pines (*Pinus* subsect. *Cembroides*) are economically, culturally, and ecologically important tree species that have been experiencing high rates of mortality as a result of climate change. Understanding their response to future climate scenarios can aid in conservation planning and act as a "canary in the coal mine" for other drought susceptible species in the Southwest. Habitat suitability models were created in Maxent for the five Southwestern pinyon pine taxa based on current and future climate scenarios. Overall, habitat suitability shifted up in elevation and latitude, with severe contractions in worst case climate scenarios. Precipitation, soil type, and temperature were the most important environmental factors limiting current distributions. The alarmingly rapid rate of projected habitat loss modeled here should serve as a warning for land managers and calls for official conservation considerations, as the loss of such ecologically important species would be devastating to the entire Southwest.

Introduction

Anthropogenic activities are changing the global climate. Increasing temperatures and altered precipitation patterns are extreme stressors on already declining populations and are causing most species' ranges to increase in elevation and move poleward

(Parmesan 2006; Young *et al.*, 2016). Recent droughts have resulted in mass mortality events in many forests around the world (Allen *et al.*, 2010; Wang *et al.*, 2012; Senf *et al.*, 2020; Hartmann *et al.*, 2022). Future droughts are predicted to increase in frequency, intensity, and length (Burke *et al.*, 2006; IPCC 2007; Mukherjee *et al*., 2018), which will likely result in more mass mortality events. These die offs usually result in vegetation shifts that alter ecosystems and can further exacerbate climatic effects (McDowell and Allen, 2015; Batllori *et al.*, 2020). Understanding and modeling these future vegetation shifts can allow managers to better prepare for future ecosystem changes and effectively distribute conservation resources (Araujo and Rahbek, 2006; Hijmans *et al.*, 2006; Schwartz *et al.*, 2012; Zhu *et al.*, 2013).

Climate change in the Southwestern US has been particularly severe, with intense droughts increasingly triggering mass forest mortality events (Van Mantgem *et al.*, 2009; Jiang *et al.*, 2013). One dominant tree type, the pinyon pine, has experienced massive rates of mortality in the past two decades due to drought (Ogle *et al.*, 2000; Mueller *et al.*, 2005; Flake and Weisberg, 2018). Pinyons are foundation species in that they support thousands of other species above and below ground (Little 1943; Wells 1983; Severson 1986; Chambers *et al.*, 1999; Paulin *et al.*, 1999; Sthultz *et al.*, 2006; Trotter *et al.*, 2008; Gehring *et al.*, 2014; Patterson *et al.*, 2018; Stone *et al.*, 2018), so their mortality has profound, cascading ecological consequences (Gitlin *et al*., 2006). Pinyon pines are also of cultural significance, as Natives have used them for thousands of years for food, shelter, and in ceremonies (Castetter *et al.*, 1935; Elmore 1944; Lanner 1984; Miller 1995, 1997). Pinyon products are economically important to many areas in the Southwest, with some sites capable of producing up to $400,000 worth of timber (Schwab 1993; Miller 1997) and 300 pounds of pine nuts per acre (Rocco 1990; Jeffers 1995) per year. For these reasons, among others, protecting pinyon woodlands should be of utmost concern to land managers and governing bodies.

One species of pinyon pine, *P. edulis*, has been described as a barometer of climate change (Gitlin *et al.*, 2006), meaning that its response to environmental changes can help determine how other species may respond. Modeling its future distribution and examining its reaction to climate change can help managers understand how future climate will drive ecological changes and serve as a gauge for the health of the entire Southwest. Previous studies have attempted to model the future distribution of *P. edulis* using climate envelopes by incorporating distribution limitations, previous migration rates, and climatic models (Cole *et al.*, 2008a). Even with a generous 100m/yr migration rate, their models predicted an approximately 75% decline in *P. edulis* distribution by 2100, with very little expansion. Unfortunately some of the areas that *P. edulis* was predicted to expand into have since experienced mass mortality events during the early 2000's drought that devastated Southwest woodlands (Breshears *et al.*, 2005; Cole *et al.*, 2008; Peterman *et al.*, 2012). Cole *et al.* (2008a) attribute this discrepancy to either the model's predicted increase of monsoon precipitation where there was actually decreased precipitation, or to an overestimation of the parameter's importance. Additionally, more sophisticated computational models that incorporate newer, high-resolution climate variables have since been developed. Further, only the future range of *P. edulis* has been modeled, but the additional pinyon pine taxa identified in this region (*P. monophylla*, *P. californiarum*, *P.*

quadrifolia, and *P.* x *fallax* - Cole *et al.*, 2008b; Buck *et al.*, 2020; Buck *et al.*, in review) participate with *P. edulis* in a complex multispecies network called a syngameon (Buck and Flores-Rentería, 2022). Modeling all species within the syngameon is essential not only to gain an accurate understanding of all the Southwestern pinyon pines' responses to climate change but also to determine the resilience of the entire complex.

Our goals are to 1) estimate the future habitat suitability of the five pinyon pine taxa and 2) understand the environmental variables that limit current distribution and could drive future expansions and contractions.

Methods

Known species occurrences were taken from Buck *et al.* (in review), in which identities of populations were genetically confirmed. Additionally, sample sites from Zavarin *et al.* (1990), as well as citizen science and herbaria data from the Consortium of California Herbaria (CCH2), Baja Flora, Global Biodiversity Information Facility (GBIF), and la Red Mundial de Información sobre Biodiversidad (REMIB) were included. Occurrence points were hand filtered to remove any erroneous points and correct any misidentified taxa based on maps generated by USGS (1999) and Cole *et al.* (2003) and genetic data from Buck *et al.* (2020; in review). This resulted in a final observation count of 1449 for *P. monophylla*, 1353 for *P. edulis*, 1343 for *P. californiarum*, 745 for *P. quadrifolia*, and 172 for *fallax*-type (Fig. 1).

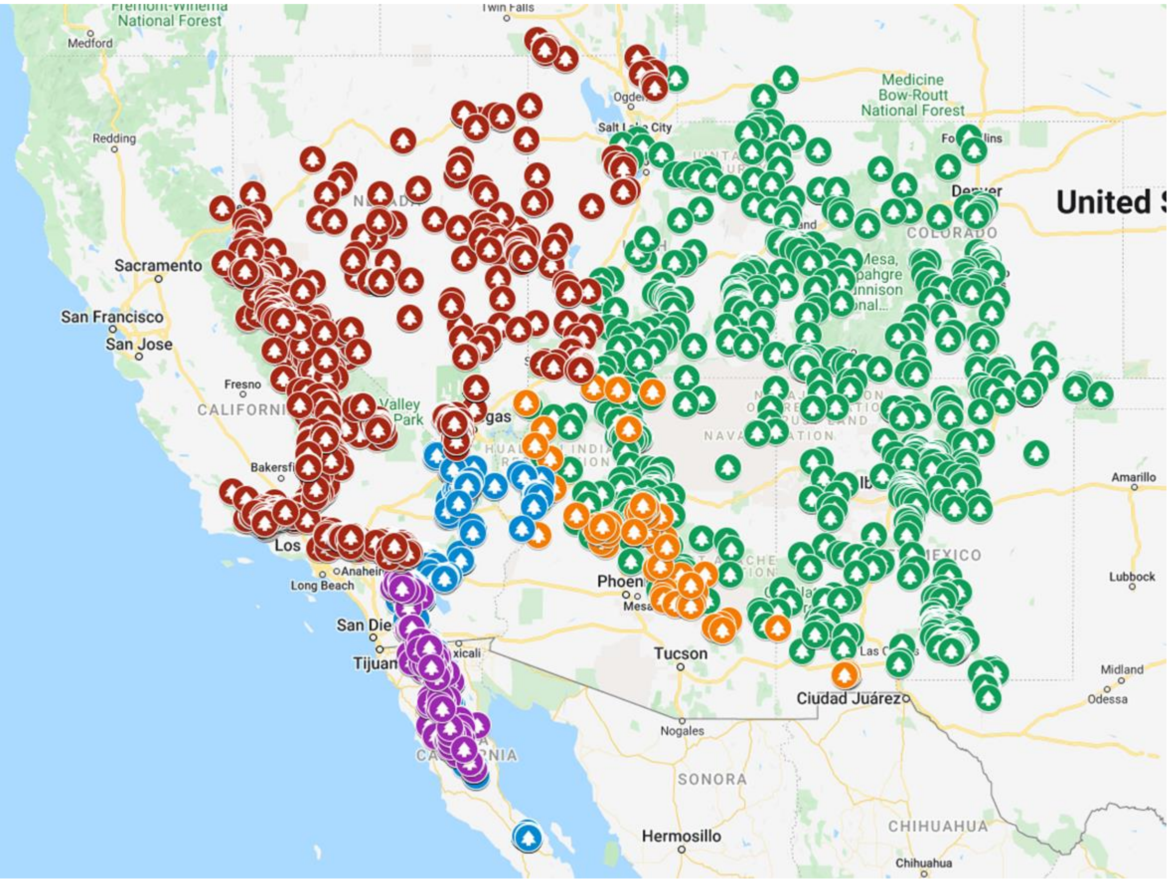


Figure 1. Occurrence points input into Maxent colored by taxon identity (*P. monophylla* = red, *P. edulis* = green, *P. californiarum* = blue, *P. quadrifolia* = purple, *P.* x *fallax* = orange).

To model current habitat suitability of the taxa in this study, environmental niche models were generated using MaxEnt v3.4.4 (Phillips *et al.*, 2006). The original 30 arc-second variables used to build the suitability models were 19 current bioclimatic variables from WorldClim1.4 (Hijmans *et al.*, 2005), and 36 categorical soil classes, 8 slope classes, and 5 slope aspect classes from the Harmonized World Soil Database (Fischer *et al.*, 2008).

To model future habitat suitability, the variables used in future projections were the same soil variables used in the current habitat suitability models, but with 19 future bioclimatic variables from WorldClim1.4 from two future time periods (2050 and 2070), three IPCC5 Representative Climate Pathways (RCPs 4.5, 6.0, 8.5), and two separate global climate projection models (BCC-CSM1-1 and IPSL-CM5A-LR) from different families to reduce interdependence (Sanderson *et al.*, 2015; Gunther 2021), for a total of 11 future projections per species (RCP6.0-2050 was not available for BCC-CSM1-1).

All rasters were cropped to latitudes 23°N and 45°N and longitudes 99°W and 125°W to encompass the study area and allow for realistic migration. Correlated variables with a Spearman's rank correlation ($|r| > 0.70$) (Merow *et al.*, 2013) and those with 0% contribution to the model were removed. Independent models were created for each species by using 10 bootstrapped replicates, jackknifing, a random test percentage of 30% (Hernández-Quiroz *et al.*, 2018), a maximum of 5000 iterations, a 1.5 regularization multiplier (Merow *et al.*, 2013), and a 10 percentile training presence. Output rasters were visualized in ArcGIS Pro2.8 using the calculated 10 percentile training presence cloglog thresholds as limits for suitable habitat. The jackknife test results were used to determine which variables contributed the most to species distributions, while the response curves were used to examine the range of those variables with the same cloglog thresholds used in suitability visualization.

Results

Models performed better than random, with AUC training and test values above 0.9 for each model. Current suitable habitat (Fig. 2) roughly matched the estimated species distribution maps generated by Cole *et al.* (2008).

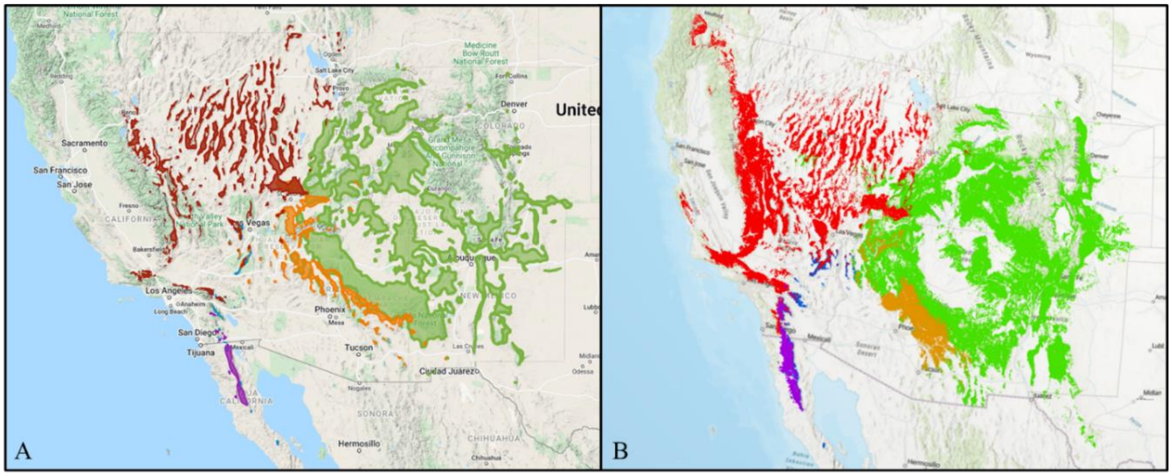


Figure 2. A) Current actual approximated ranges of pinyon species colored by taxon identity (*P. monophylla* = red, *P. edulis* = green, *P. californiarum* = blue, *P. quadrifolia* = purple, *P.* x *fallax* = orange) (based on USGS 1999; Cole *et al*., 2003). B) Current habitat suitability estimated by Maxent, visualized in ArcGIS Pro, and colored by taxon identity (*P. monophylla* = red, *P. edulis* = green, *P. californiarum* = blue, *P. quadrifolia* = purple, *P.* x *fallax* = orange).

Variable importance

Maxent model's jackknife tests determined which variables contributed most to the species distributions (Suppl. Table 1). The response curves for the top three variables for each species (Suppl. Figs. 1 - 5) identified the suitable ranges of each variable within the cloglog thresholds (*P. monophylla* > 0.34, *P. edulis* > 0.35, *P. californiarum* > 0.51, *P. quadrifolia* > 0.55, and *P.* x *fallax* > 0.36).

The most important variables contributing to *Pinus monophylla*'s habitat suitability were the mean temperature of the wettest quarter with a range of -5°C to 8°C, Luvisol and Phaeozem soil types, and the mean temperature of the coldest quarter with a range of -5°C to 7°C (Suppl. Table 1; Suppl. Fig. 1). *Pinus edulis*' habitat suitability was influenced most by Calcisol, Kastanozem, and Phaeozem soil types, the precipitation of the driest month with a range of 7.5mm to 33mm, and the mean temperature of the wettest quarter with a range of 11.5°C to 22.5°C (Suppl. Table 1; Suppl. Fig. 2). The most important variables contributing to *P. californiarum*'s habitat suitability were Luvisol soil types, the minimum temperature of the coldest month with a range of -2.5°C to 2°C, and the mean temperature of the wettest quarter with a range of -4°C to 9°C (Suppl. Table 1; Suppl. Fig. 3). *Pinus quadrifolia*'s habitat suitability was most affected by the precipitation of the driest month with a range of 0mm to 3mm, the precipitation of the coldest quarter with a range of 125mm to 400mm, and the mean temperature of the wettest quarter with a range of 5°C to 8°C (Suppl. Table 1; Suppl. Fig. 4). The most important variables contributing to *P.* x *fallax*'s habitat suitability were Kastanozem and Phaeozem soil types, the precipitation of the coldest quarter with a range of 100mm to 300mm, and the minimum temperature of the coldest month with a range of -7°C to 1°C (Suppl. Table 1; Suppl. Fig. 5).

Future Habitat Suitability

RCP 4.5 (Fig. 3)

In both global climate projection models, *P. edulis* begins to lose the lower elevation portions of its southern range by 2050, with some northeastwards expansion into the lower Rocky Mountains (Fig, 3A,C). The BCC models predict *P. monophylla* will experience a range contraction in the lower elevation portions of its Great Basin distribution and its eastern Sierra Nevada distribution (Fig 3A-B), whereas the IPSL models predict the entire lower half of Great Basin will be unsuitable for *P. monophylla* (Fig. 3C-D). The IPSL models predict a heavy expansion into the Salt Lake City, Utah area bordering the Rocky Mountains. Both models show *P. californiarum* beginning to lose the southern edge of its habitat in Baja California, while simultaneously losing its eastern Mojave Desert range. Interestingly, the IPSL model shows a loss of suitable habitat in Southern Nevada in 2050 but a regain of suitability in 2070. Similarly, the BCC model shows a more extensive loss of *P. californiarum*’s Baja California range in 2050 than in 2070. Some northwest expansions are predicted above the Mojave Desert into the eastern Sierra Nevadas. *Pinus quadrifolia* is predicted to become quite restricted, losing almost all its US distribution, moving up in elevation and existing only on higher elevation mountains. Lastly, models anticipate *P.* x *fallax* contracting in southern Arizona but beginning to expand northward into the Utah-Arizona border, filling habitat currently occupied by *P. edulis*.

