2021

Phenotypic characterization of Table Mountain (Pinus pungens) and pitch pine (Pinus rigida) hybrids along an elevational gradient in the Blue Ridge Mountains, Virginia

Alexander L. Brown
Virginia Commonwealth University

Phenotypic characterization of Table Mountain (*Pinus pungens*) and pitch pine (*Pinus rigida*) hybrids along an elevational gradient in the Blue Ridge Mountains, Virginia

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science at Virginia Commonwealth University

by

Alexander Louis Brown
Bachelor of Science in Environmental Biology
State University of New York College of Environmental Science and Forestry, 2016
Associate of Science in Environmental Science
Corning Community College, 2013

Director: Andrew J. Eckert, Ph.D.
Associate Professor, Department of Biology

Virginia Commonwealth University
Richmond, Virginia
May 2021

Acknowledgements

I would like to thank the many people who have led me to where I am today. I especially would like to thank my advisor, Dr. Andrew Eckert, who has helped guide my way throughout this project and graduate school. I also would like to thank my committee members, Dr. Chris Gough, Dr. Catherine Hulshof, and Dr. Daniel McGarvey, for their support and input throughout this project. Additionally, I would like to thank all of the teachers and mentors I have had throughout my educational career. You have shown me, and countless other students, what it means to truly love what you do. I also would like to thank all the colleagues, current and former graduate students, and friends I have had the pleasure of working with throughout my career. I would like to thank the National Park Service for allowing me the opportunity to conduct my research in the beautiful landscape of Shenandoah National Park. Of course, I would like to thank my parents and family for always supporting me and for raising me to become the best version of myself. Lastly, but far from least, I would like to thank my wife, Brooke. You have been my biggest supporter, friend, and companion. You have encouraged me to pursue my passions and inspired me to work to my highest potential. You have always been there to support me in my career and in life. Thank you.

Table of Contents

List of Figures

List of Tables

Abstract

PHENOTYPIC CHARACTERIZATION OF TABLE MOUNTAIN (*PINUS PUNGENS*) AND PITCH PINE (*PINUS RIGIDA*) HYBRIDS ALONG AN ELEVATIONAL GRADIENT IN THE BLUE RIDGE MOUNTAINS, VIRGINIA

By Alexander Louis Brown, B.S.

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science at Virginia Commonwealth University

Virginia Commonwealth University, 2021

Major Director: Andrew J. Eckert, Ph.D.
Associate Professor, Department of Biology

Hybridization has played a long-standing role in the evolution of both plant and animal species and allows for the sharing of genetic information between lineages. Here, potential hybridization of a species endemic to the Appalachian Mountains, Table Mountain pine (*Pinus pungens*), and pitch pine (*Pinus rigida*) was investigated along an elevational gradient, through the use of phenotypic measurements: cone length, cone width, and needle length. Phenotypes were used to identify hybrids in a three-tiered elevational sampling method at two sites in Shenandoah National Park with the use of linear discriminant analysis. It was found that hybridization between Table Mountain and pitch pine is relatively rare and varied by site and elevation. It was hypothesized that this lack of hybridization is due to environmental factors, which was further tested through use of climate data. The site where hybridization was highest was cooler and wetter. These factors may impact the pollen release of the focal species, causing overlap in pollen release timing and female cone receptivity, leading to increased instances of hybridization.

Introduction

Hybridization, the production of viable offspring from interspecific mating, has played a long-standing role in the evolution of plant and animal species, and allows for the sharing of genetic information among lineages (Whitney et al., 2010). Frequent hybridization can lead to introgression, the integration of genetic material from one species to another through repeated back-crossing (Mallet, 2005; Baack & Rieseberg, 2007). This recurrent gene flow can lead to increased diversity in hybrid populations, and over longer time scales can lead to local adaptation and new hybrid species (Goulet et al., 2017). Even if only a few individuals within a species are able to successfully hybridize, they can have large effects on the gene pool. These individuals provide a pathway for the transfer of genetic material from other species to enter their gene pool (Mallet, 2005).

Hybridization leads to varied evolutionary consequences such as introgression through back-crossing, hybrid speciation (i.e., hybrids evolve into a species reproductively isolated from parental species, e.g., *Pinus densata* Mast.), and the transfer of genetic material across species boundaries (Baack & Rieseberg, 2007; Mao et al., 2009). In some instances, hybrids resulting from these processes may have a fitness advantage over their parental species due to environmental selection or hybrid vigor (Buchholz, 1945; Shull, 1952; Whitney et al., 2010; Groszmann et al., 2011).

The ability to detect and study hybridization in pines is commonly observed using phenotypic traits, as even closely related species of pines usually look different phenotypically (Zobel, 1969; Smouse & Saylor, 1973; Garrett, 1979; Goulet et al., 2017). The use of phenotypic measurements to identify hybrids has been used in

previous studies using morphometric approaches with leaf size (Viscosi et al., 2012), cones and seeds (Mao et al., 2009), and needle size (Xing et al., 2014). Often, hybrid species tend to have phenotypes that are intermediate between the two parent species (Mao et al., 2009; Viscosi et al., 2012; Xing et al., 2014). A well-documented example of this can be seen in the Tibetan Plateau, where a species of pine which resulted from hybrid speciation, *Pinus densata* Mast., is phenotypically intermediate when compared to both parental species (Mao et al., 2009). Within the United States pine hybridization is widespread, with common examples including shortleaf (*Pinus echinata* Mill.) and loblolly pine (*Pinus taeda* L.) hybrids along the east coast (Tauer et al., 2012), and ponderosa pine (*Pinus ponderosa* Doug. ex Laws.) hybrids with other yellow pines along the west coast (Conkle & Critchfield, 1988).

In this particular study, I looked to identify possible hybridization through the use of phenotypic measurements between Table Mountain pine (*Pinus pungens* Lamb.), and pitch pine (*Pinus rigida* Mill.), two closely related species located throughout the eastern United States (Zobel, 1969).

Table Mountain pine is a conifer tree species native to the eastern United States which has become a target of modern conservation efforts. Its distribution range has largely declined in recent history, primarily due to fire suppression throughout the 20th century, resulting in the species becoming endemic to the central and southern portions of the Appalachian Mountains (Whittaker, 1956; Zobel, 1969; Jetton et al., 2015). Table Mountain pine is classified as a fire-dependent species because it requires fire to release its seeds and maintain its population. A key identifiable feature of Table Mountain pine is the large serotinous cones, which remain closed on the tree even after

reaching maturity and only open due to exposure to extreme heat, which typically occurs in the form of fires. Due to their serotinous cones and their relatively slow growth rates compared to other conifers, Table Mountain pine requires specific environmental conditions to persist. Table Mountain pine persists on dry, steep, exposed western and southern facing ridges, which are particularly susceptible to fire (Zobel, 1969). Without the presence of regular fire, Table Mountain pine will likely continue declining in population size (Sutherland et al., 1995).

Throughout the Appalachian Mountains, Table Mountain pine is also found growing alongside other native pine species, including pitch pine. Pitch pine is found widely throughout the eastern United States, and commonly alongside Table Mountain pine in the Appalachian Mountains (Whittaker, 1956; Callaway et al., 1987). Pitch pine is typically found on shallow soils, where other species are unable to persist, particularly along steep slopes or ridges. It occupies warm, dry environments, which can be observed through its presence on south and west facing slopes. Pitch pine can also be found throughout a variety of elevational ranges. It is commonly found at lower elevations in the northern portions of its range, and can be found at increasingly higher elevations, especially in the Appalachian Mountains, throughout the southern portion of its range (Little & Garrett, 1990).

There are distinguishable characteristics in needle and cone sizes between Table Mountain pine and pitch pine. Table Mountain pine is recognized by its two thick and twisted needles per fascicle, which range from 3-8 cm in length. This species is also identified by its large-sized cones (4-10 cm long and wide) which have broad, sharp upward curving spines, and grow in whorls of 3-4 cones. Pitch pine has been

considered one of the more variable pines in terms of phenotypic appearance (Harlow & Harrar, 1941; Ledig et al., 2015). It can be recognized by its three thin, long (up to 15 cm) needles per fascicle. This species typically has smaller cones than Table Mountain pine, ranging from 3-8 cm long and wide, which also have a flat base when opened (Gucker, 2007; Reeves, 2007). Also, pitch pine is not as fire dependent as Table Mountain pine is, as it can be found with both serotinous as well as non-serotinous (open) cones.

Table Mountain pine has been observed throughout the elevation range of 305-1220 m, although its density is greatest at higher elevations, especially in the southern and middle portions of its natural geographic range (Whittaker, 1956; Zobel, 1969). Pitch pine, however, tends to be found at lower elevations but will also grow at higher elevations, thus creating mixed stands with Table Mountain pine (Whittaker, 1956). Pitch pine can be found in abundance at elevations up to 750-950 m, but decreases in density and is replaced by Table Mountain pine at these higher elevations (Whittaker, 1956). These distribution patterns can be observed throughout the Appalachian Mountains, especially those located on the high elevation ridges of Shenandoah National Park. Here, both of these species grow in direct proximity to one another **(Figure 1)**.

Table Mountain and pitch pine are found in close spatial proximity, are closely related evolutionarily, and exhibit little divergence in climatic niche characteristics (Bolte, 2017). These two species should therefore hybridize. This has yet to be shown conclusively. I hypothesized that there is a hybrid zone at the range where these two species were found in close proximity, which could be identified using phenotypic

measurements.

Hybridization among species is not a rare occurrence, and occurs at some level in 25% of plant species (Mallet, 2005). In particular, hybridization can be observed between many different species of pines, with 95% of species within the genus *Pinus* able to successfully produce hybrid offspring in a large scale hybridization study (Critchfield, 1975). Hybridization among conifer species is common; therefore, when the ranges of two or more species overlap, especially for closely related species, it often leads to the development of hybrid zones (Smouse & Saylor, 1973; Garrett, 1979; Critchfield, 1986; Delgado et al., 2007). For some pine species, hybridization is geographically extensive, with resulting transitional zones being found outside of the primary range boundaries for the hybridizing species (Smouse & Saylor, 1973; Xing et al., 2014). However, the opposite pattern can be observed in some species of pines as well, which exhibit geographically restricted hybrid zones that are not geographically extensive (Gao et al., 2012).

Due to Table Mountain pine tending to replace pitch pine at the highest elevations, while the pitch pine tends to be more dominant at the lower elevations, it is expected that there will be a hybrid zone, if there is one at all, found between these species beginning at the middle elevations. This is the area where there will be the most overlap between elevational ranges, and both species tend to be found in relatively high abundance here. It has also been shown in other systems that hybrids tend to favor intermediate environments (Smouse & Saylor, 1973; Garrett, 1979; De La Torre et al., 2014).

In order to test this, I used phenotypes to identify hybrids in a three-tiered elevational sampling method at two sites in Shenandoah National Park. A complete lack of hybridization between these species would imply some sort of extrinsic or intrinsic barrier to hybridization, which may be related to divergent phenological schedules for pollen release (Zobel, 1969). If hybrids are found only at intermediate elevations, this would imply intermediate elevations are more favorable compared to lower and higher elevations, pointing to extrinsic selection favoring allelic combinations from each species in intermediate habitats, which has been detected in other conifer hybrids zones (Hamilton et al., 2013; De La Torre et al., 2014). These findings are of interest from a conservation standpoint as well. The analyses in this study can help provide insight to the possible hybridization zones of Table Mountain and pitch pine, which is a pivotal step in assessing conservation efforts for endemic species, ultimately leading to improved conservation efforts.