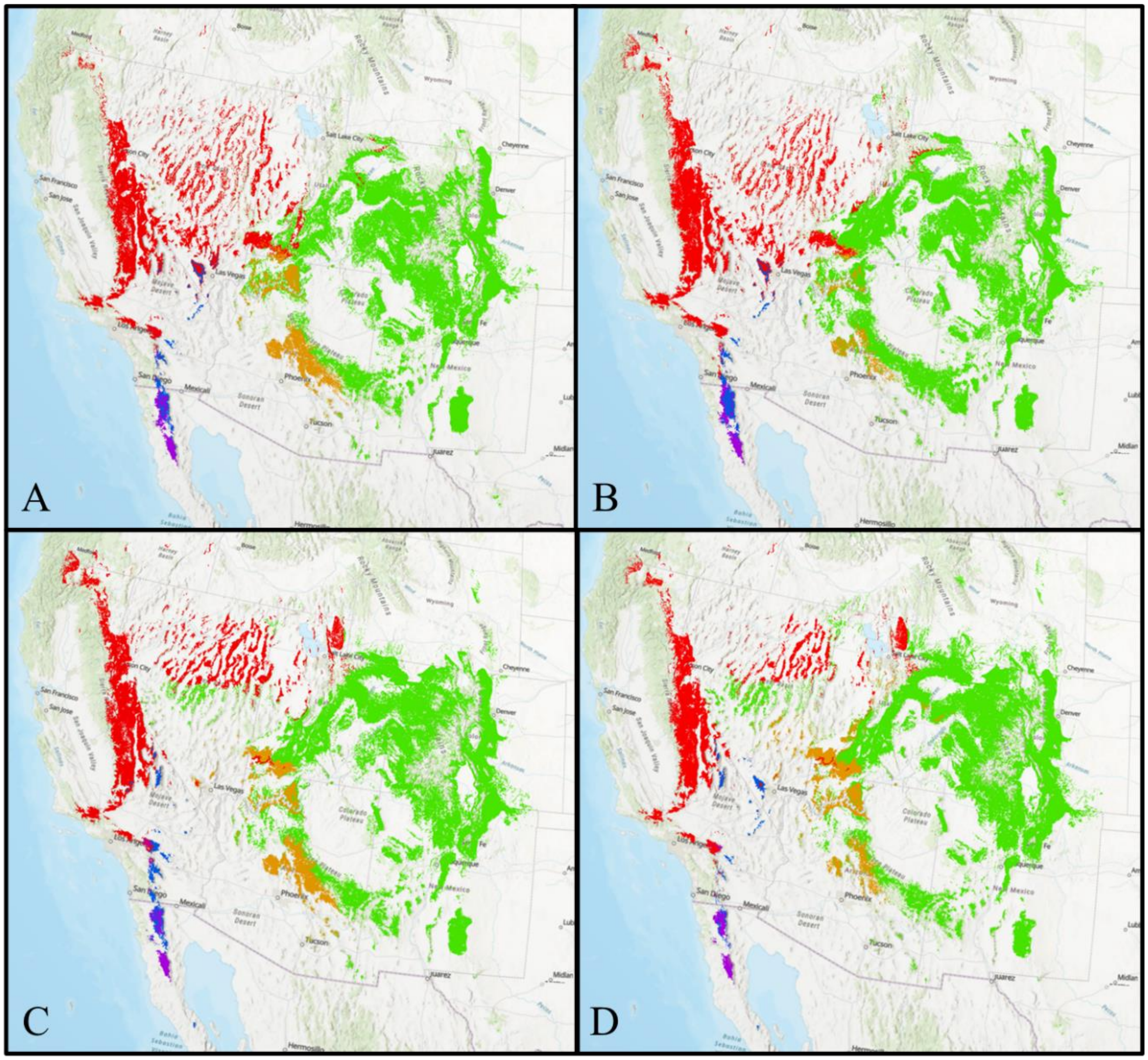


Figure 3. Future habitat suitability of *P. monophylla* (red), *P. edulis* (green), *P. californiarum* (blue), *P. quadrifolia* (purple), and *P.* x *fallax* (orange) based on BCC - RCP 4.5 - 2050 (A) and 2070 (B) and IPSL - RCP 4.5 - 2050 (C) and 2070 (D).

RCP 6.0 (Fig. 4)

Suitability models predict *P. edulis* contracting in its southern range, similar to the RCP 4.5 models, but with more suitable habitat north-east into Wyoming. The IPSL models indicate more suitable habitat in 2070 (Fig. 4D) compared to 2050 estimates (Fig. 4C). *Pinus monophylla* is modeled to lose either a western (Fig. 4A-B) or southern (Fig. 4C-D) portion of its Great Basin habitat, with minor expansion into the Salt Lake City, Utah area bordering the Rocky Mountains. There is an initial disappearance of *P. californiarum*'s eastern Mojave Desert and southern Baja California ranges in 2050, with a slight resurgence in the south and a near-current recovery in the east by 2070. Additionally, the IPSL models predict northwest expansions above the Mojave Desert into the eastern Sierra Nevadas not shown in the BCC model. In 2050, *P. quadrifolia* is modeled to lose lower elevation populations, contracting to two mountain ranges in Baja California, whereas the 2070 models predict less of an elevation gain. *Pinus* x *fallax* is again estimated to lose suitable habitat in the southern extent of its Arizona range, but with a larger increase in suitable habitat on its northern edge.

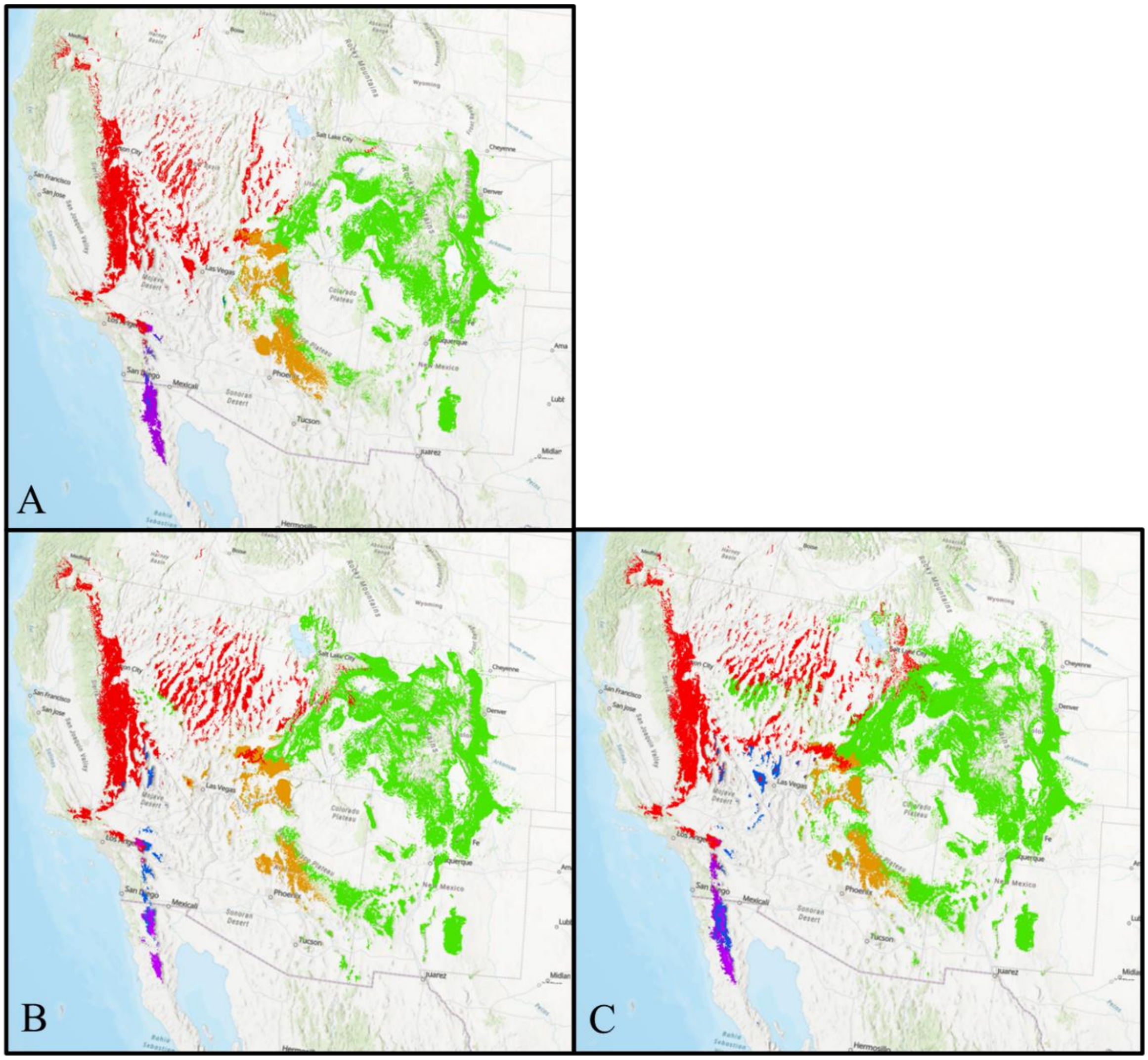


Figure 4. Future habitat suitability of *P. monophylla* (red), *P. edulis* (green), *P. californiarum* (blue), *P. quadrifolia* (purple), and *P.* x *fallax* (orange) based on BCC - RCP 6.0 - 2070 (A) and IPSL - RCP 6.0 - 2050 (C) and 2070 (D).

RCP 8.5 (Fig. 5)

The most drastic climate model shows the most loss in habitat suitability, with *P. edulis* losing almost its entire Colorado Plateau habitat. Northeast expansions into the Rocky mountains are not as large as the other RCPs, with the BCC (Fig. 5A-B) and IPSL (Fig. 5C-D) models differing in extent of expansion. The BCC models show *P. monophylla* losing a majority of its northern Great Basin range in 2050 (Fig. 5A), regaining the middle portion of it in 2070 (Fig. 5B), but still showing an overall decline in suitable habitat. The IPSL model predicts a similar decline in 2050 but in the southern half of the Great Basin, followed by an extreme loss of suitable habitat for *P. monophylla* across the entire Great Basin and most of its northern and southern Sierra Nevada ranges. *Pinus californiarum* loses almost all of its range in the BCC models, with small disconnected habitats excluded from the Mojave Desert. The IPSL models predict less contraction, with an initial decline in the Mojave Desert in 2050 (Fig. 5C), followed by an entire disappearance of suitable habitat from Southern California and Baja California with habitat suitability switching to the Mojave desert (Fig. 5D). Both models show a sharp contraction of *P. quadrifolia*, with the BCC models restricting its range to the northern mountains of Baja California and the IPSL model relegating it to two isolated mountain tops. *Pinus* x *fallax* shows similar patterns to the other RCPs with a loss in southern range and increase in northern range, but not as substantial as the other species. Some expansion of habitat suitability is predicted into the Mojave Desert where *P. californiarum* populations currently are.

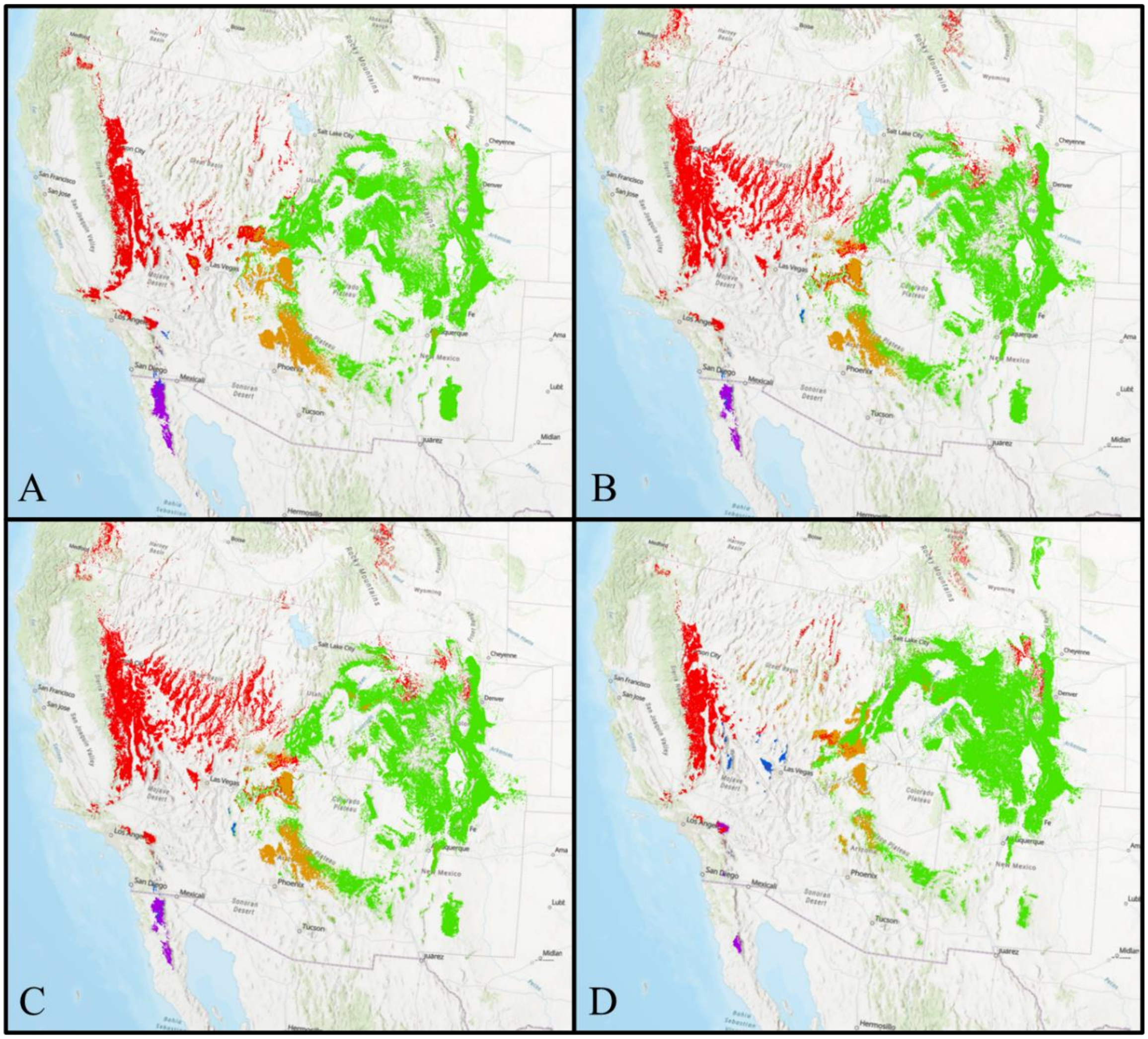


Figure 5. Future habitat suitability of *P. monophylla* (red), *P. edulis* (green), *P. californiarum* (blue), *P. quadrifolia* (purple), and *P.* x *fallax* (orange) based on BCC - RCP 8.5 - 2050 (A) and 2070 (B) and IPSL - RCP 8.5 - 2050 (C) and 2070 (D).

Discussion

Future Climate Projections

Modeling future suitable habitat of large ecosystem drivers like pinyon pines can be useful for land managers and conservation planners. The results of those models can inform policy and resource distribution by indicating areas of potential range expansion and contraction. In the Southwest pinyon pine complex, the five taxa modeled showed an overall suitable habitat shift upwards in elevation, matching projections of other species migrations under climate change (Parmesan 2006; Young *et al.*, 2016), but only some species were predicted to gain northward suitable habitat (Figs. 3-5). As expected, the habitat of most taxa contracted as more drastic RCP climate models were used. It is important to note that these model predictions are only based on abiotic factors like climate and soil variables and do not include biotic factors such as dispersal capabilities, life history traits, or hybridization; therefore these represent the potential fundamental niches, not the realized ones (Kraft *et al.*, 2015). In reality, the distributions of these foundation species are likely dependent on the complex interactions among ecosystem participants and between abiotic and biotic feedback systems. Future studies should combine suitability with dispersal capabilities, migration rates, fire history, soil microbe data, and specifically incorporate hybridization to model the full extent of the pinyon syngameon.