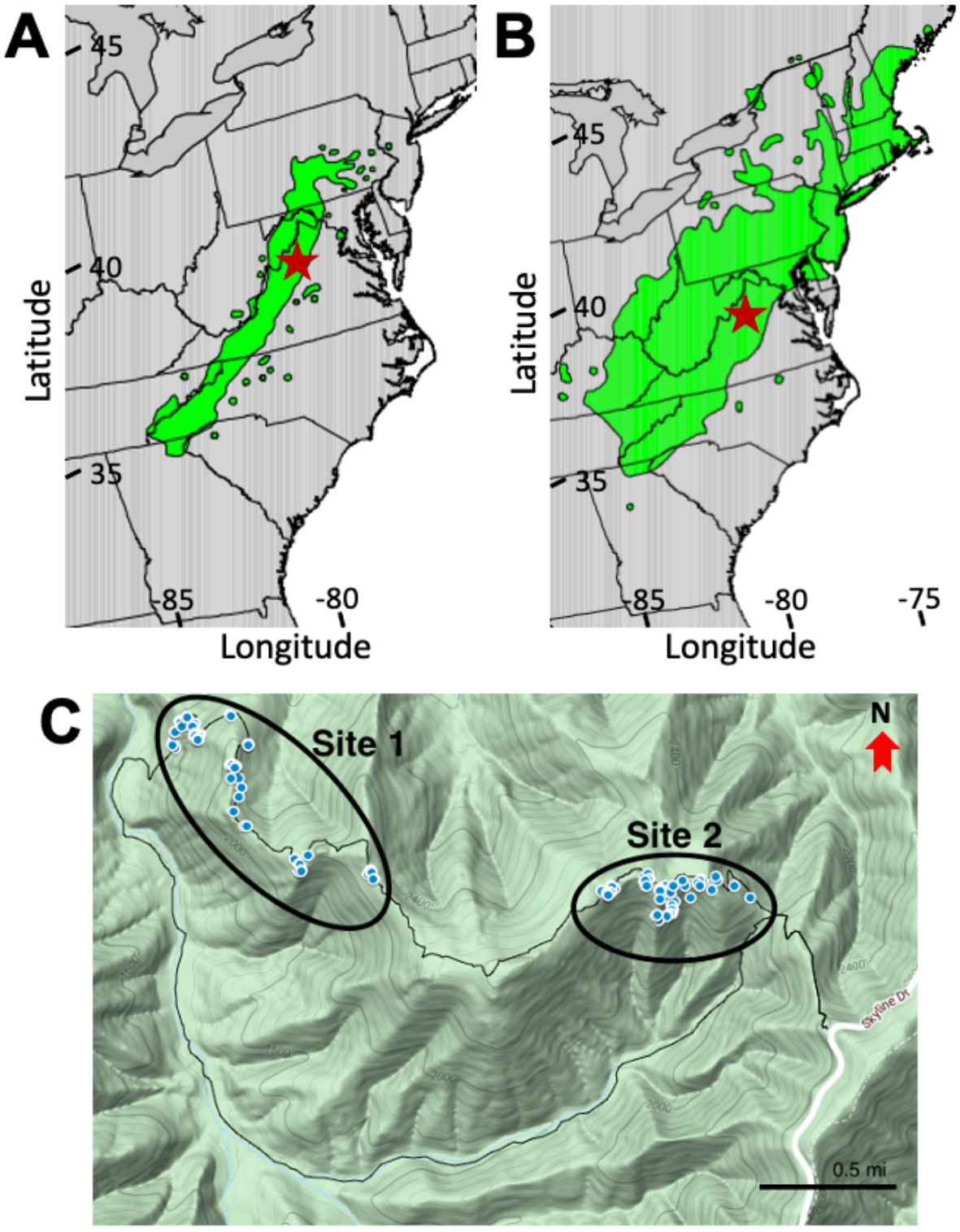

Figure 1: Geographical range of **A:** Table Mountain pine (*P. pungens*) and **B:** Pitch pine (*P. rigida*), throughout the northeastern United States. Sampling site is marked with the red star. **C)** Our two sample sites along Brown Mountain trail off of Skyline Drive within Shenandoah National Park. Collected sample locations are shown in blue circles. Trailhead Coordinates: (38.292736, -78.658073) Free Union, VA 22940.

Methods

Sampling Locations

All sampling occurred within Shenandoah National Park within the Big Run watershed. Sampling occurred along Brown Mountain trail (Trailhead coordinates: 38.29259, -78.65797), and its connection to the Rockytop trail (Trailhead coordinates: 38.30655, -78.70374). Two sample sites were used along this trail (Center coordinates: **Site 1:** 38.30530, -78.69661; **Site 2:** 38.29918, -78.66890; **Figure 1**). Both sites have a history of regular fire, which has since allowed for the abundant growth of both species to reach maturity (Lafon et al., 2017). Both species were present along steady elevational gradients across both sites. Site 1 spanned a lower elevational range than site 2 and had an aspect more west facing, but both sites had similar slope percentages **(Table 1)**. Site 1 and site 2 were separated by 1.5 miles with little to none of either species found in this area. The area between sites was exposed to a large fire in 2016 which has not allowed the full recovery of either Table Mountain or pitch pine here, providing a geographical barrier between sites (National Park Service, 2017). The study location and following study design was successfully approved by the National Park Service. (Permit #: SHEN-2020-SCI-0017).

Table 1: Sampling site characteristics.

	Site 1	**Site 2**
Elevation Range (m)	400 – 640	730 – 850
Percent Slope (%)	20 – 25	20 – 25
Aspect (0-360°)	≈ 185 – 280 (S-W)	≈ 175 – 220 (S-SW)

Study Design

Sample collection occurred in two steps: 1) Pure samples of Table Mountain and pitch pine needles and cones were collected. 2) Pines were randomly sampled at three separate elevational ranges along transects (low, mid, high). This design was then replicated at the two different sites along Brown Mountain trail.

Cone and needle samples from the trees were collected to assess three different phenotypic traits: needle length, cone length, and cone width. These traits are generally visually distinguishable between the two species, allowing for each species to be properly identified in the field, and are traits typically used for taxonomic identification.

Initial sample collection occurred during March-May 2020, where pure samples of Table Mountain and pitch pine were collected. All sampled tree coordinates and elevations were recorded while standing at the base of the tree using a GPS locating device. Approximately 3-4 fascicles of needles and 2-3 opened cones were collected per sampled tree. Cones were only taken from trees which had at least five or more cones so not all cones were removed from a tree. This also ensured sampled trees were reproductively mature. Each needle fascicle was collected from a different side of the crown to ensure representative needle samples per tree. In order for a cone to be collected, the cone had to be fully opened, with all scales fully separated from each other. This was done for 20-25 samples of each species at each site for the initial sampling. These samples spanned the entire elevational range of the respective site in order to have baseline phenotypic measurements representative of each species at each site.

After this initial collection, trees were randomly sampled along a set of elevational transects during June-August 2020 (61 total samples at site 1, 56 total samples at site 2). This was completed along three transects per site (low, middle, high). The elevational range for each transect was determined based on the initial pure sample range of collection. This resulted in transect elevations of the following ranges: **Site 1:** 420 m, 520 m, 620 m; **Site 2:** 730 m, 790 m, 850 m.

Pure Samples

Pure samples were identified based on field observations and known species identification characteristics (Gucker, 2007; Reeves, 2007). Pure Table Mountain pine samples that were collected had two thick spiral needles per fascicle, along with cones that had sharp, broad spikes along the scales, and were growing in whorls along the branches. The opened cones of Table Mountain pine are also not typically flat along the branch on which they are growing and will grow tightly around the branch.

Pure pitch pine samples were identified by the presence of three needles per fascicle, typically slender and long (≈10cm). Their cones do not typically grow in whorls along the branches and have flat bottoms when opened. Any samples in the field which did not have these distinct characteristics or that had any combination of characteristics between the two species criteria were not collected in the pure samples.

Pure sample were then analyzed using principal component analysis (PCA) to summarize multivariate differences in their phenotypic measurements and confirm it was possible to differentiate species **(Figure 2)**. PCA was completed using the *stats ver. 3.6.1* package in R with data scaled and centered (R Core Team, 2019), and 95% envelopes were used to inspect clusters.

In order to assess correspondence between taxonomy and groups observed in the PCAs, analysis of similarity (anosim) from the *vegan ver. 2.5-7* package in R was used (R Core Team, 2019; Oksanen et al., 2020). The Bray-Curtis dissimilarity measure along with 999 permutations were used, with a significance threshold of 0.05, to test the hypothesis that taxonomic groups were significantly dissimilar. This test shows the similarity between groups, or in our case, between species with respect to variation within the samples. An *R* value of 1 means that the groups are highly dissimilar, with a value of 0 meaning the groups are highly similar.

ImageJ Analysis

Cone length, cone width, and needle length were measured via ImageJ image analysis software (Rasband, 1997). An average was taken of the 2-3 cones and 3-4 fascicles of needles per tree. Photos of sampled cones and needles were taken from a consistent distance, with a scaled ruler also in the frame. A macro was used to correct the image type, sharpen the edges, and reduce shadows. Then, a threshold level was manually adjusted for each image to accurately outline each sample. The Feret's diameter (maximum caliper distance) was measured on all cones and needles, and the MinFeret (minimum caliper distance) was additionally measured on all cones. Feret's diameter is calculated by measuring the maximum distance between any two points of the selected object, which equates to the length. MinFeret then calculates the longest distance between any two points which are perpendicular to the Feret's diameter measurement, which equates to the width.

Linear Discriminant Analysis

In order to assess the presence of hybrids at each site, linear discriminant function analysis (LDA) was used. All LDAs were completed using the *MASS ver. 7.3-51.4* package in R (Ripley & Venables, 2002; R Core Team, 2019). The LDA visual output charts found in the supplement were completed using JMP Statistical Software (SAS Institute Inc., 1989). The pure samples of each species were used as training data for the LDA model, with the randomly sampled trees then used as the prediction data. This resulted in a prediction probability assignment for each sample. Samples that were assigned a prediction probability of less than 0.95 were considered hybrids. Samples assigned with a prediction probability of greater than 0.95 were considered pure species.

Climate Data

After samples were analyzed for possible hybridization, climate differences were assessed between sites using publicly available climate data from ClimateNA ver. 6.40 (Wang et al., 2016). Locally downscaled annual and monthly climate data for each sample coordinate along with its corresponding elevation was downloaded via ClimateNA (scale-free point location downscaled from 800 x 800 m resolution). Climate data were then analyzed using PCA via the *stats ver. 3.6.1* package in R with data scaled and centered (R Core Team, 2019). The climatic variable definitions for all present climate variables can be found in the supplemental section **(Table S1)**.

Results

Pure Samples

Pure species differed phenotypically at each sample site. Both sites had a high degree of variation explained by the first principal component (**Site 1:** 86.72%; **Site 2:** 76.06%; **Figure 2A and 2B**). Both site 1 and site 2 also had similar factor loading values for the phenotypic traits, with relatively equal weights applied to cone width, cone length, and needle length for the first PC axis **(Figure S1A and S2A)**. For both sites, cone length and cone width had positive loading values, with needle length having a negative loading value. The Table Mountain pine cluster was found in the positive direction, and pitch pine cluster was found in the negative direction on the first PC axis. This is consistent with typical species' identifying features (i.e., large cones and short needles for Table Mountain pine and small cones and long needles for pitch pine).