Interestingly, some species had more projected suitable habitat in 2070 than in 2050 (e.g. *P. quadrifolia* and *P. edulis* in Fig. 4C-D, *P. monophylla* in Fig. 5A-B), suggesting that combinations of future climate may vary in consequences, temporarily benefiting some species while harming others. Similar patterns are seen in *P. californiarum* where one half of its range is projected to be unsuitable in 2050 while the other half remains suitable, but then switches in 2070 where the previously unsuitable range becomes suitable again and vice versa (Figs. 3C-D, 4C-D, and 5C-D). At first glance, these patterns may seem hopeful because at least at some point in the future its range will collectively be suitable; however, *P. californiarum* may not survive the 20 years of unsuitability in between the two projections. For example, the IPSL RCP 8.5 model predicts the southern half of its range will be suitable in 2050 (Fig. 5C), while the northern half will be unsuitable. Then in 2070, the suitability switches, with the northern half being suitable and the southern half unsuitable (Fig. 5D). In the time the northern half of *P. californiarum*'s range is unsuitable, all of those populations could die off and not be able to recolonize when the habitat becomes suitable again in 2070. Collecting and storing seeds for future replanting could aid in managed recolonization efforts.

While the patterns of *P. monophylla* contraction in the Great Basin were inconsistent (e.g. Fig. 3), a general pattern of distribution change was seen, with the possibility of a 50-90% habitat suitability decrease. Great Basin land managers should begin measuring mortality events, especially after periods of warm winters, as temperature during the coldest and wettest quarter are two of the largest variables contributing to *P. monophylla* suitability (Suppl. Table 1, Suppl. Fig. 1).

Suitable habitat contraction of *P. edulis* was not as extreme as predicted by Cole *et al.* (2008a), however their models incorporated dispersal distances and historical migration rates which are not accounted for here. Further, the recent mortality events were not directly modeled in these analyses but future temperature and precipitation patterns that would cause range contraction are included. Wide-scale, multispecies, drought-induced mortality data is needed before incorporating relevant habitat loss into these niche models

Pinus californiarum already inhabits hot and arid regions (Cole *et al.*, 2008b), living at an extreme of its habitat. These models showed an alarming amount of predicted habitat loss, especially considering its already restricted range and drought tolerance abilities. Since species at the extremes of their habitats are expected to be most at risk of extinction (Hantson *et al.*, 2021), future research should assess the conservation status of *P. californiarum* and provide guidance to land managers.

The consistent pattern seen for *P.* x *fallax* was suitable range contraction on the southern edge of the Mogollon Rim, with some expansion in the northern part of its range near the Arizona-Utah border, filling current *P. edulis* habitat. The amount of habitat loss for *P.* x *fallax* is fairly concerning as it also has drought tolerant traits and lives at the extreme of its habitat (Cole *et al.*, 2008b). Further, *P.* x *fallax* is of hybrid origin from *P. edulis* and *P. californiarum* (Buck *et al.*, in review), but those two species are predicted to contract away from each other, especially in the IPSL 2050 models (Figs. 3C-D, 4C-D, and 5C-D). This loss of sympatry could prevent the creation of new F_1 hybrids and potentially end the *P.* x *fallax* lineage. However, *P.* x *fallax* may be an incipient species and its populations are mostly composed of advanced generation backcrosses (Buck *et al.*, in review) so loss of gene flow with *P. californiarum* may actually drive eventual speciation, an event not likely to happen within 50 years. If *P. californiarum* and *P. edulis* do continue to exist in sympatry (Figs. 3B, 4A,C, 5B), their hybridization could facilitate the expansion of *P.* x *fallax*. Conversely, constant introgression with *P. edulis* could result in the swamping of *P.* x *fallax*'s genome (Todesco *et al.*, 2016).

Variables important to distribution

Soil types were consistently in the top three most important variables for all taxa, with the exception of *P. quadrifolia* whose soil variables were removed from its models due to the abnormally high contribution percentage, likely a result of data discrepancies between the US and Mexico. The most important soil types found (Luvisol, Phaeozem, Calcisol, and Kastanozem) are all associated with high humus and calcium contents (Food and Agriculture Organization of the United Nations). It is not clear why pinyon pines specifically would be limited by these types of soil, but calcium is an essential nutrient in plants (White and Broadley, 2003) and deficiency has been shown to be detrimental to pine growth (Davis 1949). Alternatively, the calcium composition could be an artifact of the more important water retention abilities of the soil types, with higher pinyon mortality observed in lower water capacity soils (Peterman *et al.*, 2012). Soil is often overlooked in restoration projects and niche models (Valezco *et al.*, 2017; Zuquim *et al.*, 2019), but land managers and researchers alike should consider soil types when making restoration decisions and modeling habitat suitability.

Pinus edulis currently inhabits areas that experience monsoon precipitation (Cole *et al.*, 2008) so it is not surprising that its suitability was influenced by the precipitation of the driest month and the mean temperature of the wettest quarter (Suppl. Table 1; Suppl. Fig. 2), as it is restricted by drier winters and hotter summers to the south in *P.* x *fallax* territory. Interestingly, *P.* x *fallax* and *P. edulis* share soil preferences, so the amount of winter precipitation and the colder temperatures to the north-east likely restrict its advance. Conversely, *P. monophylla* inhabits areas that experience Mediterranean precipitation (Cole *et al.*, 2008), matching its contributing variables of mean temperature of the wettest quarter and mean temperature of the coldest quarter. *Pinus monophylla* is likely restricted to the south by hotter and drier winters, habitats in which *P. californiarum* prefers. *Pinus quadrifolia*'s suitable habitat is mostly affected by precipitation in both the driest month and coldest quarter (Suppl. Table 1, Suppl. Fig. 4), likely preventing its expansion further north and explaining its future suitable habitat's increase in elevation.

Conclusions

Some pinyon pine species are projected to lose immense amounts of suitable habitat within the next 50 years. Land managers and conservationists must begin assessing the rapid decline of these species and making active efforts to increase woodland cover. Pinyon pines are long-lived, slowly reproducing species (Ronco 1990; Weisberg and Ko, 2012; Redmond *et al.*, 2012), so the ecosystem function of old-growth pinyon stands will likely not be regained through planting seedlings alone. Further, pinyons are estimated to migrate at only 40m per year (Cole *et al.*, 2008a), so it is not likely they will be able to reach or expand into a majority of habitable ranges by 2070 beyond 2km away. Additionally, pinyon trees can live for hundreds of years (Ronco 1990) so they could be currently existing in habitat that was suitable for them hundreds of years ago, biasing our models to a shifted time frame. Future studies will need to model environments conducive to seedling and sapling growth to fully understand suitable expansion habitats (McCallum 2011). Lastly, with predicted range contractions, the future of the pinyon syngameon is at risk. The consequences of losing a syngameon participant to extinction or loss of sympatry are not yet known; however, it could result in the collapse of the syngameon and the possible extinction of any species dependent on gene flow (Buck and Flores-Rentería, 2022). Future research should focus on the conservation status of the Southwestern pinyon pines, especially those with currently limited ranges and extreme predicted habitat loss.

Supplementary Information

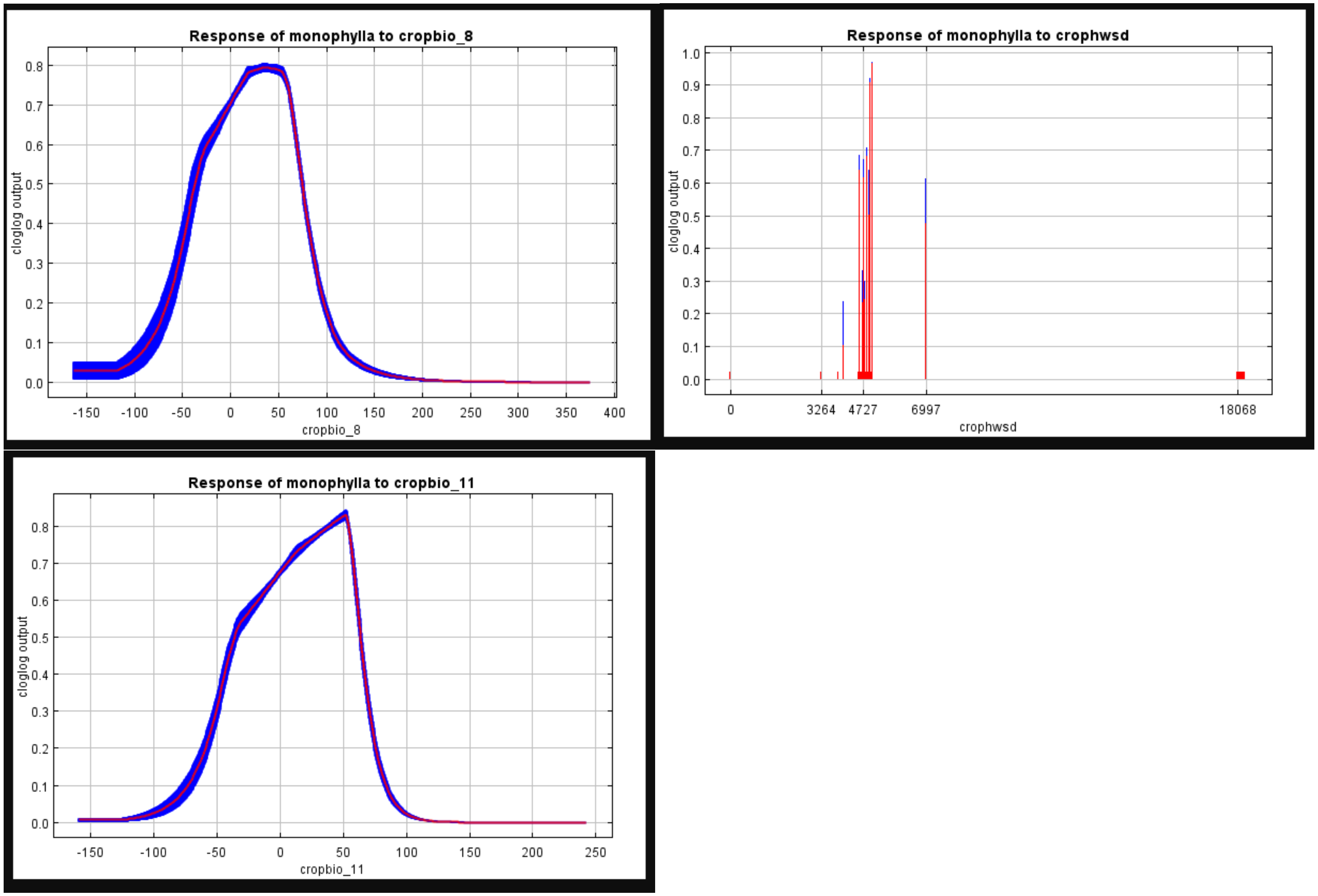
Supplementary Table 1. The contributions and permutation importance of each bioclimatic and soil variable per species. The higher contribution percentages represent higher contributions of that variable to fitting the model with all other variables present, while higher permutation importance represents a smaller drop in AUC score if models were made only using that variable. “X”s indicate the variable was removed from that species’ models due to correlation with other, higher contributing variables or a lack of contribution altogether (0%).

Symbol	Variable	*P. monophylla*		*P. edulis*		*P. californiarum*		*P. quadrifolia*		*P.* x *fallax*	
		Contribution (%)	Permutation importance	Contribution (%)	Permutation importance	Contribution (%)	Permutation importance	Contribution (%)	Permutation importance	Contribution (%)	Permutation importance
BIO2	Mean diurnal range	0.4	0.1	0.4	0.4	0.2	1.6	X	X	X	X
BIO4	Temperature seasonality	X	X	12.1	11	0.1	3.3	X	X	X	X
BIO5	Max temp of warmest month	0.8	5.6	0.6	1.8	0.9	2.2	0.2	0.1	X	X
BIO6	Min temp of coldest month	X	X	X	X	10	22.4	X	X	9.7	23.2

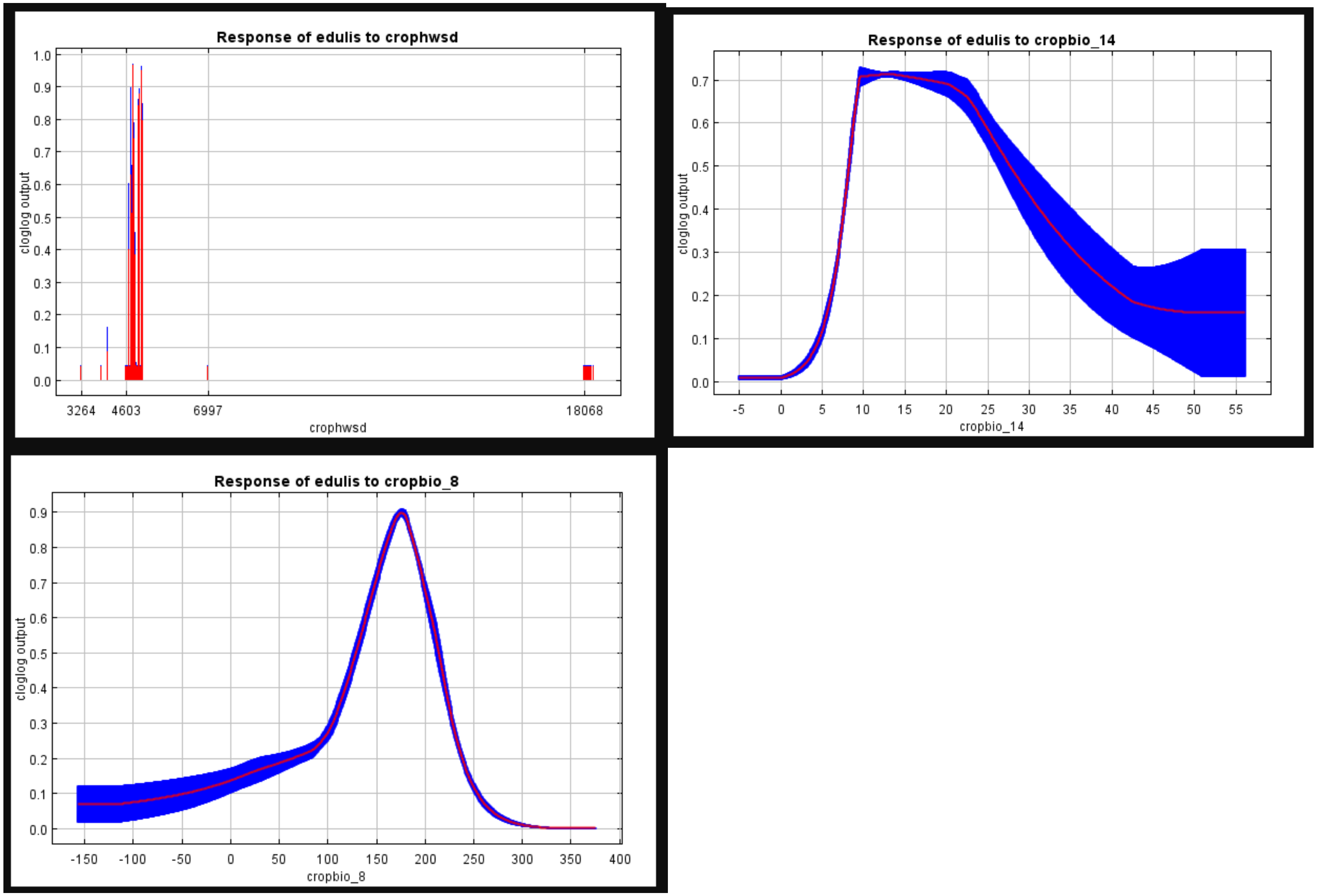
BIO8	Mean temp of wettest quarter	45.9	17.1	17.2	7	10	4	18.8	46	0.2	0.1
BIO9	Mean temp of driest quarter	0.5	2.3	0.9	1.6	X	X	X	X	X	X
BIO11	Mean temp of coldest quarter	8.4	34.5	X	X	X	X	1	1.6	X	X
BIO12	Annual precipitation	4.7	5.8	X	X	X	X	X	X	X	X
BIO13	Precipitation of wettest month	X	X	0.2	5.7	1	3.2	X	X	X	X
BIO14	Precipitation of driest month	0.3	2	19.9	4.3	0.9	35	39.5	43.2	0.9	0.7
BIO15	Precipitation seasonality	0.8	0.9	0.4	7.9	1.4	3.7	4.5	2.8	3.7	0.7