At site 1, pure species clustered into distinct groupings along the first PC axis. A similar pattern was observed for site 2, but there was a small area of overlap along the first PC, indicating more similarity between species at site 2 relative to site 1. Table Mountain pine tended to have tighter clustering than pitch pine at both sites as well, which is primarily observable along the second PC axis in both sites. For both sites, the second PC axis was comprised of negative loadings values of cone length, cone width, and needle length. The larger range along the second PC axis for pitch pine is consistent with species characteristics, as pitch pine can be quite variable in terms of phenotypes.

Analysis of similarity was used to test the hypothesis that groupings along PC1 corresponded to taxonomic identities assigned to each sample. Both sites showed high

dissimilarity values between species that were also statistically significant (**Site 1:** R = 0.9752, p = 0.001; **Site 2:** R = 0.7675, p = 0.001). Thus, Table Mountain and pitch pine pure samples were dissimilar from one another and could be distinguished in multivariate space using phenotypic measurements.

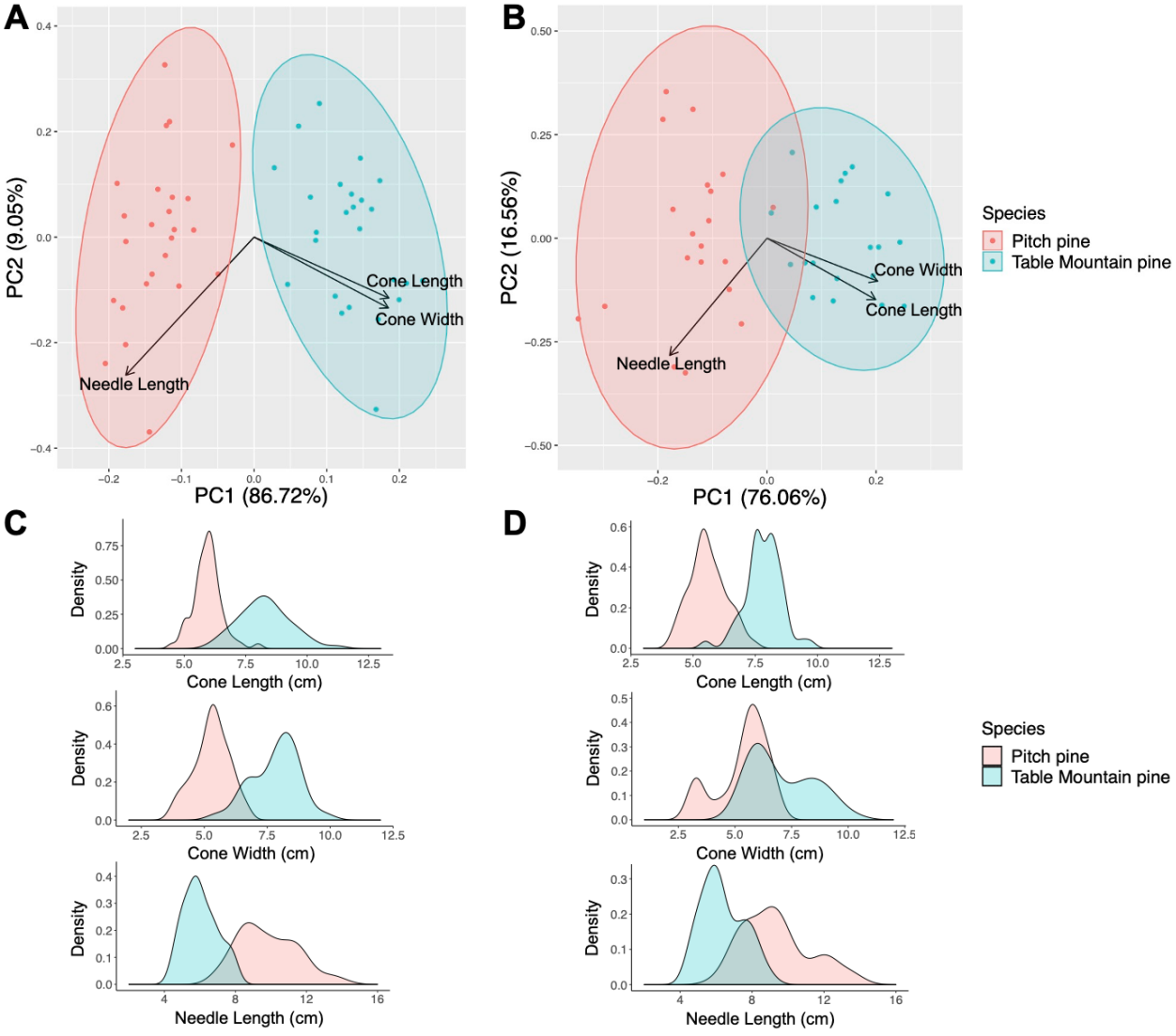

Figure 2: Principal component analysis (PCA) (**A:** Site 1; **B:** Site 2), along with kernel density plots of the phenotypic variables (**C:** Site 1; **D:** Site 2) from the pure samples.

Presence of Hybrids

The number of hybrids detected was small and varied by site **(Figure 3)**. The LDA at site 1 resulted in no hybrids, while at site 2, a few instances of hybridization were found. At site 1, all samples resulted in a prediction probability that was greater than 0.95, or less than 0.05, categorizing each as a pure Table Mountain or pitch pine sample. In fact, most of the samples had a prediction probability of 1 or 0, meaning they were accurately assigned to a species with full confidence. The full data output from each site LDA can be found in the supplement **(Table S2 and S3)**. At the low elevations most of the samples were classified as pitch pine. At the middle elevations there was an increase in the number samples assigned to Table Mountain pine. At the high elevations a relatively even number of both species were present, with more Table Mountain pine predicted. This pattern is expected, as Table Mountain pine tends to replace pitch pine at upper elevations.

The presence of hybridization was different at site 2. For the most part, samples at all elevations were accurately assigned to Table Mountain or pitch pine with probabilities close to 1 or 0 (Table Mountain pine = 1; pitch pine = 0; F1 hybrid = 0.5). In the high elevation group, all samples were accurately assigned to a species with high confidence (probability > 0.95 or < 0.05). A relatively equal number of both Table Mountain and pitch pine samples were found here. In the low elevation sample group, one sample was classified as a hybrid, with a classification probability of 0.90. The middle elevation group at site 2 had the highest number of hybrids found, with a total of 6 of the 17 samples classified as hybrids. All hybrids appeared to be advanced generation hybrids with probabilities of 0.945, 0.84, 0.39, 0.24, 0.12, and 0.09. There

does not appear to be a directional pattern of hybridization in these species, with 4 of these hybrids classified closer to pitch pine, and the other 2 classified closer to Table Mountain pine. Additionally, the elevational distribution of species at this site differed from site 1. Generally, pitch pine is more dominant at lower elevations, and is then replaced by Table Mountain pine at upper elevations. At this site however, the opposite pattern was seen. At the lower elevation points, most of the samples were classified as Table Mountain pine. There was a relatively even number of both species found at the middle elevation, and at the higher elevation is where the model identified more pitch pines. This was an interesting finding, and suggests that pitch pine was at least as prevalent, if not more prevalent, at the highest elevation points at site 2.

To assess the observed patterns of hybrid occurrence differing between sites, pure samples from each site were used to predict the pure samples of the other site. When site 1 pure data was used as training data for an LDA to predict the site 2 pure data, it was found that 3 samples of the site 2 pure samples were incorrectly predicted, and another 4 samples were predicted as hybrids **(Table S4)**. Of those 4 predicted hybrid samples, it was found that 1 sample of pitch pine was predicted with a low probability (0.72), and 3 samples of Table Mountain pine were incorrectly predicted as pitch pine. In addition, another 3 Table Mountain pine samples were predicted as hybrids (probabilities of 0.88, 0.72, 0.61). The reverse of this was tested as well. When site 2 pure data was used as training data to predict site 1 pure samples, 3 of the pitch pine samples were found to be predicted as hybrids (probabilities of 0.86, 0.78, 0.59) **(Table S5)**. All Table Mountain pine samples were accurately predicted in this instance. Overall, these findings in combination with the PCA of the pure data for both sites and

the analysis of similarity findings, suggest that the site 2 pure samples tend to be more phenotypically similar between species than site 1 samples, but that this similarity is unlikely to produce artifactual detection of hybrids.

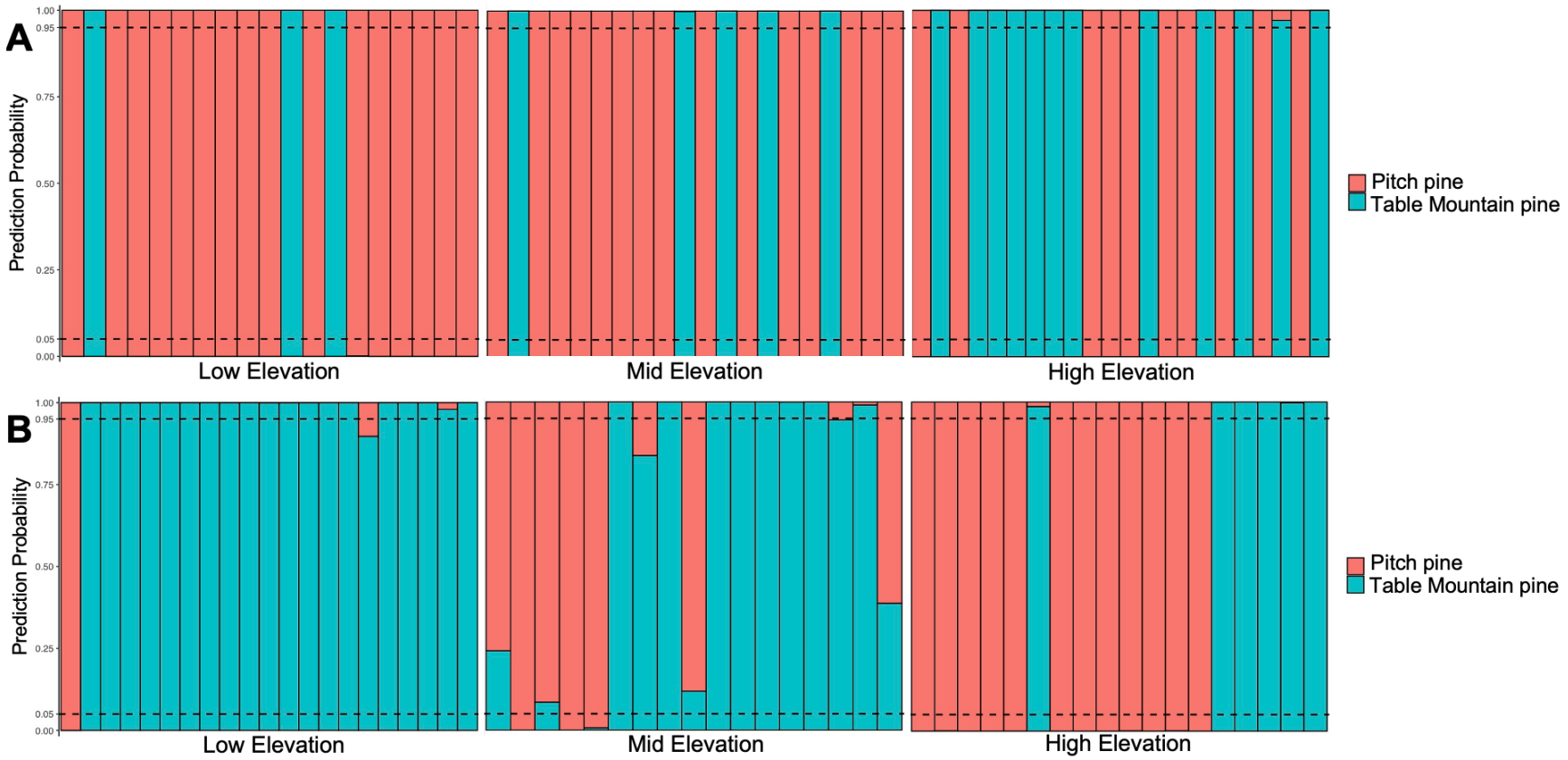

Figure 3: Sample prediction probabilities from the linear discriminant analysis test (**A:** Site 1; **B:** Site 2). Each individual bar reflects a sample's predicted probability assignment to either pitch pine (red) or Table Mountain pine (blue). Probability assignments which fell between the range of 0.05 and 0.95 were considered hybrids.

Climate Analysis

Annual climate data extracted from ClimateNA for each sample's coordinates from both sites were analyzed using PCA **(Figure 4)**. The first PC axis explains nearly all variation in the data (99.7%). Site 1 and site 2 data clustered separately from each other, with the majority of clustering explained by the first PC. The first PC was primarily determined by the annual climate factors: degree days above 5°C (DD5), mean annual precipitation (MAP), degree days below 18°C (DD_18), and degree days above 10°C and below 40°C (DD1040) **(Figure S3A)**. The second PC axis was largely determined by the climatic moisture deficit (CMD), reference evaporation (Eref), and mean annual precipitation (MAP) **(Figure S3B)**. Many of the site 2 samples, along with the samples which were classified as hybrids according to the LDA, tended to be found in the area of PC space defined by increasing degree days below 18°C (DD_18), mean annual precipitation (MAP), and May to September precipitation (MSP), meaning site 2 is exposed to colder and wetter annual conditions than site 1.

A similar PCA analysis was conducted using monthly climate data and can be found in the supplemental section **(Figure S4)**. In this case, the first PC axis explained nearly all the variation in the data (99.75%). The first PC axis was largely determined by an aggregation of monthly precipitation values, with the months of May, July, and October contributing the highest loading values **(Figure S5A)**. The samples from site 2 tended to be clustered towards the increasing precipitation values for all months compared to site 1 samples. Some of samples which were classified as hybrids were located in the portion of PC space defined by increasing precipitation for the months of September and October, while others were located towards the increasing loading

values for the months of January and May precipitation. Seasonal precipitation did not appear to predict the occurrence of hybrids, although increased annual precipitation at site 2 relative to site 1 may partially explain the increase in hybridization found at this site.

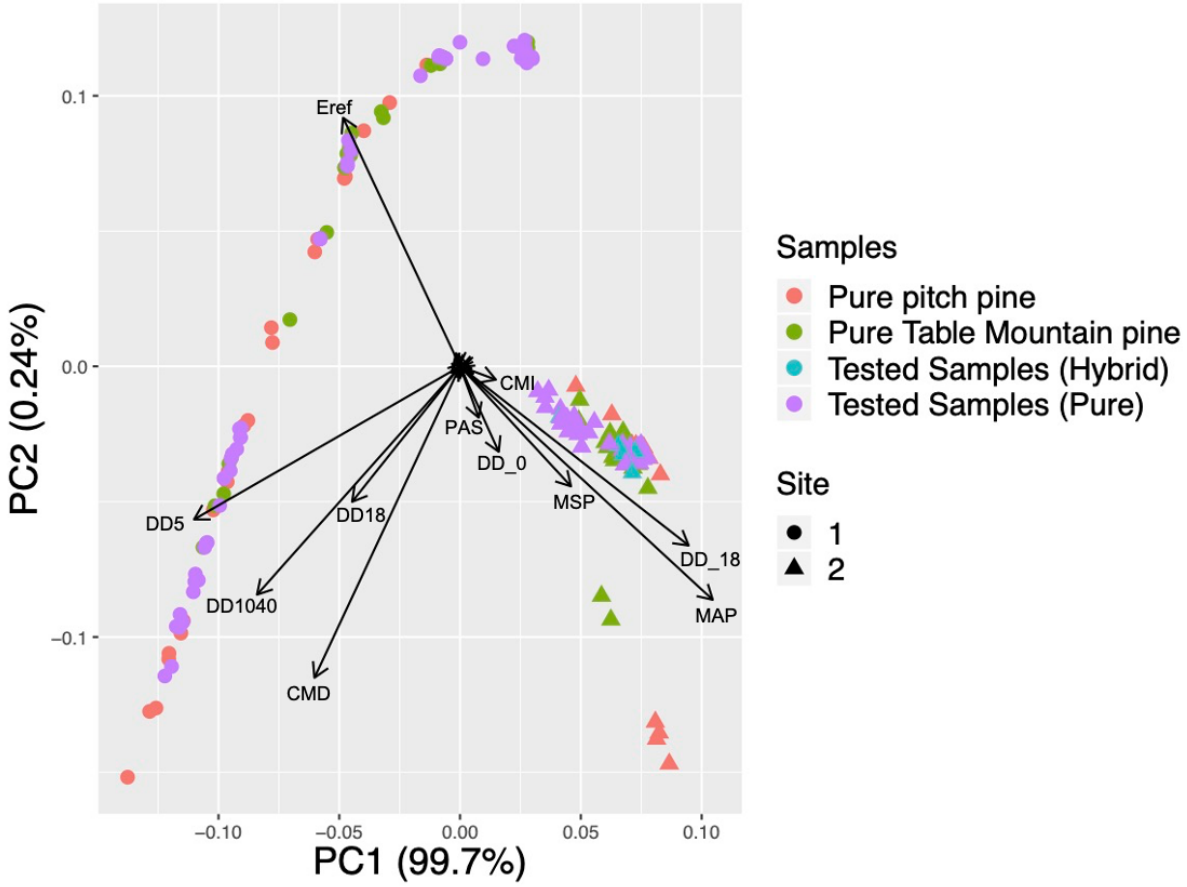

Figure 4: Principal component analysis (PCA) of annual climatic variables (ClimateNA) for each sample coordinate location. Variable definitions are located in Table S1, along with loadings values in Figure S3.

Discussion

Hybridization

Hybridization between Table Mountain and pitch pine was relatively rare and varied by location and elevation. The few instances of hybridization found appeared to be site and environment dependent and were all classified as being more consistent with advanced generation hybrids. Recent hybrids (e.g., F1 hybrids) were not found, as samples with prediction probabilities that were close to 0.5 were not found. If recent hybrids were present it would be expected this would have been the case as first generation hybrids often have phenotypic characteristics intermediate between the parental phenotypes (Zobel, 1969; Smouse & Saylor, 1973; Thompson et al., 2021). This pattern is common in hybrid zones for long-lived plants (Field et al., 2011), where advanced generation hybrids, including extensive introgression, is common (Hodges et al., 1996; Lexer et al., 2004). The data was consistent with rare hybridization events, likely including extensive back-crossing, under certain site conditions. There was also no directionality to these events, with hybrids almost equally likely to be more like either parent species. Taken together, these results are most consistent with site-dependent, rare hybridization events.

As time since divergence from a common ancestor increases between species, so does hybrid incompatibility (Orr & Turelli, 2001). It is possible that post-zygotic barriers have remained present and are preventing the hybridization of the focal species. However, post-zygotic barriers seem unlikely, as they are rarely identified in pines when experimentally crossed (Critchfield, 1975). Furthermore, Table Mountain and pitch pine are sister species, recently diverged from a common ancestor

(approximately 1.5 million years ago), which is relatively recent in terms of the timescale which barriers to gene flow typically evolve (Hernández-León et al., 2013). Many examples of pines which actively hybridize have much deeper divergence times and are often less phylogenetically clustered (Critchfield, 1975; Hernández-León et al., 2013). The lack of extensive hybridization between Table Mountain and pitch pine is thus likely due to environmental factors contributing to the establishment of pre-zygotic barriers and that when these factors vary or change, so does the existence of the barrier, which was implied by Zobel (1969).

Environmental Factors

Table Mountain pine typically releases pollen in mid-April, beginning slightly earlier than pitch pine (Zobel, 1969). Pitch pine typically releases pollen towards the end of April into early May (Zobel, 1969; Cho et al., 2003; Gucker, 2007). The early release of pollen from Table Mountain pine has been suggested as a primary factor in the lack of extensive hybridization between these species (Zobel, 1969). However, pollen release timing can be variable from year to year, hard to predict, and can even be variable across populations of trees (Whittet et al., 2017). I used historical climate data at the coordinate points from the collected samples in order to quantify environmental trends that may be impacting pollen release timing between Table Mountain and pitch pine. I was able to determine that site 2, where there appeared to be signs of hybridization, had greater annual precipitation as well as cooler temperatures relative to site 1. Temperature has been shown to play a significant role in determining the timing of pollen release in trees, while precipitation also plays a role (Fuhrmann et al., 2016). This can be seen in other pine populations which were exposed to warmer

temperatures, where trees shed pollen earlier than those which were exposed to cooler temperatures (Whittet et al., 2017). In addition, increased pollen concentrations are also associated with increased temperatures, and increased precipitation in the months leading up to pollen release (Fuhrmann et al., 2016).

It is possible that the cooler temperatures at site 2 are causing a delay in the typically early pollen release of Table Mountain pine, aligning it more with the timing of pitch pine pollen release, thus allowing greater chances for hybridization. For example, a delay in pollen release from Table Mountain pine could lead to longer periods of time when both species are actively shedding and receiving pollen, as pitch pine pollen receptivity and pollen release typically begin later than Table Mountain. As the climate continues to shift, the pollen release timings of these two species will likely change as well, which would then impact chances of further gene exchange. Thus, changing climates, especially changes to patterns of annual variation, could result in pulses of hybridization that vary wildly from year to year. An in-depth analysis of pollen release timing across multiple years would be needed to further test this idea.

Conservation

When assessing the conservation of endemic species, standing levels of gene diversity play an important role in the determination of the genetic future of a species (Ellstrand, 1992). Understanding the possible effects of hybridization in species such as this is highly important in terms of conservation practices, as hybridization leads to increased levels of genetic diversity, increased adaptation capabilities in changing environments, and increased genetic variation for fitness (Kremer et al., 2012; Aitken & Whitlock, 2013). However, hybridization can also have negative effects within

conservation contexts. For instance, concerns can arise when invasive species hybridize with local species, causing adverse effects on the local pure species (Huxel, 1999). This can become particularly problematic if hybridization occurs between a species that is rare, and one that is widespread, and can result in hybridization among a previously isolated (potentially protected) species (Rhymer & Simberloff, 1996).

Table Mountain pine's serotinous cones have persisted throughout its species history due to repeated fire (Zobel, 1969; Radeloff et al., 2004). Currently, contemporary fire suppression potentially makes this adaptation disadvantageous. Fire reduction has been shown to increase rates of hybridization in a similar system, including a fire-adapted (*P. echinata Mill*.) and less fire-adapted (*P. taeda L*.) pine species (Tauer et al., 2012; Stewart et al., 2015). This is of concern due to the loss of genetic integrity in these species, with increased instances such as this eventually leading to a loss in the amount of genetic diversity seen among species. In general, I did not observe patterns of converging phenotype similarities between the two species in either direction, showing that hybridization continues to be relatively rare. However, this could change as climate continues to change.