BIO16	Precipitation of wettest quarter	X	X	X	X	X	X	X	X	1.7	23
BIO18	Precipitation of warmest quarter	2	6.1	1.2	24.7	6.9	10.8	16.9	5.7	5	22.8
BIO19	Precipitation of coldest quarter	X	X	X	X	1.3	1.6	19	0.4	13.2	20.3
CI1	0-0.5% slope	X	X	9.7	1.7	X	X	X	X	X	X
CI3	2-5% slope	0.5	0.7	0.2	0.6	0.3	0.1	X	X	0.1	0
CI4	5-10% slope	0.2	0.9	1.9	0.8	X	X	X	X	0.2	0.1
CI5	10-15% slope	0.2	0.4	5.5	0.2	X	X	X	X	5.3	0.2
CI6	15-30% slope	3.6	0.2	5.5	0.9	5.7	0.1	X	X	0.4	0.2
CI7	30-45% slope	2.5	0.7	0.6	0.5	X	X	0.2	0.1	X	X

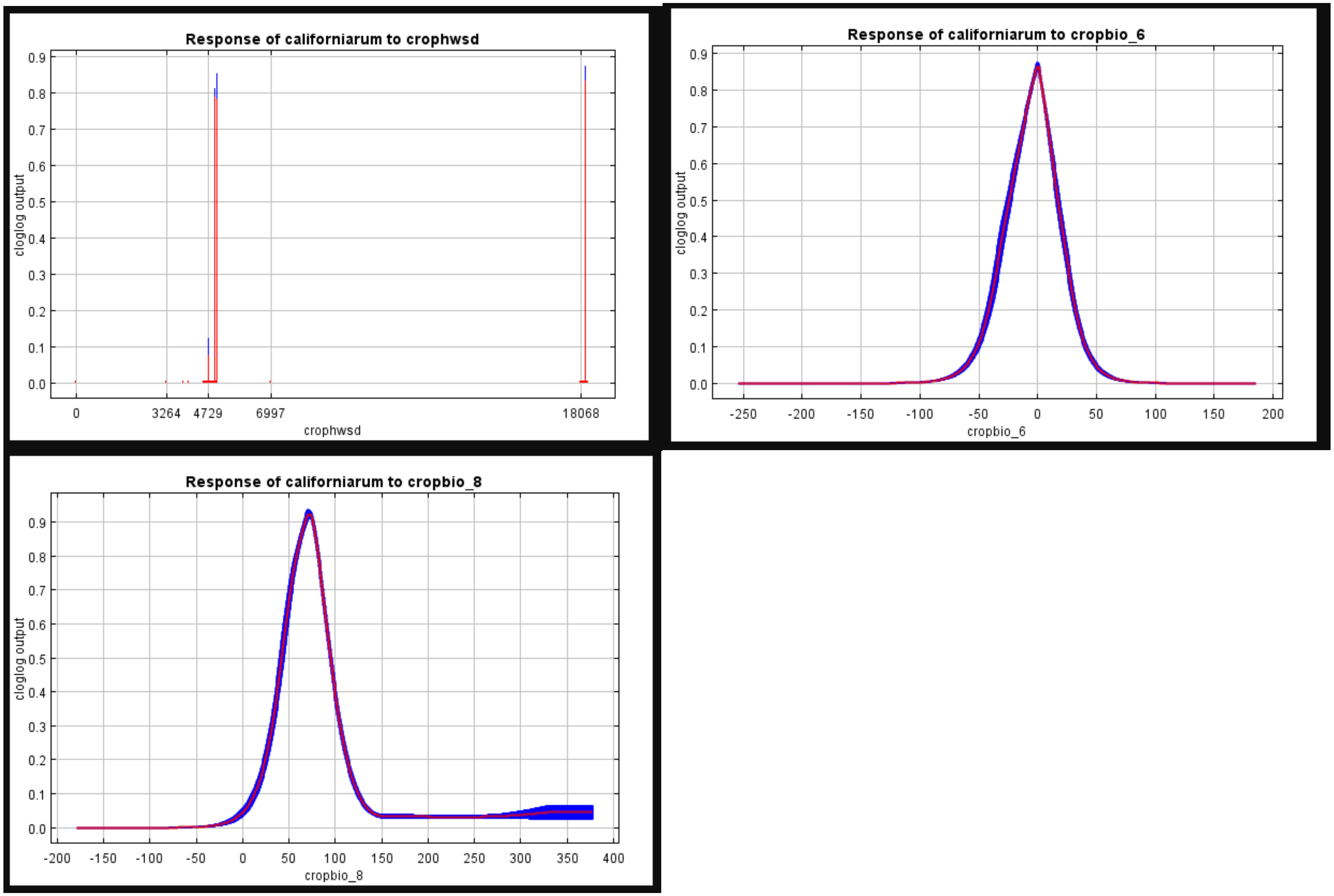
CI8	>45% slope	0.1	0.3	0.8	1	X	X	X	X	0.2	0.1
CIN	North facing slope (0°-45°)	0.2	0.4	0.1	0.7	X	X	X	X	0.2	0.1
CIE	East facing slope (46°-135°)	0.1	0.2	0.1	0.6	X	X	X	X	0.1	0.1
CIS	South facing slope (136°-225°)	0.1	0.4	0.1	0.9	X	X	X	X	0.4	0.1
CIW	West facing slope (226°-315°)	0.4	0.3	0.3	1.2	X	X	X	X	X	X
CIU	Undefined or less than 2% grade	7.2	3	X	X	3.5	2.4	X	X	3.6	7.6
HWSD	Soil category	21	18.1	22.3	26.5	57.8	9.6	X	X	54.9	0.8



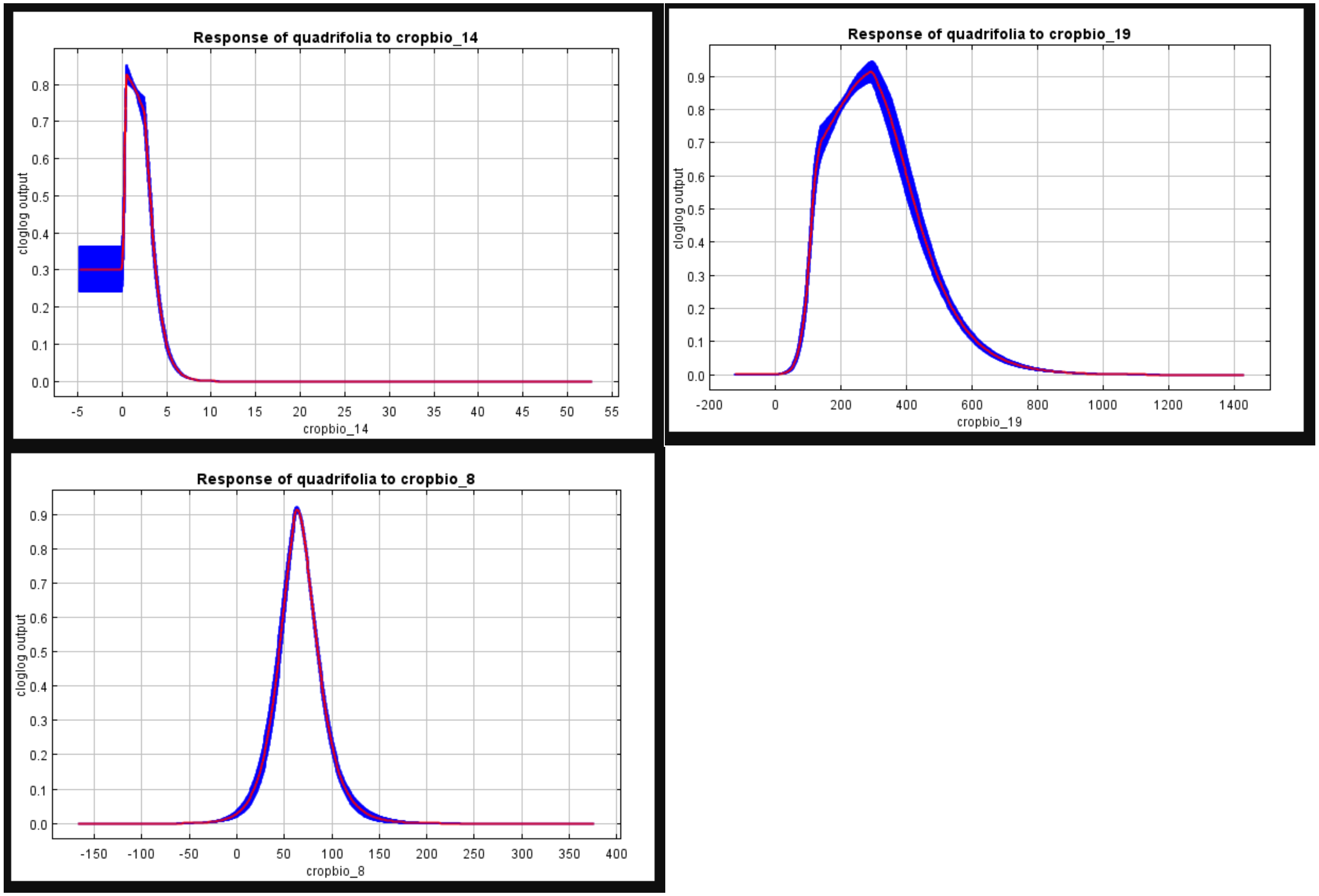
Supplementary Figure 1. Response curves for the top three contributing variables to *P. monophylla*'s suitable habitat (bio8 = mean temperature of wettest quarter, hwsd = soil type, bio11 = mean temperature of coldest quarter). Variable ranges were evaluated with a cloglog threshold > 0.34 (y-axis). Temperature units are x 10°C (x-axis).



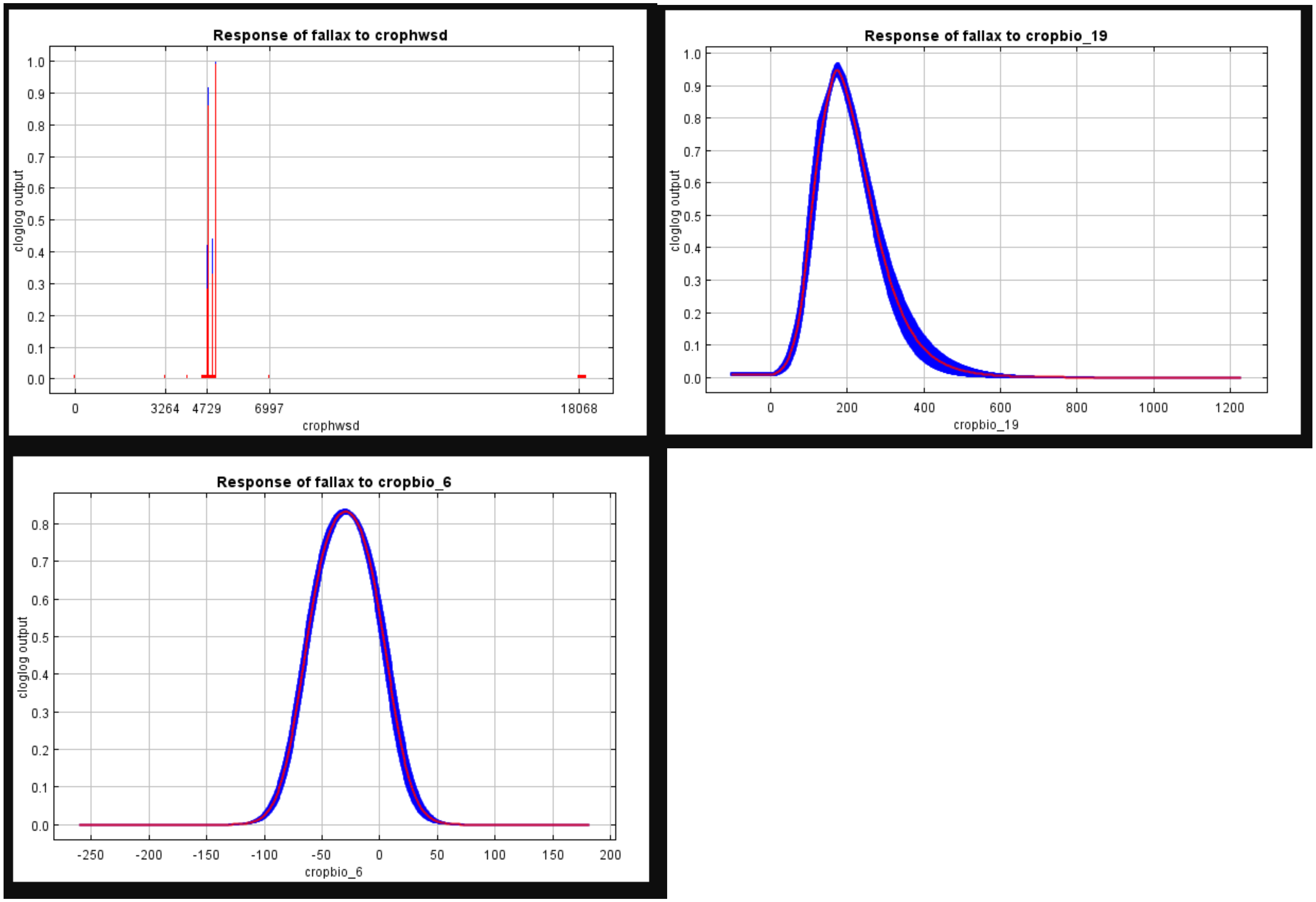
Supplementary Figure 2. Response curves for the top three contributing variables to *P. edulis'* suitable habitat (hwsd = soil type, bio14 = precipitation of driest month, bio8 = mean temperature of wettest quarter). Variable ranges were evaluated with a cloglog threshold > 0.35 (y-axis). Precipitation units are in mm, temperature units are x 10°C (x-axis).



Supplementary Figure 3. Response curves for the top three contributing variables to *P. californiarum*'s suitable habitat (hwsd = soil type, bio6 = minimum temperature of coldest month, bio8 = mean temperature of wettest quarter). Variable ranges were evaluated with a cloglog threshold > 0.51 (y-axis). Temperature units are x 10°C (x-axis).



Supplementary Figure 4. Response curves for the top three contributing variables to *P. quadrifolia*'s suitable habitat (bio14 = precipitation of driest month, bio19 = precipitation of coldest quarter, bio8 = mean temperature of wettest quarter). Variable ranges were evaluated with a cloglog threshold > 0.55 (y-axis). Precipitation units are in mm, temperature units are x 10°C (x-axis).



Supplementary Figure 5. Response curves for the top three contributing variables to *P.* x *fallax*'s suitable habitat (hwsd = soil type, bio19 = precipitation of coldest quarter, bio6 = minimum temperature of coldest month) Variable ranges were evaluated with a cloglog threshold > 0.36 (y-axis). Precipitation units are in mm, temperature units are x 10°C (x-axis).

References

Allen, C.D., Macalady, A.K., Chenchouni, H., Bachelet, D., McDowell, N., Vennetier, M., Kitzberger, T., Rigling, A., Breshears, D.D., Hogg, E.T. and Gonzalez, P., 2010. A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests. *Forest ecology and management*, *259*(4), pp.660-684.

Araújo, M.B. and Rahbek, C., 2006. How does climate change affect biodiversity?. *Science*, 313(5792), pp.1396-1397.

Batllori, E., Lloret, F., Aakala, T., Anderegg, W.R., Aynekulu, E., Bendixsen, D.P., Bentouati, A., Bigler, C., Burk, C.J., Camarero, J.J. and Colangelo, M., 2020. Forest and woodland replacement patterns following drought-related mortality. *Proceedings of the National Academy of Sciences*, *117*(47), pp.29720-29729.

Breshears, D.D., Cobb, N.S., Rich, P.M., Price, K.P., Allen, C.D., Balice, R.G., Romme, W.H., Kastens, J.H., Floyd, M.L., Belnap, J. and Anderson, J.J., 2005. Regional vegetation die-off in response to global-change-type drought. *Proceedings of the National Academy of Science*s, *102*(42), pp.15144-15148.

Buck, R., Ortega-Del Vecchyo, D., Gehring, C., Michelson, R., Flores-Rentería, D., Klein, B., Whipple, A.V., Flores-Rentería, L. 2022. Sequential homoploid hybridization may have facilitated ecological transitions in the Southwestern pinyon pine syngameon. *New Phytologist*, in review.

Burke, E.J., Brown, S.J. and Christidis, N., 2006. Modeling the recent evolution of global drought and projections for the twenty-first century with the Hadley Centre climate model. *Journal of Hydrometeorology*, *7*(5), pp.1113-1125.

Castetter, E.F., Underhill, R.M., Opler, M.E., Bell, W.H. and Grove, A.R., 1935. Ethnobiological studies in the American Southwest (Vol. 1). University of New Mexico Press.