Conclusion

Table Mountain and pitch pine are found in close spatial proximity, are closely related evolutionarily, and exhibit little divergence in climatic niche characteristics (Bolte, 2017). Therefore, these two species should be hybridizing to some degree. I found that hybridization between these two species is relatively rare, with environmental factors driving this rarity. Historical climate data showed that the site which had signs of hybridization was found to be colder and wetter, on average. These environmental

factors may be impacting the pollen release timing of both species, leading to increased instances of hybridization. Moving forward, in order to properly assess and make the most informed conservation decisions with regards to endemic species, the possibility of introgression through ongoing hybridization between species should be considered.

Literature Cited

Aitken, S. N., & Whitlock, M. C. (2013). Assisted Gene Flow to Facilitate Local Adaptation to Climate Change. *Annual Review of Ecology, Evolution, and Systematics*, *44*(1), 367–388. https://doi.org/10.1146/annurev-ecolsys-110512-135747

Baack, E. J., & Rieseberg, L. H. (2007). A genomic view of introgression and hybrid speciation. *Current Opinion in Genetics & Development*, *17*(6), 513–518. https://doi.org/10.1016/j.gde.2007.09.001

Bolte, C. E. (2017). *Niche Conservatism Or Divergence: Insights Into the Evolutionary Histories of Pinus Taeda, Pinus Rigida, and Pinus Pungens*. https://doi.org/10.25772/4PYQ-0821

Buchholz, J. T. (1945). Embryological Aspects of Hybrid Vigor in Pines. *Science*, *102*(2641), 135–142. https://doi.org/10.1126/science.102.2641.135

Callaway, R. M., Clebsch, E. E. C., & White, P. S. (1987). A Multivariate Analysis of Forest Communities in the Western Great Smoky Mountains National Park. *American Midland Naturalist*, *118*(1), 107–120. https://doi.org/10.2307/2425633

Cho, Y. J., Kim, I. S., Kim, P., & Lee, E. J. (2003). Deposition of airborne pine pollen in a temperate pine forest. *Grana*, *42*(3), 178–182. https://doi.org/10.1080/00173130310016158

Conkle, M. T., & Critchfield, W. B. (1988). Genetic Variation and Hybridization of Ponderosa Pine. *In: Ponderosa Pine: The Species and Its Management, Washington State University Cooperative Extension*, 27–43.

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*, 99–105.

Critchfield, W. B. (1986). Hybridization and Classification of the White Pines (Pinus section Strobus). *Taxon*, *35*(4), 647–656. https://doi.org/10.2307/1221606

De La Torre, A. R., Wang, T., Jaquish, B., & Aitken, S. N. (2014). Adaptation and exogenous selection in a Picea glauca × Picea engelmannii hybrid zone: Implications for forest management under climate change. *New Phytologist*, *201*(2), 687–699. https://doi.org/10.1111/nph.12540

Delgado, P., Salas-Lizana, R., Vázquez-Lobo, A., Wegier, A., Anzidei, M., Alvarez-Buylla, E. R., Vendramin, G. G., & Piñero, D. (2007). Introgressive Hybridization in *Pinus montezumae* Lamb and *Pinus pseudostrobus* Lindl. (Pinaceae): Morphological and Molecular (cpSSR) Evidence. *International Journal of Plant Sciences*, *168*(6), 861–875. https://doi.org/10.1086/518260

Ellstrand, N. C. (1992). Gene Flow by Pollen: Implications for Plant Conservation Genetics. *Oikos*, 77–86.

Field, D. L., Ayre, D. J., Whelan, R. J., & Young, A. G. (2011). Patterns of hybridization and asymmetrical gene flow in hybrid zones of the rare Eucalyptus aggregata and common E. rubida. *Heredity*, *106*(5), 841–853. https://doi.org/10.1038/hdy.2010.127

Fuhrmann, C. M., Sugg, M. M., & Konrad, C. E. (2016). Airborne pollen characteristics and the influence of temperature and precipitation in Raleigh, North Carolina, USA (1999–2012). *Aerobiologia*, *32*(4), 683–696. https://doi.org/10.1007/s10453-016-9442-6

Gao, J., Wang, B., Mao, J. F., Ingvarsson, P., Zeng, Q. Y., & Wang, X. R. (2012). Demography and speciation history of the homoploid hybrid pine *Pinus densata* on the Tibetan Plateau: Speciation History of Pinus Densata. *Molecular Ecology*, *21*(19), 4811–4827. https://doi.org/10.1111/j.1365-294X.2012.05712.x

Garrett, P. W. (1979). *Species Hybridization in the Genus Pinus* (Vol. 436). Dept. of Agriculture, Forest Service, Northeastern Forest Experiment Station.

Goulet, B. E., Roda, F., & Hopkins, R. (2017). Hybridization in Plants: Old Ideas, New Techniques. *Plant Physiology*, *173*(1), 65–78. https://doi.org/10.1104/pp.16.01340

Groszmann, M., Greaves, I. K., Albertyn, Z. I., Scofield, G. N., Peacock, W. J., & Dennis, E. S. (2011). Changes in 24-nt siRNA levels in Arabidopsis hybrids suggest an epigenetic contribution to hybrid vigor. *Proceedings of the National Academy of Sciences*, *108*(6), 2617–2622. https://doi.org/10.1073/pnas.1019217108

Gucker, C. L. (2007). Pinus rigida. *In: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station, Fire Sciences Laboratory*. https://www.fs.fed.us/database/feis/plants/tree/pinrig/all.html

Hamilton, J. A., Lexer, C., & Aitken, S. N. (2013). Genomic and phenotypic architecture of a spruce hybrid zone (*Picea sitchensis* × *P. glauca*). *Molecular Ecology*, *22*(3), 827–841. https://doi.org/10.1111/mec.12007

Harlow, W. M., & Harrar, E. S. (1941). Textbook of dendrology. Covering the important forest trees of the United States and Canada. *Textbook of Dendrology. Covering the Important Forest Trees of the United States and Canada*.

Hernández-León, S., Gernandt, D. S., Pérez de la Rosa, J. A., & Jardón-Barbolla, L. (2013). Phylogenetic Relationships and Species Delimitation in Pinus Section Trifoliae Inferrred from Plastid DNA. *PLoS ONE*, *8*(7), e70501. https://doi.org/10.1371/journal.pone.0070501

Hodges, S. A., Burke, J. H., & Arnold, M. L. (1996). Natural Formation of Iris Hybrids: Experimental Evidence on the Establishment of Hybrid Zones. *Evolution*, *50*(6), 2504–2509.

Huxel, G. R. (1999). Rapid displacement of native species by invasive species: Effects of hybridization. *Biological Conservation*, *89*, 143–152.

Jetton, R. M., Crane, B. S., Whittier, W. A., & Dvorak, W. S. (2015). Genetic Resource Conservation of Table Mountain Pine (Pinus pungens) in the Central and Southern Appalachian Mountains. *Tree Planters' Notes*, *58*(1), 42–52.

Kremer, A., Ronce, O., Robledo-Arnuncio, J. J., Guillaume, F., Bohrer, G., Nathan, R., Bridle, J. R., Gomulkiewicz, R., Klein, E. K., Ritland, K., Kuparinen, A., Gerber, S., & Schueler, S. (2012). Long-distance gene flow and adaptation of forest trees to rapid climate change. *Ecology Letters*, *15*(4), 378–392. https://doi.org/10.1111/j.1461-0248.2012.01746.x

Lafon, C. W., Naito, A. T., Grissino-Mayer, H. D., Horn, S. P., & Waldrop, T. A. (2017). Fire history of the Appalachian region: A review and synthesis. *Gen. Tech. Rep. SRS-219. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station*, *219*, 1–97. https://doi.org/10.2737/SRS-GTR-219

Ledig, F. T., Smouse, P. E., & Hom, J. L. (2015). Postglacial migration and adaptation for dispersal in pitch pine (Pinaceae). *American Journal of Botany*, *102*(12), 2074–2091. https://doi.org/10.3732/ajb.1500009

Lexer, C., Heinze, B., Alia, R., & Rieseberg, L. H. (2004). Hybrid zones as a tool for identifying adaptive genetic variation in outbreeding forest trees: Lessons from wild annual sunflowers (Helianthus spp.). *Forest Ecology and Management*, *197*(1–3), 49–64. https://doi.org/10.1016/j.foreco.2004.05.004

Little, S., & Garrett, P. W. (1990). Pinus rigida Mill. Pitch pine. *Silvics of North America*, *1*, 456–462.

Mallet, J. (2005). Hybridization as an invasion of the genome. *Trends in Ecology & Evolution*, *20*(5), 229–237. https://doi.org/10.1016/j.tree.2005.02.010

Mao, J. F., Li, Y., & Wang, X. R. (2009). Empirical assessment of the reproductive fitness components of the hybrid pine Pinus densata on the Tibetan Plateau. *Evolutionary Ecology*, *23*(3), 447–462. https://doi.org/10.1007/s10682-008-9244-6

National Park Service. (2017). *Fire Ecology*. https://www.nps.gov/shen/learn/nature/fire-ecology.htm

Oksanen, J., Guillaume Blanchet, F., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., H. Stevens, M. H., Szoecs, E., & Wagner, H. (2020). *vegan: Community Ecology Package* (R package version 2.5-7) [Computer software]. https://CRAN.R-project.org/package=vegan

Orr, H. A., & Turelli, M. (2001). The Evolution of Postzygotic Isolation: Accumulating Dobzhansky-Muller Incompatibilities. *Evolution*, *55*(6), 1085–1094. https://doi.org/10.1111/j.0014-3820.2001.tb00628.x

R Core Team. (2019). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Radeloff, V. C., Mladenoff, D. J., Guries, R. P., & Boyce, M. S. (2004). Spatial patterns of cone serotiny in Pinus banksiana in relation to fire disturbance. *Forest Ecology and Management*, *189*(1–3), 133–141. https://doi.org/10.1016/j.foreco.2003.07.040

Rasband, W. S. (1997). ImageJ. *U.S. National Institutes of Health, Bethesda, Maryland, USA*. https://imagej.nih.gov/ij/

Reeves, S. L. (2007). *Pinus pungens*. https://www.fs.fed.us/database/feis/plants/tree/pinpun/all.html

Rhymer, J. M., & Simberloff, D. (1996). Extinction by hybridization and introgression. *Annual Review of Ecology, Evolution, and Systematics*, *27*, 83–109.

Ripley, B. D., & Venables, W. N. (2002). Modern Applied Statistics with S. *Fourth Edition. Springer, New York. ISBN 0-387-95457-0*. http://www.stats.ox.ac.uk/pub/MASS4

SAS Institute Inc. (1989). *JMP Pro, Version 15.2.0*.

Shull, G. H. (1952). Beginnings of the Heterosis Concept. *Heterosis*, 14–48.