Chambers, J.C., Vander Wall, S.B. and Schupp, E.W., 1999. Seed and seedling ecology of pinon and juniper species in the pygmy woodlands of western North America. *The Botanical Review*, *65*(1), pp.1-38.

Cole, K.L., Fisher, J., Arundel, S.T., Cannella, J. and Swift, S., 2008b. Geographical and climatic limits of needle types of one-and two-needled pinyon pines. *Journal of Biogeography*, *35*(2), pp.257-269.

Cole, Kenneth L., Kirsten E. Ironside, Samantha T. Arundel, Philip Duffy, and John Shaw. "Modeling future plant distributions on the Colorado Plateau: an example using *Pinus edulis*." *In: van Riper, C. III; Sogge, MK, eds. The Colorado Plateau III. Integrating Research and Resources Management for Effective Conservation. Tucson, AZ: The University of Arizona Press. p. 319-330.* (2008a).

Davis, D.E., 1949. Some effects of calcium deficiency on the anatomy of *Pinus taeda*. American *Journal of Botany*, pp.276-282.

Elmore, F.H., 1944. Ethnobotany of the Navajo. In Monographs of the School of American Research (pp. 9-15).

Fick, S.E. and Hijmans, R.J., 2017. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International journal of climatology*, *37*(12), pp.4302-4315

Fischer, G., Nachtergaele, F., Prieler, S., Van Velthuizen, H.T., Verelst, L. and Wiberg, D., 2008. Global agro-ecological zones assessment for agriculture (GAEZ 2008). *IIASA, Laxenburg, Austria and FAO, Rome, Italy*, *10*.

Flake, S.W. and Weisberg, P.J., 2019. Fine-scale stand structure mediates drought-induced tree mortality in pinyon–juniper woodlands. *Ecological Applications*, *29*(2), p.e01831.

Gehring, C., Flores-Rentería, D., Sthultz, C.M., Leonard, T.M., Flores-Rentería, L., Whipple, A.V. and Whitham, T.G., 2014. Plant genetics and interspecific competitive interactions determine ectomycorrhizal fungal community responses to climate change. *Molecular ecology*, *23*(6), pp.1379-1391.

Gitlin, Alicyn R., Christopher M. Sthultz, Matthew A. Bowker, Stacy Stumpf, Kristina L. Paxton, Karla Kennedy, Axhel Muñoz, Joseph K. Bailey, and Thomas G. Whitham. "Mortality gradients within and among dominant plant populations as barometers of ecosystem change during extreme drought." *Conservation Biology* 20, no. 5 (2006): 1477-1486.

Gunther, K. L. 2021. *An integrative approach reveals evolutionary significant units in a not-so monotypic genus from a biodiversity hotspot*. Master's thesis, San Diego State University, California, USA.

Hartmann, H., Bastos, A., Das, A.J., Esquivel-Muelbert, A., Hammond, W.M., Martínez-Vilalta, J., McDowell, N.G., Powers, J.S., Pugh, T.A., Ruthrof, K.X. and Allen, C.D., 2022. Climate Change Risks to Global Forest Health: Emergence of Unexpected Events of Elevated Tree Mortality Worldwide. *Annual Review of Plant Biology*, 73.

Hernández-Quiroz, N.S., Badano, E.I., Barragán-Torres, F., Flores, J. and Pinedo-Álvarez, C., 2018. Habitat suitability models to make conservation decisions based on areas of high species richness and endemism. *Biodiversity and Conservation*, *27*(12), pp.3185-3200.

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

Hijmans, R.J. and Graham, C.H., 2006. The ability of climate envelope models to predict the effect of climate change on species distributions. *Global change biology*, 12(12), pp.2272-2281.

Jeffers, R.M., 1995. Pinon pine seed production, collection, and storage. Desired Future Conditions for Piñon-Juniper Ecosystems, edited by DW Shaw, EF Aldon, and C. LoSapio, pp.191-197.

Jiang, X., Rauscher, S.A., Ringler, T.D., Lawrence, D.M., Williams, A.P., Allen, C.D., Steiner, A.L., Cai, D.M. and McDowell, N.G., 2013. Projected future changes in vegetation in western North America in the twenty-first century. *Journal of Climate*, *26*(11), pp.3671-3687.

Kraft, N.J., Adler, P.B., Godoy, O., James, E.C., Fuller, S. and Levine, J.M., 2015. Community assembly, coexistence and the environmental filtering metaphor. *Functional ecology*, 29(5), pp.592-599.

Lanner, R.M., 1984. Trees of the Great Basin. *A Natural History. U Nevada Press, Reno*.

Little, E.L., 1943. Common insects on pinyon (*Pinus edulis*). *Journal of the New York Entomological Society*, *51*(4), pp.239-252.

McCallum, K.C., 2011. *Regeneration in a changing climate: Modeling and experimental investigations of* Pinus edulis *recruitment* (Doctoral dissertation, Northern Arizona University).

McDowell, N.G. and Allen, C.D., 2015. Darcy's law predicts widespread forest mortality under climate warming. *Nature Climate Change*, 5(7), pp.669-672.

Merow, C., Smith, M.J. and Silander Jr, J.A., 2013. A practical guide to MaxEnt for modeling species' distributions: what it does, and why inputs and settings matter. *Ecography*, *36*(10), pp.1058-1069.

Miller, R.K., 1995. Responding to Tribal Voices in Managing Woodland Resources. Desired Future Conditions for Piñon-juniper Ecosystems, August 8-12, 1994, Flagstaff, Arizona, (258), p.146.

Miller, R.K., 1997. Southwest Woodlands: Cultural Uses of the" Forgotten Forest". *Journal of Forestry*, 95(11), pp.24-28.

Mueller, R.C., Scudder, C.M., Porter, M.E., Talbot Trotter III, R., Gehring, C.A. and Whitham, T.G., 2005. Differential tree mortality in response to severe drought: evidence for long-term vegetation shifts. *Journal of Ecology*, 93(6), pp.1085-1093.

Mukherjee, S., Mishra, A. and Trenberth, K.E., 2018. Climate change and drought: a perspective on drought indices. *Current Climate Change Reports*, *4*(2), pp.145-163.

Ogle, K., Whitham, T.G. and Cobb, N.S., 2000. Tree-ring variation in pinyon predicts likelihood of death following severe drought. *Ecology*, 81(11), pp.3237-3243.

Parmesan, C., 2006. Ecological and evolutionary responses to recent climate change. *Annu. Rev. Ecol. Evol. Syst*., 37, pp.637-669.

Parry, M.L., Canziani, O., Palutikof, J., Van der Linden, P. and Hanson, C. eds., 2007. Climate change 2007-impacts, adaptation and vulnerability: Working group II contribution to the fourth assessment report of the IPCC (Vol. 4). Cambridge University Press.

Patterson, A., Flores-Rentería, L., Whipple, A., Whitham, T. and Gehring, C., 2018. Common garden experiments disentangle plant genetic and environmental contributions to ectomycorrhizal fungal community structure. *New Phytologist*, *221*(1).

Paulin, K.M., Cook, J.J. and Dewey, S.R., 1999. Pinyon-juniper woodlands as sources of avian diversity. *Proceedings: ecology and management of pinyon-juniper communities within the Interior West. USDA Forest Service Proceedings, Ogden, UT*, pp.240-243.

Peterman, W., Waring, R.H., Seager, T. and Pollock, W.L., 2013. Soil properties affect pinyon pine–juniper response to drought. *Ecohydrology*, 6(3), pp.455-463.

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

Redmond, M.D., Forcella, F. and Barger, N.N., 2012. Declines in pinyon pine cone production associated with regional warming. *Ecosphere*, *3*(12), pp.1-14.

Ronco Jr, F.P., 1990. *Pinus edulis* Engelm. pinyon. *Silvics of North America*, 1, pp.327-337.

Sanderson, B. M., Knutti, R., & Caldwell, P. (2015). A representative democracy to reduce interdependency in a multimodel ensemble. *Journal of Climate*, 28(13), 5171-5194.

Schwartz, M.W., 2012. Using niche models with climate projections to inform conservation management decisions. *Biological Conservation*, 155, pp.149-156.

Senf, C., Buras, A., Zang, C.S., Rammig, A. and Seidl, R., 2020. Excess forest mortality is consistently linked to drought across Europe. *Nature communications*, 11(1), pp.1-8.

Severson, K. E. "Small mammals in modified pinyon-juniper woodlands, New Mexico." *Rangeland Ecology & Management/Journal of Range Management Archives* 39, no. 1 (1986): 31-34.

Sthultz, C.M., Gehring, C.A. and Whitham, T.G., 2006. Shifts from competition to facilitation between a foundation tree and a pioneer shrub across spatial and temporal scales in a semiarid woodland. *New Phytologist*, *173*, pp.135-145.

Stone, A.C., Gehring, C.A., Cobb, N.S. and Whitham, T.G., 2018. Genetic-based susceptibility of a foundation tree to herbivory interacts with climate to influence arthropod community composition, diversity, and resilience. *Frontiers in Plant Science*, p.1831.

Todesco, M., Pascual, M.A., Owens, G.L., Ostevik, K.L., Moyers, B.T., Hübner, S., Heredia, S.M., Hahn, M.A., Caseys, C., Bock, D.G. and Rieseberg, L.H., 2016. Hybridization and extinction. *Evolutionary applications*, 9(7), pp.892-908.

Trotter III, R.T., Cobb, N.S. and Whitham, T.G., 2008. Arthropod community diversity and trophic structure: a comparison between extremes of plant stress. *Ecological Entomology*, *33*(1), pp.1-11.

Velazco, S.J.E., Galvao, F., Villalobos, F. and De Marco Junior, P., 2017. Using worldwide edaphic data to model plant species niches: An assessment at a continental extent. *PLoS One*, *12*(10), p.e0186025.

Van Mantgem, P.J., Stephenson, N.L., Byrne, J.C., Daniels, L.D., Franklin, J.F., Fulé, P.Z., Harmon, M.E., Larson, A.J., Smith, J.M., Taylor, A.H. and Veblen, T.T., 2009. Widespread increase of tree mortality rates in the western United States. *Science*, 323(5913), pp.521-524.

Wang, W., Peng, C., Kneeshaw, D.D., Larocque, G.R. and Luo, Z., 2012. Drought-induced tree mortality: ecological consequences, causes, and modeling. *Environmental Reviews*, 20(2), pp.109-121.

Weisberg, P.J. and Ko, D.W., 2012. Old tree morphology in singleleaf pinyon pine (*Pinus monophylla*). *Forest Ecology and Management*, *263*, pp.67-73.

Wells, P.V., 1983. Paleobiogeography of montane islands in the Great Basin since the last glaciopluvial. *Ecological Monographs*, *53*(4), pp.341-382.

White, P.J. and Broadley, M.R., 2003. Calcium in plants. Annals of botany, 92(4), pp.487-511.

Young, D.J., Stevens, J.T., Earles, J.M., Moore, J., Ellis, A., Jirka, A.L. and Latimer, A.M., 2017. Long-term climate and competition explain forest mortality patterns under extreme drought. *Ecology letters*, 20(1), pp.78-86.

Zavarin, E., Snajberk, K. and Cool, L., 1990. Chemical differentiation in relation to the morphology of the single-needle pinyons. *Biochemical Systematics and Ecology*, *18*(2-3), pp.125-137.

Zhu, G., Liu, G., Bu, W. and Gao, Y., 2013. Ecological niche modeling and its applications in biodiversity conservation. *Biodiversity Science*, 21(1), p.90.

Zuquim, G., Costa, F.R., Tuomisto, H., Moulatlet, G.M. and Figueiredo, F.O., 2020. The importance of soils in predicting the future of plant habitat suitability in a tropical forest. *Plant and Soil*, *450*(1), pp.151-170.

CHAPTER V

For the conservation of a recently delineated pinyon species, *Pinus californiarum*

Ryan Buck and Lluvia Flores-Rentería

Abstract

Human-induced extinctions are disproportionately affecting newly and undescribed species, highlighting the importance of conservation assessment studies that focus on these taxa. *Pinus californiarum*, a recently delimited species of pinyon pine, has a restricted range and is projected to lose a majority of its suitable habitat in the near future. Using next-generation sequencing nuclear data, the genetic health and conservation status of *P. californiarum* populations were assessed by estimating population structure, genetic diversity, and inbreeding coefficients. Low genetic diversity, low population structure, and high inbreeding were detected, suggesting that *P. californiarum* may be of conservation concern. Management recommendations are discussed with a goal of increasing genetic diversity, including facilitated migration, translocation, and interspecific gene flow.

Introduction

Anthropogenic activities, including greenhouse gas emission, land use, and species introductions are currently the largest drivers of species extinction and biodiversity loss on a global scale (Parmesan and Yohe, 2003; Hughes 2008; Tylianakis *et al.*, 2008; Bellard *et al.*, 2012; Pecl *et al.*, 2017). Current rates of extinction are estimated to be 1,000 times higher than the background rate due to climate change (Thomas *et al.*, 2004; Pimm *et al.*, 2014; Urban 2015). It is estimated that there are five to 11 million undescribed species

(Chapman 2009; Mora *et al.*, 2011; Appeltans *et al.*, 2012; Costello *et al*., 2013; Pimm *et al.*, 2014), with an estimated 50,000 species of land plants undiscovered (Joppa *et al.*, 2011) and 100,000 taxonomically unresolved (Paton 2013). Around 11-30% of these undescribed species are thought to be under threat of extinction (Tollefson 2019; Liu *et al.*, 2022) and many species have gone or will go extinct before even being described (Pimm *et al*., 2006, 2014; Costello *et al.*, 2013; Todesco *et al.*, 2014).

In order to preserve biodiversity, massive global conservation efforts are required (Hannah 2010; Rands *et al.*, 2010). Unfortunately, scientific research rarely leads to implementable management strategies due to inaccessible jargon, vague recommendations, and non-conservation focused research questions (Ottewell *et al.*, 2016). One method to bridge the gap between researchers and land managers is through conservation genetics (Ottewell *et al.*, 2016). Conservationists can use genetic estimates to gauge species' health with population genetic measures such as inbreeding, effective population sizes, genetic diversity, and population differentiation (Ottewell *et al.*, 2016). Ideal populations are generally characterized as having large effective population sizes and ample gene flow to maintain diversity and offset inbreeding depression (Ottewell *et al.*, 2016). High inbreeding estimates are usually as a result of small or declining populations, which can result in the accumulation of deleterious recessive alleles and therefore decreased diversity (Lowe *et al.*, 2005; Charlesworth and Willis, 2009). Genetic diversity is the fuel for adaptation and is directly related to fitness and extinction risks (Spielman *et al.*, 2014; Rajora and Zinck, 2021). Low genetic diversity is usually a result of inbreeding or a bottleneck event (Allendorf *et al.*, 2013) and can prevent species from evolving in response to climate change (Spielman *et al.*, 2004; Hoban *et al.*, 2020). Lastly, examining population structure can indicate how isolated or divergent populations are and determine the extent of gene flow among populations. Low population differentiation measures typically result from high connectivity among populations or recent divergence (Ottewell *et al.*, 2016). These estimates can be used to identify conservation units and focus conservation efforts on specific populations (Crandall *et al.*, 2000; Fraser and Bernatchez, 2001; Ottewell *et al.*, 2016). Quantifying these measures will help managers understand the trajectory of their populations and determine which genetic processes they should be targeting (Ottewell *et al.*, 2016).

One group of species potentially in need of conservation attention is the Southwestern pinyon pine complex (Buck and Flores-Rentería, in prep). Unlike many undiscovered species, pinyon pines (*Pinus* subsect. *Cembroides*) are neither small nor rare, but are instead long-lived, dominant species (Ronco 1990; Weisberg and Ko, 2012). Pinyon-juniper woodlands are the most abundant forest in Southwestern North America (National Park Service) and the third largest vegetation type in the United States (West 1984). Pinyon pines are anthropologically significant, as they are an important part of Native culture and diet (Simms 1985; Rhode *et al.*, 1998; Bettinger 2021). They are considered foundation species, supporting and structuring their communities above and below ground. Some species have been associated with >1,000 species of arthropods (Little 1943; Trotter *et al.*, 2008; Stone *et al.*, 2018), birds (Wells 1983; Paulin *et al.*, 1999), rodents (Severson 1986), plants (Chambers *et al.*, 1999; Sthultz *et al.*, 2006), and fungi (Gehring *et al.*, 2014; Patterson *et al.*, 2018) collectively. Unfortunately, pinyon pines are especially at risk due to increased fires, droughts, and the intensified presence of bark beetles (Greenwood and Weisberg, 2008; Clifford *et al.*, 2013; Meddens and Hicke, 2014), all of which are projected to increase in frequency and intensity under climate change scenarios (Sheffield and Wood, 2008; Dai 2011; Williams *et al.*, 2013; Trenberth *et al.*, 2014). Pinyons have experienced massive rates of mortality and seed reduction after recent drought events (Ogle *et al.*, 2000; Mueller *et al.*, 2005; Redmond *et al.*, 2012; Flake and Weisberg, 2018), having profound impacts on their ecological communities. Further, future habitat suitability models have predicted large range contractions within the next 50 years (Cole *et al.*, 2008a; Buck and Flores-Rentería, in prep)

Recently, conclusive genetic studies have delineated the morphologically cryptic and commonly misidentified species, *Pinus californiarum* (Buck *et al.*, 2020; Buck *et al.*, in review). Historically, there has been a lot of controversy surrounding the taxonomic status of the single-needled pinyon pines in the Southwest (Lanner 1974; Bailey 1987; Cole *et al.*, 2008b), resulting in decades of research that have mistakenly grouped *P. californiarum* with the other single-needled, more widely distributed species, *P. monophylla*. Consequently, land managers are unaware that the pinyons they oversee are different species and therefore they do not allocate funding/resources appropriately. Further, we do not have accurate assessments of genetic estimates, population census numbers, or on the undocumented mortality occurring in *P. californiarum* (Thomas Oberbauer, pers. comm.). Compared to other closely related species, *P. californiarum* has the most restricted range (5240km^2) (Cole *et al.*, 2008b), distributing in fragmented populations from Southern Nevada to Baja California, Mexico. It tends to occupy hotter and drier habitats than other related species (Cole *et al.*, 2008; Buck *et al.*, 2022), putting it at the extreme of its environment and indicating that it may be at even greater risk for extinction (Hantson *et al.*, 2021). Moreover, Buck and Flores-Rentería (in prep) predicted that *P. californiarum* would lose more than half of its suitable habitat by 2070. As the extinction risk may be higher for newly described species (Liu *et al.*, 2022), our goals are to 1) assess the genetic health of *P. californiarum* populations and 2) provide conservation recommendations to land managers and conservationists.

Methods

Data set - We used a portion of the dataset generated by Buck *et al.* (2020; in review), selecting next-generation sequence data from 69 pure individuals (as identified in Buck *et al.*, 2022's fastStructure results) from nine locations across *P. californiarum*'s range in Arizona, Nevada, southern California, and Baja California (Table 1, Fig. 1). The DArTseq™ dataset contained 15,264 SNPs that were then subsequently filtered to removed loci with lower than 100% reproducibility, missing data lower than 20%, all loci departing from Hardy-Weinberg equilibrium, all monomorphs, and all secondaries, resulting in a final dataset of 1,651 SNPs.

Table 1. Sampling locations with population names, coordinates, and number of trees used in genetic analyses.

Location	**Population Name**	**Latitude**	**Longitude**	**No. pure trees used**
Hualapai Peak, AZ	HU	35.124	-113.914	5
Spirit Mountain, NV	CPT	35.264	-114.751	2
Mojave Desert, CA	MJV	35.174	-115.407	4
Joshua Tree National Park, CA	JTNP	34.016	-116.168	5
San Jacinto, CA	JAC	33.559	-116.610	18
Jacumba Hot Springs, CA	DVT	32.876	-116.410	7
La Rumorosa, Baja CA	RMR	32.522	-116.041	10
Lazaro Cardenas, Baja CA	MRO	31.258	-115.599	11
San Pedro Martir, Baja CA	SPM	31.035	-115.465	7

Figure 1. Distribution map of *P. californiarum* (blue) with sampling locations represented as triangles and labeled for identification (based on USGS 1999; Cole *et al.*, 2003).

To assess the genetic health of *P. californiarum* populations, we examined genetic differentiation, effective population sizes, heterozygosity, and inbreeding coefficients (F_{IS}). To calculate the fixation index (F_{ST}), the "stamppFst" command was implemented from the R package *adegenet* (Jombart 2008) using 1000 bootstraps and a 95% confidence interval. A principal coordinates analysis (PCoA) was performed in R using the package *dartR* and the command "gl.pcoa.plot" to examine genetic distance among populations of *P.*

californiarum by considering differences in allele frequencies between individuals (Gruber *et al.,* 2019). The software fastStructure was used to conduct a Bayesian analysis of K-means population clustering (Raj *et al.*, 2014). The logistic prior was run with five cross-validations, with model complexity (K) selected via the chooseK command. The online application pophelper (Francis 2017) was used to visualize the resulting Q mean bar plots. The effective population size (N_e) of the entire *P. californiarum* range was estimated using NeEstimator v2.1 (Do *et al.*, 2014) via the linkage disequilibrium method (Waples 2006; Jones *et al.*, 2016). In order to calculate the heterozygosity (H_o and H_e) of each population, the command "gl.report.heterozygosity" was implemented from the R package *adegenet*. The inbreeding coefficients (F_{IS}) for each population were calculated with the *hierfstat* package (Goudet 2005) via the "wc" command. For this analysis, the Hualapai Peak (CPT) and Spirit Mountain (HU) populations were grouped together due to the low sample size (n=2) in CPT and the negligible genetic differentiation between the two (see Results).

Results

The fixation index estimates (Table 2) show very low population differentiation with no clear patterns that would suggest geographically distant populations are more differentiated. Estimated values ranged from 0 to 0.085, with an average F_{ST} of 0.045. The PCoA shows no indication of genetic population differentiation, with very minor structuring by geographic distance (Fig. 2). PCoA axis one represents 3.5% of variation in genetic distance and PCoA axis two represents 2.7%. The subsequent axes incrementally represent less variation, with PCoA axis three representing 2.6%, four representing 2.6%, five representing 2.4%, and so on. The fastStructure results (Fig. 3) show one genetic cluster (K=1), with no indication of population substructuring. The effective population size *P. californiarum* was estimated to be fairly low, with an average of 361 individuals and 95% confidence intervals ranging from 280 to 509 individuals. The observed heterozygosity levels (Table 3) for every sampled population were estimated to be lower than the expected heterozygosity levels under Hardy-Weinberg equilibrium, with the exception of the population from Spirit Mountain, NV (CPT). H_o estimates ranged from 0.036 to 0.063, with an average of 0.052. The estimated inbreeding coefficients for each *P. californiarum* population were very high (Table 3), indicating a large amount of mating among related individuals. F_{IS} estimates ranged from 0.360 to 0.545, with an average of 0.430.

Table 2. Pairwise F-statistic estimates per sampled population, with lower numbers indicating less differentiation (on a scale from 0-1).

Population	**JTNP**	**DVT**	**JAC**	**MJV**	**MRO**	**RMR**	**SPM**	**CPT**
DVT	0.039							
JAC	0.026	0.026						
MJV	0.035	0.047	0.027					
MRO	0.080	0.044	0.057	0.078				
RMR	0.048	-0.002	0.033	0.044	0.057			
SPM	0.070	0.037	0.049	0.070	0.062	0.056		
CPT	0.030	0.050	0.043	0.029	0.085	0.034	0.063	
HU	0.036	0.042	0.039	0.013	0.077	0.0456	0.050	0.000

Table 3. Heterozygosity and inbreeding coefficient estimates per population

Population	**Observed Heterozygosity**	**Expected Heterozygosity**	**Inbreeding Coefficient**
CPT	0.063	0.061	(combined with HU)
DVT	0.056	0.079	0.360
HU	0.049	0.081	0.459
JAC	0.061	0.091	0.361
JTNP	0.059	0.084	0.407
MJV	0.057	0.077	0.395
MRO	0.046	0.076	0.436
RMR	0.048	0.082	0.475
SPM	0.036	0.068	0.545
Average	0.052	0.078	0.430

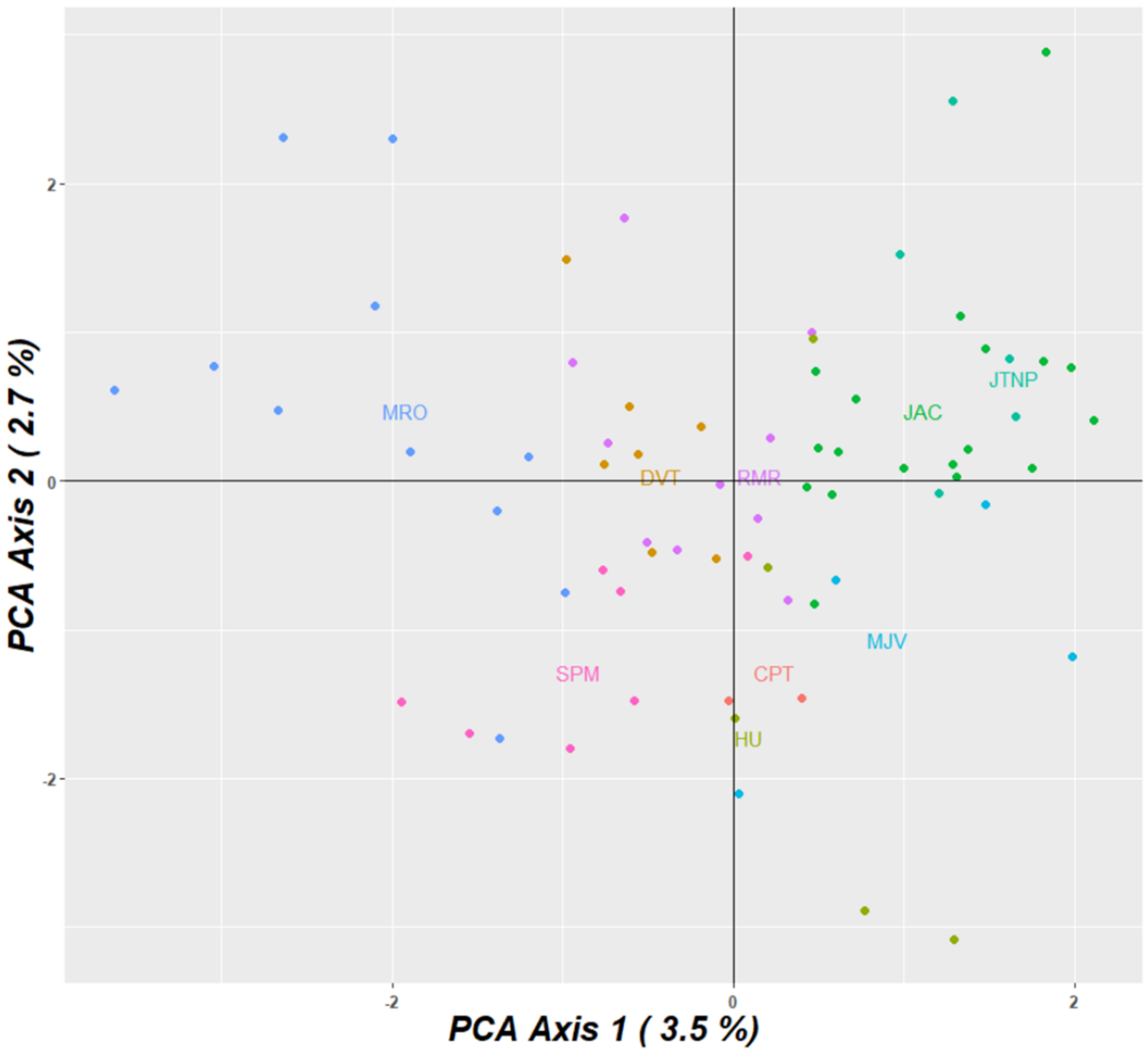


Figure 2. Principal coordinates analysis showing the variation in genetic distance, with each dot representing an individual colored by the nine population's identities.

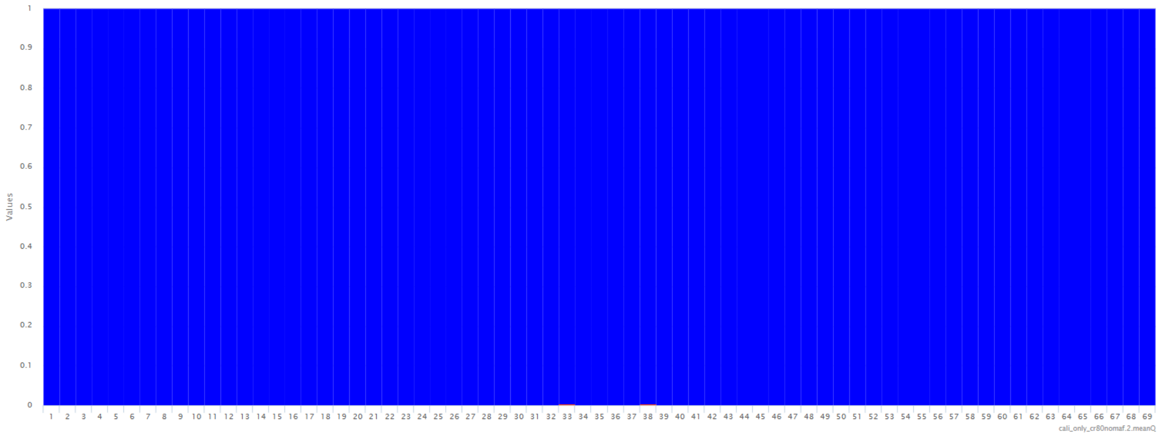


Figure 3. fastStructure plot showing one genetic cluster (K=1) colored by genetic identity. Each tick mark on the x-axis represents an individual, with the probability of those individuals belonging to the genetic cluster indicated by the proportion of color on the y-axis.

Discussion

Conservation genetics can bridge the gap between research and applicable management recommendations (Ottewell *et al.*, 2016). Quantifying population genetic estimates can help identify populations at risk of extinction and target specific management approaches. Conservation genetic studies are critically important in the light of climate change as species are rapidly disappearing, some before they are even described or discovered (Pimm *et al.*, 2006, 2014; Costello *et al.*, 2013; Todesco *et al.*, 2014). Extremely low measures of genetic differentiation and diversity were seen in populations of *P. californiarum*. Coupled with high inbreeding estimates and low effective population sizes, the trajectory of *P. californiarum* appears to be declining.

Genetic health of P. californiarum

The lack of population structure seen among *P. californiarum* populations (Fig. 2-3, Table 2) was quite low, suggesting high gene flow among populations or recent divergence. The patterns seen could be due to long distance, wind-mediated pollen dispersal, which can viably travel at least 41 km (Williams 2010) and corvid-mediated seed dispersal up to 22 km (Wells 1983), both of which are sufficient to reach neighboring populations, with the exception of the unsampled population in El Crucero, Baja California which is separated from San Pedro Martir by 62 km. Alternatively, Buck *et al.* (2022) estimated a relatively recent divergence time (70,000-150,000ya) for *P. californiarum*, so it is possible that enough time has passed for speciation, but not for fixed interpopulation differences to accumulate (Ottewell *et al.*, 2016). However, when compared to estimates in other pine species (Suppl. Table 1), *P. californiarum* has extraordinarily low F_{ST} values, only coming close to that of *P. maximartinezii* ($F_{ST} = 0.081$) which is a rare and endangered Mexican pinyon pine (Ledig *et al.*, 1999). Similar F_{ST} values were seen between a single *P. edulis* population and its descendants within a 12km radius (Premoli *et al.*, 1994), implying that *P. californiarum* populations have comparable genetic structuring to recently diverged populations.

The observed heterozygosity levels of each population were also abnormally low, indicating a lack of genetic diversity in *P. californiarum*. Due to the lack of population structure found, the low heterozygosity observed is likely not a result of Wahlund effects (Wahlund 1928; Zhivotovsky 2015). Compared to other species' population-level heterozygosity (Suppl. Table 1), the populations of *P. californiarum* have anywhere from

43 to 95% less diversity. While *P. californiarum* has elevated species-level heterozygosity compared to other species in the subsection (Buck *et al.*, 2022), this is likely a result of its hybrid origin and should be scrutinized with other qualifications such as *P. californiarum*'s restricted range, high inbreeding coefficients, and potential future habitat loss (Buck and Flores-Rentería, 2022). The lower observed heterozygosity of each population compared to the expected heterozygosity (with the exception of Spirit Mountain, NV) may provide a more accurate scale of missing diversity, but still indicates potential inbreeding. As low genetic diversity is directly related to evolutionary potential (Spielman *et al.*, 2004; Rajora and Zinck, 2021), *P. californiarum* may not be able to adapt to future climatic conditions.