Smouse, P. E., & Saylor, L. C. (1973). Studies of the Pinus rigida-Serotina Complex II. Natural Hybridization Among the Pinus rigida-Serotina Complex, P. Taeda and P. echinata. *Annals of the Missouri Botanical Garden*, *60*(2), 192. https://doi.org/10.2307/2395085

Stewart, J. F., Will, R. E., Robertson, K. M., & Nelson, C. D. (2015). Frequent fire protects shortleaf pine (Pinus echinata) from introgression by loblolly pine (P. taeda). *Conservation Genetics*, *16*(2), 491–495. https://doi.org/10.1007/s10592-014-0669-x

Sutherland, E. K., Grissino-Mayer, H., Woodhouse, C. A., Covington, W. W., Horn, S., Huckaby, L., Kerr, R., Kush, J., Moorte, M., & Plumb, T. (1995). Two Centuries of Fire in a Southwestern Virginia Pinus pungens Community. *In: Inventory and Management Techniques in the Context of Catastrophic Events: Altered States of the Forest; University Park, PA. Penn State University, Center for Statistical Ecology & Environmental Statistics*.

Tauer, C. G., Stewart, J. F., Will, R. E., Lilly, C. J., Guldin, J. M., & Nelson, C. D. (2012). Hybridization Leads to Loss of Genetic Integrity in Shortleaf Pine: Unexpected Consequences of Pine Management and Fire Suppression. *Journal of Forestry*, *110*(4), 216–224. https://doi.org/10.5849/jof.11-044

Thompson, K. A., Urquhart-Cronish, M., Whitney, K. D., Rieseberg, L. H., & Schluter, D. (2021). Patterns, Predictors, and Consequences of Dominance in Hybrids. *The American Naturalist*, *197*(3), E72–E88.

Viscosi, V., Antonecchia, G., Lepais, O., Fortini, P., Gerber, S., & Loy, A. (2012). Leaf Shape and Size Differentiation in White Oaks: Assessment of Allometric Relationships among Three Sympatric Species and Their Hybrids. *International Journal of Plant Sciences*, *173*(8), 875–884. https://doi.org/10.1086/667234

Wang, T., Hamann, A., Spittlehouse, D., & Carroll, C. (2016). Locally Downscaled and Spatially Customizable Climate Data for Historical and Future Periods for North America. *PLOS ONE*, *11*(6), e0156720. https://doi.org/10.1371/journal.pone.0156720

Whitney, K. D., Ahern, J. R., Campbell, L. G., Albert, L. P., & King, M. S. (2010). Patterns of hybridization in plants. *Perspectives in Plant Ecology, Evolution and Systematics*, *12*(3), 175–182. https://doi.org/10.1016/j.ppees.2010.02.002

Whittaker, R. H. (1956). Vegetation of the Great Smoky Mountains. *Ecological Monographs*, *26*(1), 1–80.

Whittet, R., Cavers, S., Cottrell, J., Rosique-Esplugas, C., & Ennos, R. (2017). Substantial variation in the timing of pollen production reduces reproductive synchrony between distant populations of *Pinus sylvestris* L. in Scotland. *Ecology and Evolution*, *7*(15), 5754–5765. https://doi.org/10.1002/ece3.3154

Xing, F., Mao, J., Meng, J., Dai, J., Zhao, W., Liu, H., Xing, Z., Zhang, H., Wang, X., & Li, Y. (2014). Needle morphological evidence of the homoploid hybrid origin of Pinus densata based on analysis of artificial hybrids and the putative parents, *Pinus tabuliformis* and *Pinus yunnanensis* . *Ecology and Evolution*, *4*, 1890–1902. https://doi.org/10.1002/ece3.1062

Zobel, D. B. (1969). Factors Affecting the Distribution of Pinus pungens, an Appalachian Endemic. *Ecological Monographs*, *39*(3), 303–333.

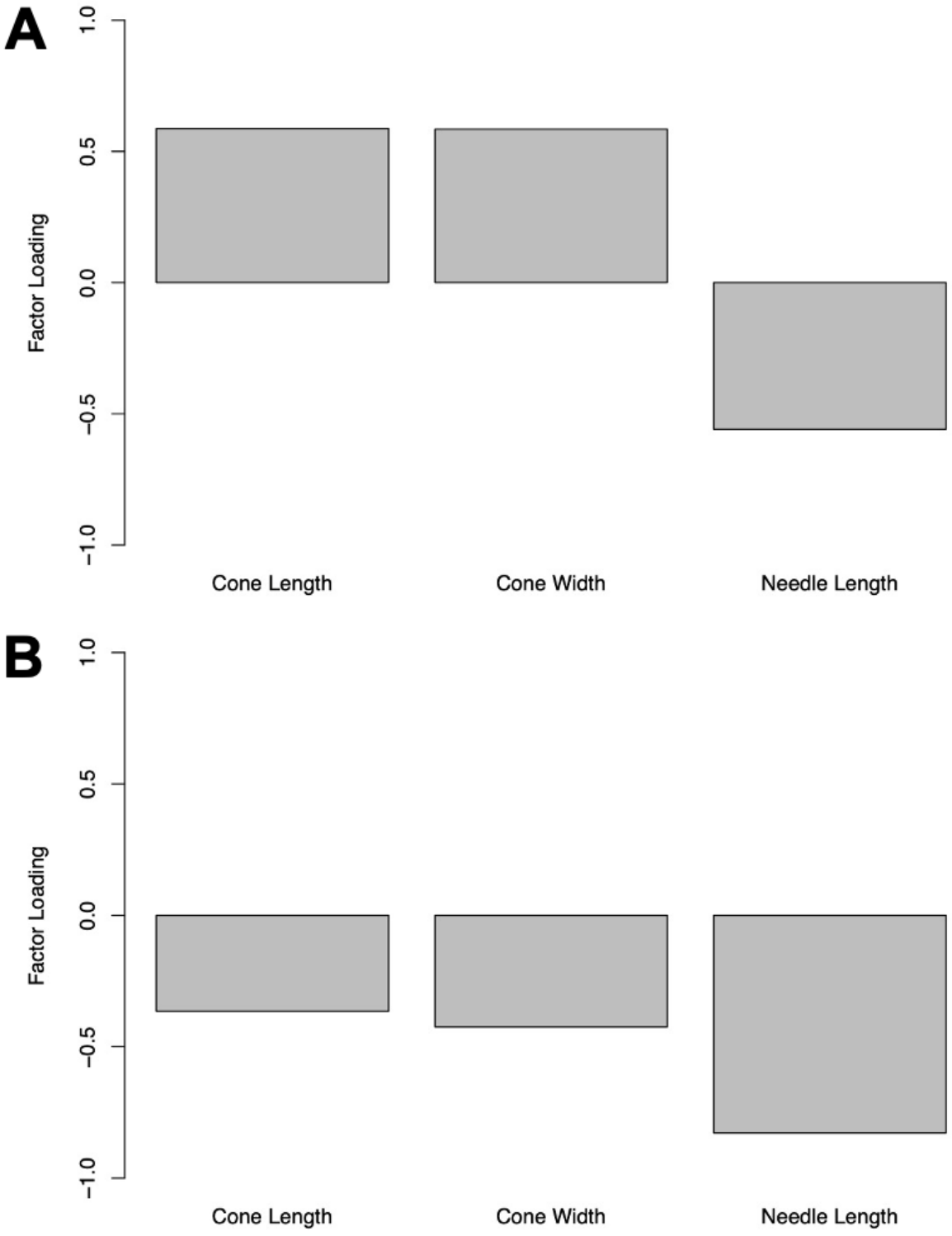

Figure S1: Loadings plots for site 1 pure sample PCA variables from Figure 2A. **A:** PC1 axis; **B:** PC2 axis.

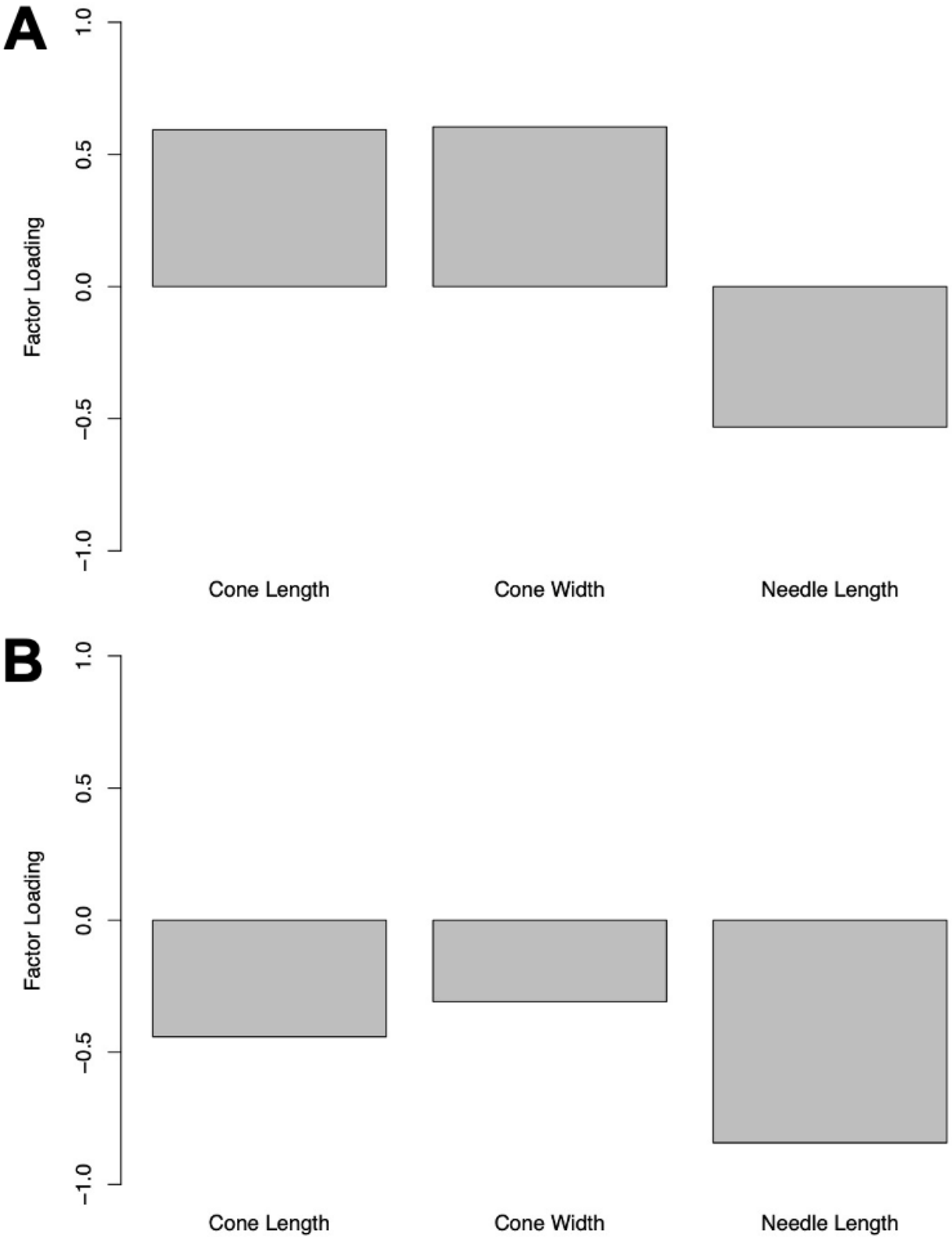

Figure S2: Loadings plots for site 2 pure sample PCA variables from Figure 2B. **A:** PC1 axis; **B:** PC2 axis.

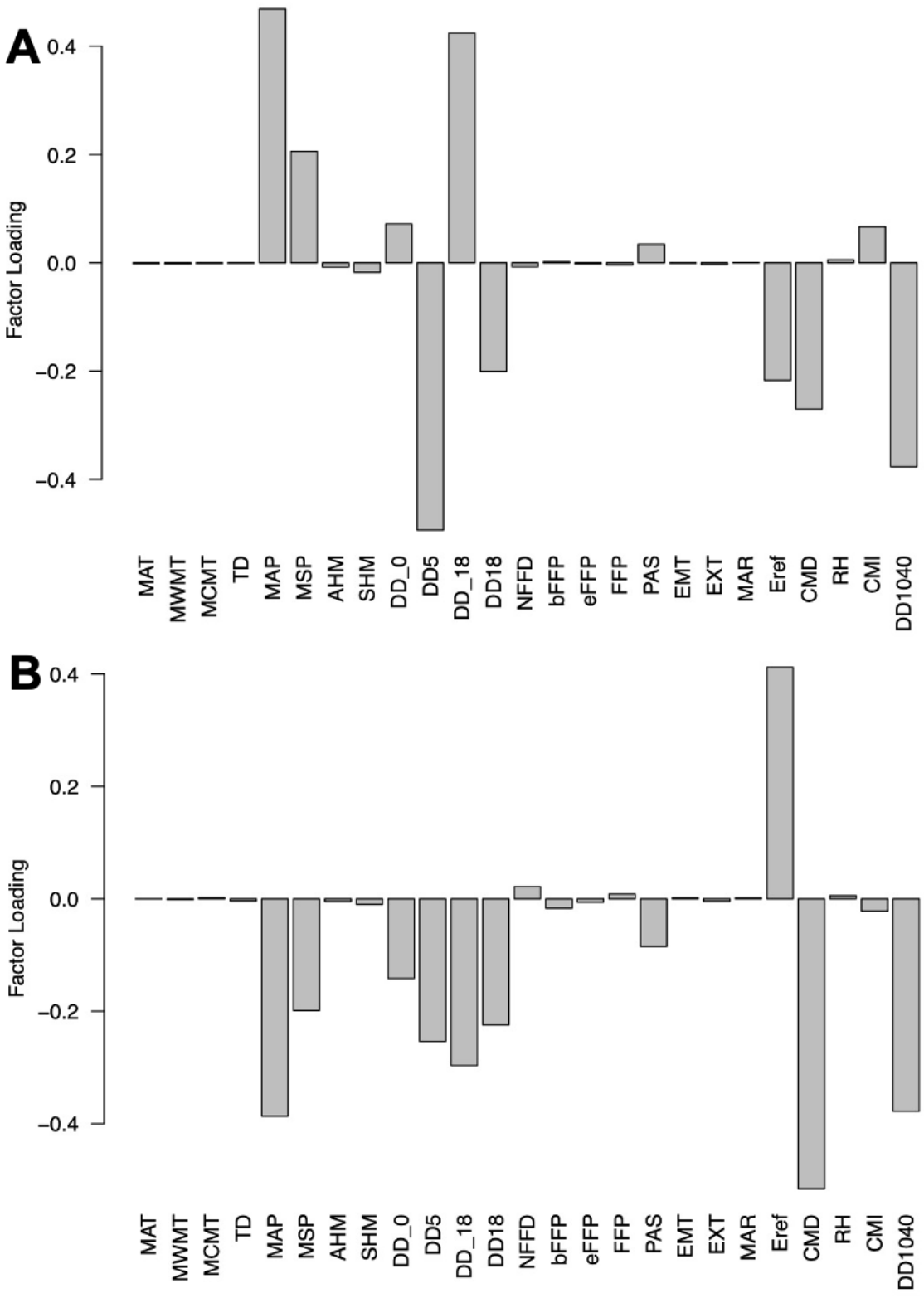

Figure S3: Loadings plots for annual climate variables from Figure 4.
A: PC1 axis; **B:** PC2 axis.

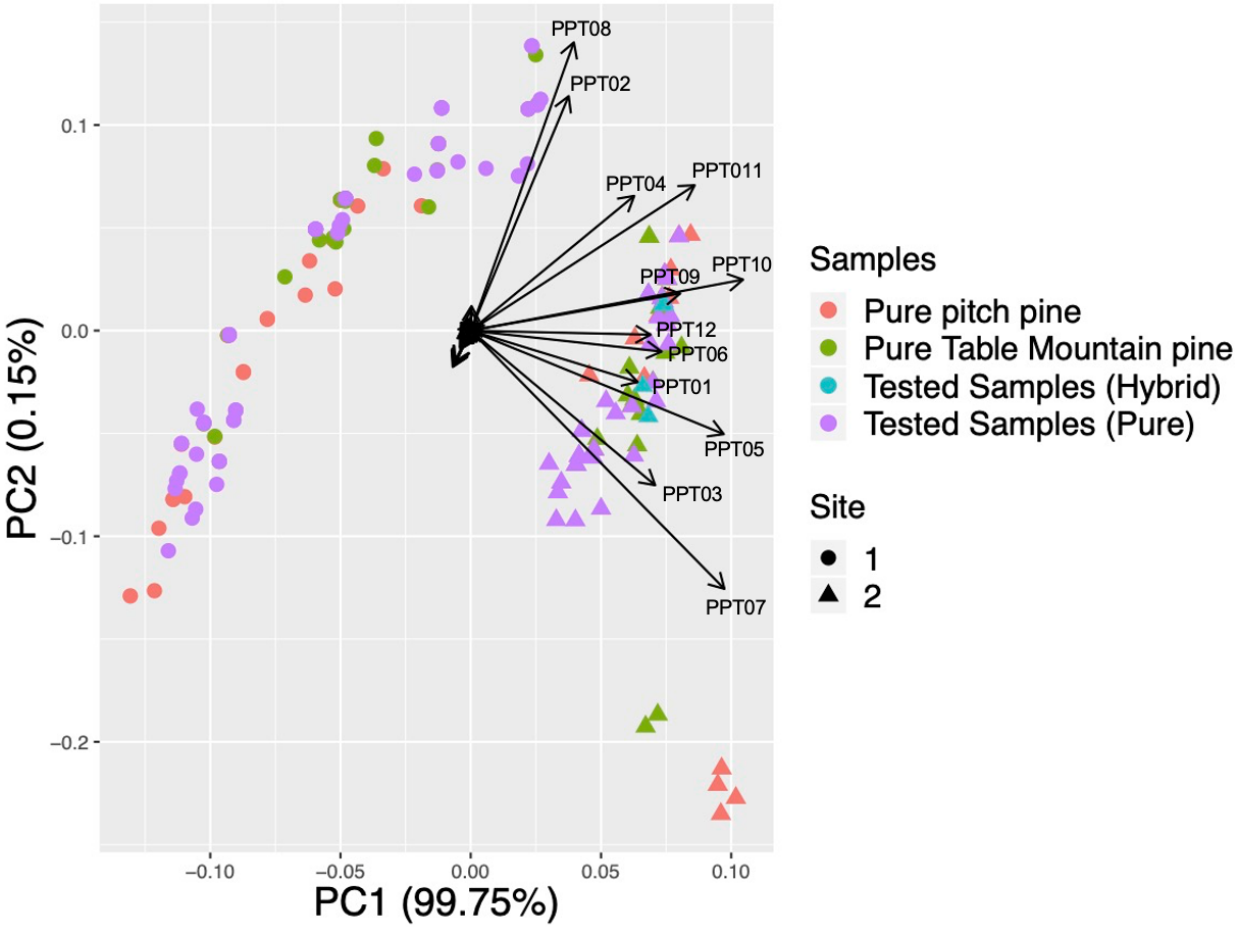

Figure S4: Principal component analysis (PCA) of monthly climatic variables (ClimateNA) for each sample coordinate location. Variable definitions are located in Table S1, along with loadings values in Figure S5.

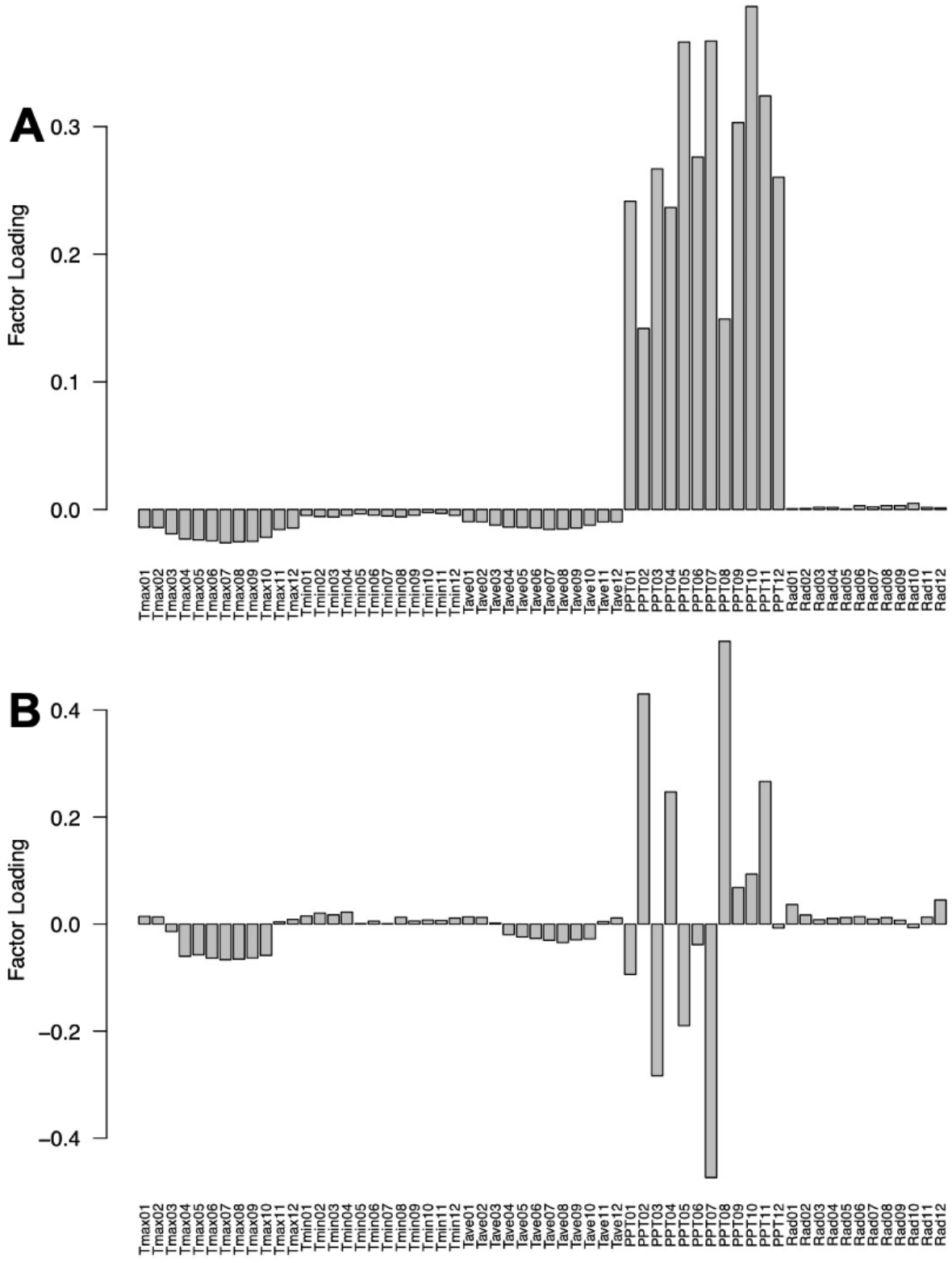

Figure S5: Loadings plots for monthly climate variables from Figure S4.
A: PC1 axis; **B:** PC2 axis.