The high inbreeding coefficients calculated (Table 3; Fig. 4) suggest mating among related individuals is prevalent in these populations. In comparison, other pine species generally show lower inbreeding coefficients (Suppl. Table 1), with the exception of rare Mexican pines (*P. pinceana*, *P. lagunae*, and *P. muricata* in Molina-Freaner *et al.*, 2001). With inbreeding levels this high, *P. californiarum* is likely experiencing or will soon experience inbreeding depression, but more research measuring fitness in this species must be done. As high inbreeding exposes recessive deleterious alleles to selection, it is possible that *P. californiarum* has already purged these alleles and reduced its genetic load (Charlesworth and Charlesworth, 1998; Vogl *et al.*, 2002). However, pines are known to have high outcrossing rates (Lanner 1980; Molina-Freaner *et al.*, 2001), with most pines having unsurvivably high genetic load (Vogl *et al.*, 2002). Yet some studies show that inbreeding in pines only results in a reduced seed set (Snyder 1968; Franklin 1970;

Sorensen 1970), which could lead to a further reduction in population size and even more inbreeding, but is not fatal.

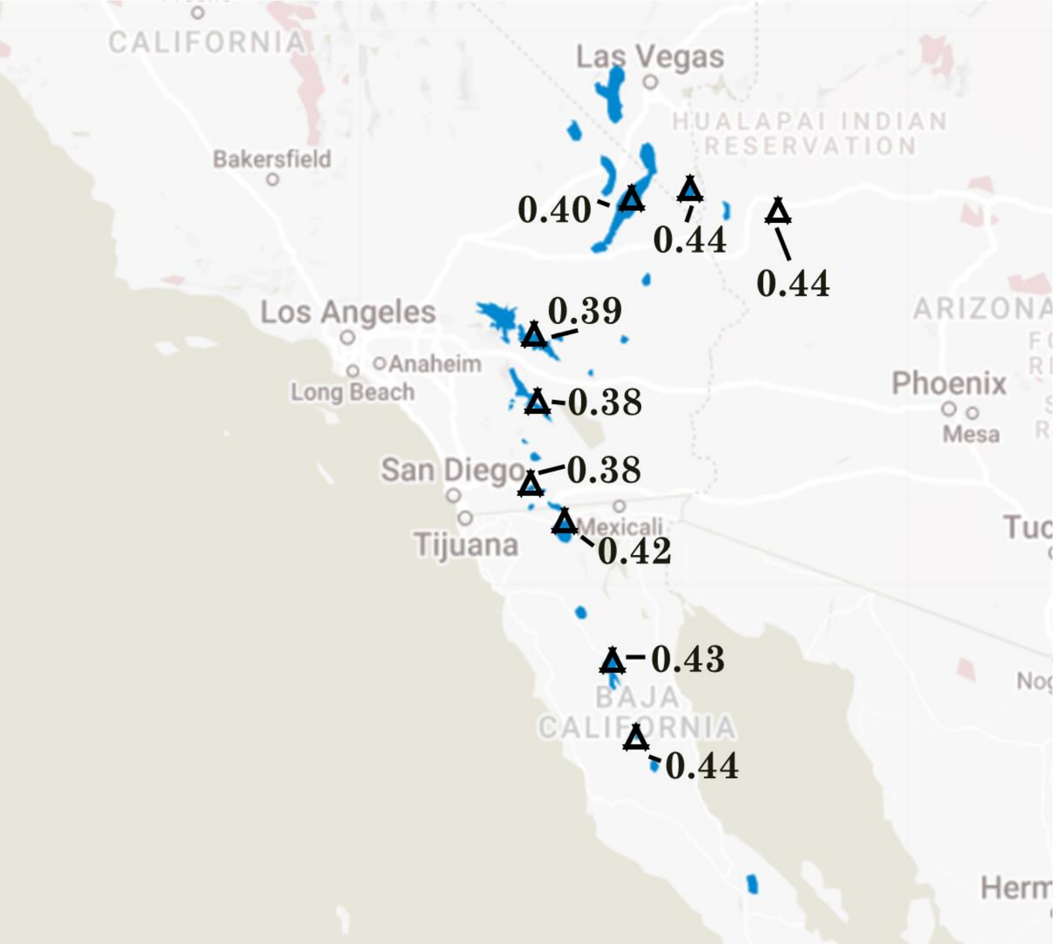


Figure 4. Inbreeding coefficients overlain on the distribution map

The somewhat counterintuitive patterns seen here could be due to the small range size of *P. californiarum* with some tree stands only carrying a dozen or so individuals scattered across the landscape. However, because of the lack of population structure seen, high inbreeding is not likely from populations being too isolated. While low population differentiation can indicate connectivity, the high inbreeding coefficients suggest interpopulation gene flow is either uncommon or insufficient enough to increase genetic diversity. It is also not likely due to sampling gaps, as sampling occurred in almost every known location of *P. californiarum*, with samples taken at least 30m apart. Another explanation of the low heterozygosity and high inbreeding observed is that most of the *P. californiarum* populations overlap in distribution with at least one other species (usually *P. quadrifolia* and sometimes *P. monophylla*; Buck *et al.*, 2020; Buck *et al.*, in review). These sympatric populations are typically made up of admixed individuals with a few pure individuals, so a majority of the pure individuals may have been replaced by hybrids, a process known as genetic swamping (Todesco *et al.*, 2016). Alternatively, *P. californiarum* could have recently experienced a large bottleneck event, resulting in substantial decreases of genetic diversity and population sizes and increases of inbreeding (Nei *et al*., 1975; Chakraborty and Nei, 1977). This would have likely been followed either by an expansion into *P. californiarum*'s current habitat or constant interpopulation gene flow to produce the low population differentiation seen, as long-term inbreeding in isolated populations would produce high population structure (Jain 1976).

Conservation recommendations

Ottewell *et al.* (2016) established a decision-making framework based on measurements of genetic differentiation, genetic diversity, and inbreeding to aid researchers in providing relevant conservation recommendations to managers. Based on the genetic measures in this study: low genetic differentiation, low genetic diversity, and high inbreeding, strategy four is recommended (Ottewell *et al.*, 2016). This strategy consists of the active introduction of new genetic material through facilitated migration or translocation from genetically diverse populations. Additionally, they suggest that if inbreeding is recent, further research should be conducted to recover diversity from within the populations, giving the example of a soil seedbank as a potential source. The latter guidance is not recommended as seedbanks in pinyons are typically restricted to corvid and rodent caches (Balda and Kamil, 1989; Vander Wall, 1997; Pearson and Theimer, 2004) that would not be practical to locate. However, collecting, storing, and growing seed of pure *P. californiarum* individuals is recommended to curb the rapid decline seen in some populations. Some potential issues are foreseen with choosing a source population for facilitated migration, as little genetic differentiation was seen among populations and most populations showed low genetic diversity. Similarly, translocations of large pinyon trees may not be pragmatic and removing individuals from already small populations could result in further declines. However, we recognize that some of these solutions are still more realistic than protecting large swaths of land or curbing carbon emissions so informed efforts should still be made.

An alternative, but somewhat controversial method of introducing genetic variation into a population is through interspecific gene flow. Gene flow is frequently viewed as detrimental to species purity among traditional conservationists; however, the potential benefits including increased genetic variation, population size, and adaptive introgression, may outweigh the potential outbreeding risks (Hamilton and Miller, 2016; Chan *et al.*, 2018; Quilodran *et al.*, 2020). With anthropogenic activities rapidly driving global species extinction events, gene flow may be the only tool that acts quickly enough to save biodiversity, and thus must seriously be considered in conservation work. In *P. californiarum*, hybridization may be the most effective way to increase genetic variation, especially given the uniformly high levels of inbreeding and low genetic diversity across populations. Moreover, hybridization is already occurring naturally in these species, with a majority of *P. californiarum* populations containing individuals with admixture (Buck *et al.*, 2020; Buck *et al.*, in review). Perhaps the current populations are only existing not in spite of, but because they exchange genes with other species. Future studies should focus on hybrid fitness to test for any potential outbreeding depression before substantial recommendations for the creating and planting of hybrid seeds can be made. However, seeds from naturally occuring admixed individuals could be stored in seed banks until conclusive results are produced.

With human-induced climate change threatening the future distribution of species worldwide (Oreskes 2004), slowly migrating species with low connectivity between populations like *P. californiarum* are at increased risk of extinction (Fahrig 2002; Pearson 2006; Cole *et al.*, 2013). Researchers and those working with pinyons must recognize *P. californiarum*'s taxonomic status as an independent species in order to properly allocate resources and direct conservation funding. Land managers and conservationists should begin enacting protocols to ensure the future of this species, including facilitated migration, translocations, and potentially interspecific crosses.

Supplementary Information:

Supplementary Table 1. Genetic estimates gathered from other pine species

Species	**Data type**	**Fst**	**Ho**	**He**	**Fis**	**Ne**	**Source**
P. strobus	microsatellites and SNPs	-	0.670	0.720	0.075	-	Rajora and Zinck, 2021
P. nelsonii	chloroplast	0.130	0.727	-	-	10,000 to 23,500	Cuenca *et al.*, 2003
P. pinceana	allozyme	0.140	0.144	0.174	0.116	-	Ledig *et al.*, 2001
P. pinceana	allozyme	0.247	0.216	0.374	0.458	-	Molina-Freaner *et al.*, 2001
P. rzedowskii	isozyme	0.175	0.162	0.219	0.247	-	Delgado *et al.*, 1999
P. maximartinezii	allozyme	0.081	0.110	0.122	0.101	-	Ledig *et al.*, 1999
P. lagunae	allozyme	0.188	0.188	0.386	0.534	-	Molina-Freaner *et al.*, 2001
P. muricata	allozyme	0.161	0.253	0.346	0.307	-	Molina-Freaner *et al.*, 2001
P. californiarum	nuclear SNPs	0.045	0.052	0.078	0.430	280 to 509	this study

References

Allendorf, F.W., Funk, W.C., Aitken, S.N., Byrne, M. and Luikart, G., 2022. Conservation and the genomics of populations. Oxford University Press.

Appeltans, W., Ahyong, S.T., Anderson, G., Angel, M.V., Artois, T., Bailly, N., Bamber, R., Barber, A., Bartsch, I., Berta, A. and Błażewicz-Paszkowycz, M., 2012. The magnitude of global marine species diversity. *Current biology*, 22(23), pp.2189-2202.

Bailey, D.K., 1987. A study of *Pinus* subsection *Cembroides*. I: The single-needle pinyons of the Californias and the Great Basin. *Notes from the Royal Botanic Garden, Edinburgh*, *44*(2), pp.275-310.

Balda, R.P. and Kamil, A.C., 1989. A comparative study of cache recovery by three corvid species. *Animal Behaviour*, *38*(3), pp.486-495.

Bellard, C., Bertelsmeier, C., Leadley, P., Thuiller, W. and Courchamp, F., 2012. Impacts of climate change on the future of biodiversity. *Ecology letters*, *15*(4), pp.365-377.

Bettinger, R.L., 2021. 13 Native Land Use: Archaeology and Anthropology. In *Natural History of the White-Inyo Range, Eastern California* (pp. 463-486). University of California Press.

Buck, R. Flores-Rentería, L. 2022. Predicting the Future: Habitat Suitability of the Southwestern Pinyon Pines. *In prep*.

Buck, R., Hyasat, S., Hossfeld, A. and Flores-Rentería, L., 2020. Patterns of hybridization and cryptic introgression among one-and four-needled pinyon pines. *Annals of Botany*, 126(3), pp.401-411.

Buck, R., Ortega-Del Vecchyo, D., Gehring, C., Michelson, R., Flores-Rentería, D., Klein, B., Whipple, A.V., Flores-Rentería, L. 2022. Sequential homoploid hybridization may have facilitated ecological transitions in the Southwestern pinyon pine syngameon. *New Phytologist*, in review.

Chakraborty, R. and Nei, M., 1977. Bottleneck effects on average heterozygosity and genetic distance with the stepwise mutation model. *Evolution*, pp.347-356.

Chapman, A.D., 2009. Numbers of living species in Australia and the world. *Report for the Australian Biodiversity Resources Study*, 2nd edition, pp. 1-78.

Chambers, J.C., Vander Wall, S.B. and Schupp, E.W., 1999. Seed and seedling ecology of pinon and juniper species in the pygmy woodlands of western North America. *The Botanical Review*, *65*(1), pp.1-38.

Chan, W.Y., Hoffmann, A.A. and van Oppen, M.J., 2019. Hybridization as a conservation management tool. *Conservation Letters*, *12*(5), p.e12652

.

Charlesworth, B. and Charlesworth, D., 1998. Some evolutionary consequences of deleterious mutations. *Genetica,* 102, pp.3-19.

Charlesworth, D. and Willis, J.H., 2009. The genetics of inbreeding depression. *Nature reviews genetics*, 10(11), pp.783-796.

Clifford, M.J., Royer, P.D., Cobb, N.S., Breshears, D.D. and Ford, P.L., 2013. Precipitation thresholds and drought-induced tree die-off: insights from patterns of *Pinus edulis* mortality along an environmental stress gradient. *New Phytologist*, *200*(2), pp.413-421.

Cole, K.L., Ferguson, G., Cannella, J., Spellenberg, R., Saunders, A., Arundel, S. & Riser, J. (2003) *Digital rangemaps for North American one and two-needled pinyon pines (*Pinus monophylla, P. edulis, *and* fallax *and* californiarum*-types)*. Available at: http://sbsc.wr.usgs.gov/cprs/research/projects/global_change/RangeMaps.asp.

Cole, K.L., Fisher, J., Arundel, S.T., Cannella, J. and Swift, S., 2008b. Geographical and climatic limits of needle types of one-and two-needled pinyon pines. *Journal of Biogeography*, 35(2), pp.257-269.

Cole, K.L., Fisher, J.F., Ironside, K., Mead, J.I. and Koehler, P., 2013. The biogeographic histories of *Pinus edulis* and *Pinus monophylla* over the last 50,000 years. *Quaternary International*, *310*, pp.96-110.

Cole, Kenneth L., Kirsten E. Ironside, Samantha T. Arundel, Philip Duffy, and John Shaw. "Modeling future plant distributions on the Colorado Plateau: an example using *Pinus edulis*." *In: van Riper, C. III; Sogge, MK, eds. The Colorado Plateau III. Integrating Research and Resources Management for Effective Conservation. Tucson, AZ: The University of Arizona Press. p. 319-330.* (2008a).

Crandall, K.A., Bininda-Emonds, O.R., Mace, G.M. and Wayne, R.K., 2000. Considering evolutionary processes in conservation biology. *Trends in ecology & evolution*, 15(7), pp.290-295.

Cuenca, A., Escalante, A.E. and Pinero, D., 2003. Long-distance colonization, isolation by distance, and historical demography in a relictual Mexican pinyon pine (*Pinus nelsonii* Shaw) as revealed by paternally inherited genetic markers (cpSSRs). *Molecular Ecology,* 12(8), pp.2087-2097.

Dai, A., 2011. Drought under global warming: a review. *Wiley Interdisciplinary Reviews: Climate Change*, *2*(1), pp.45-65.

Delgado, P., Pin~ ero, D., Chaos, A., Pe´ rez-Nasser, N. and Alvarez-Buylla, E.R., 1999. High population differentiation and genetic variation in the endangered Mexican pine *Pinus rzedowskii* (Pinaceae). *American Journal of Botany*, *86*(5), pp.669-676.

Do, C., Waples, R.S., Peel, D., Macbeth, G.M., Tillett, B.J. and Ovenden, J.R., 2014. NeEstimator v2: Re-implementation of software for the estimation of contemporary effective population size (Ne) from genetic data. *Molecular ecology resources*, 14(1), pp.209-214.

Fahrig, L., 2002. Effect of habitat fragmentation on the extinction threshold: a synthesis. *Ecological applications*, *12*(2), pp.346-353.

Flake, S.W. and Weisberg, P.J., 2019. Fine-scale stand structure mediates drought-induced tree mortality in pinyon–juniper woodlands. *Ecological Applications*, *29*(2), p.e01831.

Francis RM. 2017. pophelper: an R package and web app to analyse and visualize population structure. *Molecular Ecology Resources* 17: 27–32.

Franklin, E. C. 1970. Survey of mutant forms and inbreeding depression in species of the family Pinaceae. U.S. Forest Service Research Paper SE-61.

Fraser, D.J. and Bernatchez, L., 2001. Adaptive evolutionary conservation: towards a unified concept for defining conservation units. *Molecular ecology*, 10(12), pp.2741-2752.

Gehring, C., Flores-Rentería, D., Sthultz, C.M., Leonard, T.M., Flores-Rentería, L., Whipple, A.V. and Whitham, T.G., 2014. Plant genetics and interspecific competitive interactions determine ectomycorrhizal fungal community responses to climate change. *Molecular ecology*, *23*(6), pp.1379-1391.

Goudet, J., 2005. Hierfstat, a package for R to compute and test hierarchical F-statistics. *Molecular Ecology Notes*, *5*(1), pp.184-186.

Greenwood, D.L. and Weisberg, P.J., 2008. Density-dependent tree mortality in pinyon-juniper woodlands. *Forest Ecology and Management*, *255*(7), pp.2129-2137.

Gruber B, Unmack P, Berry O, Georges A. 2019. Introduction to dartR. https://rdrr.io/cran/dartR/f/inst/doc/IntroTutorial_dartR.pdf. 29 November 2019.

Hamilton, J.A. and Miller, J.M., 2016. Adaptive introgression as a resource for management and genetic conservation in a changing climate. *Conservation Biology*, *30*(1), pp.33-41.

Hannah, L., 2010. A global conservation system for climate-change adaptation. *Conservation Biology*, 24(1), pp.70-77.

Hantson, S., Huxman, T.E., Kimball, S., Randerson, J.T. and Goulden, M.L., 2021. Warming as a driver of vegetation loss in the Sonoran Desert of California. *Journal of Geophysical Research: Biogeosciences*, *126*(6), p.e2020JG005942.

Hoban, S., Bruford, M., Jackson, J.D.U., Lopes-Fernandes, M., Heuertz, M., Hohenlohe, P.A., Paz-Vinas, I., Sjögren-Gulve, P., Segelbacher, G., Vernesi, C. and Aitken, S., 2020. Genetic diversity targets and indicators in the CBD post-2020 Global Biodiversity Framework must be improved. *Biological Conservation*, 248, p.108654.

Hughes, L., 2000. Biological consequences of global warming: is the signal already apparent?. *Trends in ecology & evolution, 15*(2), pp.56-61.

Jain, S.K., 1976. The evolution of inbreeding in plants. *Annual review of ecology and systematics*, *7*(1), pp.469-495.

Jombart T. 2008. adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics (Oxford, England)* 24: 1403–1405.

Jones, A.T., Ovenden, J.R. and Wang, Y.G., 2016. Improved confidence intervals for the linkage disequilibrium method for estimating effective population size. *Heredity*, 117(4), pp.217-223.

Joppa, L.N., Roberts, D.L. and Pimm, S.L., 2011. How many species of flowering plants are there?. *Proceedings of the Royal Society B: Biological Sciences*, 278(1705), pp.554-559.

Lanner, R.M., 1974. Natural hybridization between Pinus edulis and Pinus monophylla in the American Southwest. *Silvae Genetica*.

Lanner, R.M., 1980. A self-pollination experiment in *Pinus edulis*. *The Great Basin Naturalist*, pp.265-267.

Ledig, F.T., Capó-Arteaga, M.A., Hodgskiss, P.D., Sbay, H., Flores-López, C., Thompson Conkle, M. and Bermejo-Velázquez, B., 2001. Genetic diversity and the mating system of a rare Mexican piñon, *Pinus pinceana*, and a comparison with *Pinus maximartinezii* (Pinaceae). *American Journal of Botany*, 88(11), pp.1977-1987.

Ledig, F.T., Conkle, M.T., Bermejo-Velázquez, B., Eguiluz-Piedra, T., Hodgskiss, P.D., Johnson, D.R. and Dvorak, W.S., 1999. Evidence for an extreme bottleneck in a rare Mexican pinyon: genetic diversity, disequilibrium, and the mating system in *Pinus maximartinezii*. *Evolution*, 53(1), pp.91-99.

Little, E.L., 1943. Common insects on pinyon (Pinus edulis). *Journal of the New York Entomological Society*, *51*(4), pp.239-252.

Liu, J., Slik, F., Zheng, S. and Lindenmayer, D.B., 2022. Undescribed species have higher extinction risk than known species. *Conservation Letters*, p.e12876.

Lowe, A.J., Boshier, D., Ward, M., Bacles, C.F.E. and Navarro, C., 2005. Genetic resource impacts of habitat loss and degradation; reconciling empirical evidence and predicted theory for neotropical trees. *Heredity*, 95(4), pp.255-273.

Meddens, A.J. and Hicke, J.A., 2014. Spatial and temporal patterns of Landsat-based detection of tree mortality caused by a mountain pine beetle outbreak in Colorado, USA. *Forest Ecology and Management*, *322*, pp.78-88.

Molina-Freaner, F., Delgado, P., Piñero, D., Perez-Nasser, N. and Alvarez-Buylla, E., 2001. Do rare pines need different conservation strategies? Evidence from three Mexican species. *Canadian Journal of Botany*, 79(2), pp.131-138.

Mora, C., Tittensor, D.P., Adl, S., Simpson, A.G. and Worm, B., 2011. How many species are there on Earth and in the ocean?. *PLoS biology*, 9(8), p.e1001127.

Mueller, R.C., Scudder, C.M., Porter, M.E., Talbot Trotter III, R., Gehring, C.A. and Whitham, T.G., 2005. Differential tree mortality in response to severe drought: evidence for long-term vegetation shifts. *Journal of Ecology*, 93(6), pp.1085-1093.

Nei, M., Maruyama, T. and Chakraborty, R., 1975. The bottleneck effect and genetic variability in populations. *Evolution*, pp.1-10.

Ogle, K., Whitham, T.G. and Cobb, N.S., 2000. Tree-ring variation in pinyon predicts likelihood of death following severe drought. *Ecology,* 81(11), pp.3237-3243.

Ottewell, K.M., Bickerton, D.C., Byrne, M. and Lowe, A.J., 2016. Bridging the gap: A genetic assessment framework for population-level threatened plant conservation prioritization and decision-making. *Diversity and Distributions,* 22(2), pp.174-188.

Parmesan, C., 2006. Ecological and evolutionary responses to recent climate change. *Annu. Rev. Ecol. Evol. Syst.*, *37*, pp.637-669.

Parmesan, C. and Yohe, G., 2003. A globally coherent fingerprint of climate change impacts across natural systems. *Nature*, *421*(6918), pp.37-42.

Paton, A.J., 2013. From Working List to Online Flora of All Known Plants—Looking Forward with Hindsight. *Annals of the Missouri Botanical Garden,* 99(2), pp.206-213.

Patterson, A., Flores-Rentería, L., Whipple, A., Whitham, T. and Gehring, C., 2018. Common garden experiments disentangle plant genetic and environmental contributions to ectomycorrhizal fungal community structure. *New Phytologist*, *221*(1).

Paulin, K.M., Cook, J.J. and Dewey, S.R., 1999. Pinyon-juniper woodlands as sources of avian diversity. *Proceedings: ecology and management of pinyon-juniper communities within the Interior West. USDA Forest Service Proceedings, Ogden, UT*, pp.240-243.

Pearson, R.G., 2006. Climate change and the migration capacity of species. *Trends in ecology & evolution*, *21*(3), pp.111-113.

Pearson, K.M. and Theimer, T.C., 2004. Seed-caching responses to substrate and rock cover by two *Peromyscus* species: implications for pinyon pine establishment. *Oecologia*, *141*(1), pp.76-83.

Pecl, G.T., Araújo, M.B., Bell, J.D., Blanchard, J., Bonebrake, T.C., Chen, I.C., Clark, T.D., Colwell, R.K., Danielsen, F., Evengård, B. and Falconi, L., 2017. Biodiversity redistribution under climate change: Impacts on ecosystems and human well-being. *Science*, 355(6332), p.eaai9214.

Pimm, S.L., Jenkins, C.N., Abell, R., Brooks, T.M., Gittleman, J.L., Joppa, L.N., Raven, P.H., Roberts, C.M. and Sexton, J.O., 2014. The biodiversity of species and their rates of extinction, distribution, and protection. *Science*, *344*(6187), p.1246752.

Pimm, S., Raven, P., Peterson, A., Şekercioğlu, Ç.H. and Ehrlich, P.R., 2006. Human impacts on the rates of recent, present, and future bird extinctions. *Proceedings of the National Academy of Sciences*, 103(29), pp.10941-10946.

Premoli, A.C., Chischilly, S. and Mitton, J.B., 1994. Levels of genetic variation captured by four descendant populations of pinyon pine (*Pinus edulis* Engelm.). *Biodiversity & Conservation*, 3(4), pp.331-340.

Quilodrán, C.S., Montoya-Burgos, J.I. and Currat, M., 2020. Harmonizing hybridization dissonance in conservation. *Communications Biology*, *3*(1), pp.1-10.

Raj A, Stephens M, Pritchard JK. 2014. fastSTRUCTURE: variational inference of population structure in large SNP data sets. *Genetics* 197: 573–589.

Rajora, O.P. and Zinck, J.W., 2021. Genetic Diversity, Structure and Effective Population Size of Old-Growth vs. Second-Growth Populations of Keystone and Long-Lived Conifer, Eastern White Pine (*Pinus strobus*): Conservation Value and Climate Adaptation Potential. *Frontiers in genetics*, *12*.

Rands, M.R., Adams, W.M., Bennun, L., Butchart, S.H., Clements, A., Coomes, D., Entwistle, A., Hodge, I., Kapos, V., Scharlemann, J.P. and Sutherland, W.J., 2010. Biodiversity conservation: challenges beyond 2010. *Science*, 329(5997), pp.1298-1303.

Redmond, M.D., Forcella, F. and Barger, N.N., 2012. Declines in pinyon pine cone production associated with regional warming. *Ecosphere*, *3*(12), pp.1-14.

Rhode, D. and Madsen, D.B., 1998. Pine nut use in the early Holocene and beyond: the Danger Cave archaeobotanical record. *Journal of Archaeological Science*, *25*(12), pp.1199-1210.

Ronco, Frank P., Jr. 1990. *Pinus edulis* Engelm. pinyon. In: Burns, Russell M.; Honkala, Barbara H., technical coordinators. Silvics of North America. Volume 1. Conifers. Agric. Handb. 654. Washington, DC: U.S. Department of Agriculture, Forest Service: 327-337.

Severson, K. E. "Small mammals in modified pinyon-juniper woodlands, New Mexico." *Rangeland Ecology & Management/Journal of Range Management Archives* 39, no. 1 (1986): 31-34.

Sheffield, J. and Wood, E.F., 2008. Projected changes in drought occurrence under future global warming from multi-model, multi-scenario, IPCC AR4 simulations. *Climate dynamics*, *31*(1), pp.79-105.

Simms, S.R., 1985. Pine nut use in three Great Basin cases: Data, theory, and a fragmentary material record. *Journal of California and Great Basin Anthropology*, *7*(2), pp.166-175.

Snyder, E.B., 1968. Seed yield and nursery performance of self-pollinated slash pines. Forest Science, 14(1), pp.68-74.

Sorensen, F.C., 1970. Self-fertility of a central Oregon source of ponderosa pine (Vol. 109). Pacific Northwest Forest and Range Experiment Station, US Department of Agriculture, Forest Service.

Spielman, D., Brook, B.W. and Frankham, R., 2004. Most species are not driven to extinction before genetic factors impact them. *Proceedings of the National Academy of Sciences*, 101(42), pp.15261-15264.

Sthultz, C.M., Gehring, C.A. and Whitham, T.G., 2006. Shifts from competition to facilitation between a foundation tree and a pioneer shrub across spatial and temporal scales in a semiarid woodland. *New Phytologist, 173*, pp.135-145.

Stone, A.C., Gehring, C.A., Cobb, N.S. and Whitham, T.G., 2018. Genetic-based susceptibility of a foundation tree to herbivory interacts with climate to influence arthropod community composition, diversity, and resilience. *Frontiers in Plant Science*, p.1831.

Thomas, C.D., Cameron, A., Green, R.E., Bakkenes, M., Beaumont, L.J., Collingham, Y.C., Erasmus, B.F., De Siqueira, M.F., Grainger, A., Hannah, L. and Hughes, L., 2004. Extinction risk from climate change. *Nature*, 427(6970), pp.145-148.

Todesco, M., Pascual, M.A., Owens, G.L., Ostevik, K.L., Moyers, B.T., Hübner, S., Heredia, S.M., Hahn, M.A., Caseys, C., Bock, D.G. and Rieseberg, L.H., 2016. Hybridization and extinction. *Evolutionary applications*, *9*(7), pp.892-908.

Tollefson, J., 2019. Humans are driving one million species to extinction. *Nature*, 569(7755), pp.171-172.

Trenberth, K.E., Dai, A., Van Der Schrier, G., Jones, P.D., Barichivich, J., Briffa, K.R. and Sheffield, J., 2014. Global warming and changes in drought. *Nature Climate Change*, *4*(1), pp.17-22.

Trotter III, R.T., Cobb, N.S. and Whitham, T.G., 2008. Arthropod community diversity and trophic structure: a comparison between extremes of plant stress. *Ecological Entomology*, *33*(1), pp.1-11.

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

Urban, M.C., 2015. Accelerating extinction risk from climate change. *Science*, 348(6234), pp.571-573.

USGS. 1999. Digital representation of 'Atlas of United States Trees'. https://archive.usgs.gov/archive/sites/www.usgs.gov/science/cite-view.php-cite=22.html. 29 November 2019.

Vander Wall, S.B., 1997. Dispersal of singleleaf pinon pine (*Pinus monophylla*) by seed-caching rodents. *Journal of Mammalogy*, *78*(1), pp.181-191.

Vogl, C., Karhu, A., Moran, G. and Savolainen, O., 2002. High resolution analysis of mating systems: inbreeding in natural populations of *Pinus radiata*. *Journal of Evolutionary Biology*, 15(3), pp.433-439.

Wahlund, S., 1928. Zusammensetzung von Populationen und Korrelationserscheinungen vom Standpunkt der Vererbungslehre aus betrachtet. *Hereditas*, *11*(1), pp.65-106.

Waples, R.S., 2006. A bias correction for estimates of effective population size based on linkage disequilibrium at unlinked gene loci. *Conservation Genetics*, 7(2), pp.167-184.

Weisberg, P.J. and Ko, D.W., 2012. Old tree morphology in singleleaf pinyon pine (*Pinus monophylla*). *Forest Ecology and Management*, *263*, pp.67-73.

Wells, P.V., 1983. Paleobiogeography of montane islands in the Great Basin since the last glaciopluvial. *Ecological Monographs*, *53*(4), pp.341-382.

West, N.E., 1984. Successional patterns and productivity potentials of pinyon-juniper ecosystems. *Wildland Resources Faculty Publications*, 1688.

Williams, A.P., Allen, C.D., Macalady, A.K., Griffin, D., Woodhouse, C.A., Meko, D.M., Swetnam, T.W., Rauscher, S.A., Seager, R., Grissino-Mayer, H.D. and Dean, J.S., 2013. Temperature as a potent driver of regional forest drought stress and tree mortality. *Nature climate change*, 3(3), pp.292-297.

Williams, C.G., 2010. Long-distance pine pollen still germinates after meso-scale dispersal. *American Journal of Botany*, *97*(5), pp.846-855.

Zhivotovsky, L.A., 2015. Relationships between Wright's F ST and F IS statistics in a context of Wahlund effect. *Journal of Heredity*, *106*(3), pp.306-309.

CONCLUSION

The use of next-generation sequencing and the inclusion of multiple species in population-level studies resulted in the discovery of the Southwestern pinyon pine syngameon. Within the syngameon, cryptic hybrids were detected, in which advanced generation backcrossed individuals appeared morphologically similar to their parental species. This syngameon appears to be an edge-range syngameon whose participants remain distinct at species' range cores but hybridize in areas of sympatry at range margins. Admixed individuals have likely replaced a majority of the parental species in these sympatric populations through a process called genetic swamping. Syngameon participation has potentially resulted in adaptive introgression and the sequential creation of two drought-tolerant lineages, one of which is now a confirmed independent species that may be of conservation concern. Future studies will need to measure hybrid fitness to determine if they have a selective advantage. Additionally, once the pinyon genome is annotated, researchers can explore which genes are being introgressed across species barriers and if they are conferring drought tolerance. With the predicted contractions of suitable pinyon habitat, the future of this syngameon is uncertain. Syngameonic gene flow could allow for the introgression of adaptive traits and save declining species from climate change. However, if ranges contract out of sympatry or if species go extinct, it is not known how this loss would affect the existence of the other species. Conservation managers should seriously consider interspecific gene flow as a conservation tool to increase genetic diversity, but careful studies on potential outbreeding depression need to be conducted before introducing heterospecific gametes. Conservation laws must change to protect hybridizing species and should recognize syngameons as a whole, while also acknowledging their participants as individual species.