Table S1: Climatic variable definitions.

Annual Variables		Monthly Variables	
CMD	Hargreaves climatic moisture deficit	**PPT01 – PPT12**	January – December precipitation
CMI	Climate moisture index	**RAD01 – RAD12**	January – December solar radiation
DD_0	Degree-days below 0°C	**Tave01 – Tave12**	January – December mean temperatures
DD_18	Degree-days below 18°C	**TMX01 – TMX12**	January – December maximum mean temperatures
DD1040	Degree-days above 10°C and below 40°C	**TMN01 – TMN12**	January – December minimum mean temperatures
DD18	Degree-days above 18°C		
DD5	Degree-days above 5°C		
Eref	Hargreaves reference evaporation		
MAP	Mean annual precipitation		
MSP	May to September precipitation		
PAS	Precipitation as snow		

Table S2: Linear discriminant analysis output from site 1. The Prob(Pred) values are on a scale from 0 to 1 and are respective to the species that is labeled in the Predicted column.

Sample ID	Low/Mid/High	Actual	SqDist(Actual)	Prob(Actual)	-Log(Prob)	Predicted	Prob(Pred)
PPP_001	-	Pitch	8.94667	1	0	Pitch	1
PPP_002	-	Pitch	1.0019	1	0	Pitch	1
PPP_003	-	Pitch	0.37696	1	0	Pitch	1
PPP_004	-	Pitch	0.19502	1	0	Pitch	1
PPP_005	-	Pitch	0.66996	1	0	Pitch	1
PPP_006	-	Pitch	2.52983	1	0	Pitch	1
PPP_007	-	Pitch	1.59979	1	0	Pitch	1
PPP_008	-	Pitch	1.46951	1	0	Pitch	1
PPP_009	-	Pitch	1.32854	1	0	Pitch	1
PPP_010	-	Pitch	0.36586	1	0	Pitch	1
PPP_011	-	Pitch	3.40049	0.9996	0	Pitch	0.9996
PPP_012	-	Pitch	2.41498	1	0	Pitch	1
PPP_013	-	Pitch	0.73457	1	0	Pitch	1
PPP_014	-	Pitch	0.64359	1	0	Pitch	1
PPP_015	-	Pitch	7.54299	0.9706	0.03	Pitch	0.9706
PPP_016	-	Pitch	6.71453	1	0	Pitch	1
PPP_017	-	Pitch	1.98947	1	0	Pitch	1
PPP_018	-	Pitch	3.44405	1	0	Pitch	1
PPP_019	-	Pitch	1.12792	1	0	Pitch	1
PPP_020	-	Pitch	5.58356	1	0	Pitch	1
PPP_021	-	Pitch	3.06165	1	0	Pitch	1
PPP_022	-	Pitch	0.86273	1	0	Pitch	1

PPP_023	-	Pitch	0.48882	1	0	Pitch	1
PPP_024	-	Pitch	0.99855	1	0	Pitch	1
PPP_025	-	Pitch	7.16696	1	0	Pitch	1
PTMP_001	-	TMP	4.08012	1	0	TMP	1
PTMP_002	-	TMP	16.26512	1	0	TMP	1
PTMP_003	-	TMP	2.13956	1	0	TMP	1
PTMP_004	-	TMP	1.48523	1	0	TMP	1
PTMP_005	-	TMP	8.12807	1	0	TMP	1
PTMP_006	-	TMP	3.83594	1	0	TMP	1
PTMP_007	-	TMP	0.6847	1	0	TMP	1
PTMP_008	-	TMP	1.29675	1	0	TMP	1
PTMP_009	-	TMP	1.35885	1	0	TMP	1
PTMP_010	-	TMP	0.54644	1	0	TMP	1
PTMP_011	-	TMP	0.7141	1	0	TMP	1
PTMP_012	-	TMP	5.11768	0.9997	0	TMP	0.9997
PTMP_013	-	TMP	3.46837	1	0	TMP	1
PTMP_014	-	TMP	5.20676	1	0	TMP	1
PTMP_015	-	TMP	5.94333	0.9918	0.008	TMP	0.9918
PTMP_016	-	TMP	1.8482	1	0	TMP	1
PTMP_017	-	TMP	0.99852	1	0	TMP	1
PTMP_018	-	TMP	0.81745	1	0	TMP	1
PTMP_019	-	TMP	0.99483	1	0	TMP	1
PTMP_020	-	TMP	3.69074	0.9998	0	TMP	0.9998
PTMP_021	-	TMP	1.23719	1	0	TMP	1
PTMP_022	-	TMP	3.68237	1	0	TMP	1
PTMP_023	-	TMP	1.17114	1	0	TMP	1
PTMP_024	-	TMP	1.68422	1	0	TMP	1

PTMP_025	-	TMP	2.94541	1	0	TMP	1
1-H-TMP	H	-	-	-	-	TMP	1
4-H-TMP	H	-	-	-	-	TMP	0.9997
1-H-PP	H	-	-	-	-	Pitch	1
2-H-PP	H	-	-	-	-	Pitch	1
3-H-PP	H	-	-	-	-	Pitch	1
4-H-PP	H	-	-	-	-	Pitch	1
6-H-TMP	H	-	-	-	-	TMP	1
5-H-PP	H	-	-	-	-	Pitch	1
7-H-TMP	H	-	-	-	-	TMP	1
8-H-TMP	H	-	-	-	-	TMP	0.9713
9-H-TMP	H	-	-	-	-	TMP	1
10-H-TMP	H	-	-	-	-	TMP	1
12-H-TMP	H	-	-	-	-	TMP	1
6-H-PP	H	-	-	-	-	Pitch	1
7-H-PP	H	-	-	-	-	Pitch	1
8-H-PP	H	-	-	-	-	Pitch	0.9995
9-H-PP	H	-	-	-	-	Pitch	1
10-H-PP	H	-	-	-	-	Pitch	1
15-H-TMP	H	-	-	-	-	TMP	0.9996
16-H-TMP	H	-	-	-	-	TMP	1
17-H-TMP	H	-	-	-	-	TMP	1
18-H-TMP	H	-	-	-	-	TMP	1
1-L-PP	L	-	-	-	-	Pitch	1
2-L-PP	L	-	-	-	-	Pitch	1
3-L-PP	L	-	-	-	-	Pitch	1
4-L-PP	L	-	-	-	-	Pitch	0.9986

5-L-PP	L	-	-	-	-	Pitch	1
6-L-PP	L	-	-	-	-	Pitch	1
7-L-PP	L	-	-	-	-	Pitch	1
8-L-PP	L	-	-	-	-	Pitch	1
9-L-PP	L	-	-	-	-	Pitch	1
10-L-PP	L	-	-	-	-	Pitch	1
11-L-PP	L	-	-	-	-	Pitch	1
12-L-PP	L	-	-	-	-	Pitch	1
1-L-TMP	L	-	-	-	-	TMP	1
2-L-TMP	L	-	-	-	-	TMP	1
13-L-PP	L	-	-	-	-	Pitch	0.9999
14-L-PP	L	-	-	-	-	Pitch	1
15-L-PP	L	-	-	-	-	Pitch	1
3-L-TMP	L	-	-	-	-	TMP	1
16-L-PP	L	-	-	-	-	Pitch	1
1-M-PP	M	-	-	-	-	Pitch	1
2-M-PP	M	-	-	-	-	Pitch	0.9998
3-M-PP	M	-	-	-	-	Pitch	1
4-M-PP	M	-	-	-	-	Pitch	1
5-M-PP	M	-	-	-	-	Pitch	0.9997
6-M-PP	M	-	-	-	-	Pitch	1
7-M-PP	M	-	-	-	-	Pitch	1
8-M-PP	M	-	-	-	-	Pitch	1
9-M-PP	M	-	-	-	-	Pitch	1
10-M-PP	M	-	-	-	-	Pitch	1
11-M-PP	M	-	-	-	-	Pitch	1
12-M-PP	M	-	-	-	-	Pitch	1

13-M-PP	M	-	-	-	-	Pitch	1
14-M-PP	M	-	-	-	-	Pitch	1
15-M-PP	M	-	-	-	-	Pitch	1
1-M-TMP	M	-	-	-	-	TMP	1
2-M-TMP	M	-	-	-	-	TMP	0.9992
3-M-TMP	M	-	-	-	-	TMP	1
4-M-TMP	M	-	-	-	-	TMP	1
6-M-TMP	M	-	-	-	-	TMP	1

Table S3: Linear discriminant analysis output from Site 2. The Prob(Pred) values are on a scale from 0 to 1 and are respective to the species that is labeled in the Predicted column.

Sample ID	Low/Mid/High	Training/Test Data	Actual	SqDist(Actual)	Prob(Actual)	-Log(Prob)	Predicted	Prob(Pred)
PPP_2_001	-	Trained Data (Pure Samples)	Pitch	0.296919	1	0	Pitch	1
PPP_2_002	-	Trained Data (Pure Samples)	Pitch	0.245529	0.9999	0	Pitch	0.9999
PPP_2_003	-	Trained Data (Pure Samples)	Pitch	4.656726	0.9617	0.039	Pitch	0.9617
PPP_2_004	-	Trained Data (Pure Samples)	Pitch	4.468687	1	0	Pitch	1
PPP_2_005	-	Trained Data (Pure Samples)	Pitch	9.433665	1	0	Pitch	1
PPP_2_006	-	Trained Data (Pure Samples)	Pitch	5.054485	0.9672	0.033	Pitch	0.9672
PPP_2_007	-	Trained Data (Pure Samples)	Pitch	7.270038	0.9999	0	Pitch	0.9999
PPP_2_008	-	Trained Data (Pure Samples)	Pitch	5.553804	1	0	Pitch	1
PPP_2_009	-	Trained Data (Pure Samples)	Pitch	5.218662	1	0	Pitch	1
PPP_2_010	-	Trained Data (Pure Samples)	Pitch	0.556584	1	0	Pitch	1
PPP_2_011	-	Trained Data (Pure Samples)	Pitch	1.302749	1	0	Pitch	1
PPP_2_012	-	Trained Data (Pure Samples)	Pitch	1.551719	0.9996	0	Pitch	0.9996
PPP_2_013	-	Trained Data (Pure Samples)	Pitch	0.006481	1	0	Pitch	1
PPP_2_014	-	Trained Data (Pure Samples)	Pitch	1.558321	0.9999	0	Pitch	0.9999
PPP_2_015	-	Trained Data (Pure Samples)	Pitch	5.126154	1	0	Pitch	1
PPP_2_016	-	Trained Data (Pure Samples)	Pitch	0.755371	1	0	Pitch	1
PPP_2_017	-	Trained Data (Pure Samples)	Pitch	5.833862	0.5466	0.604	Pitch	0.5466
PPP_2_018	-	Trained Data (Pure Samples)	Pitch	4.684069	1	0	Pitch	1
PPP_2_019	-	Trained Data (Pure Samples)	Pitch	0.073452	1	0	Pitch	1
PPP_2_020	-	Trained Data (Pure Samples)	Pitch	0.895429	0.9995	0.001	Pitch	0.9995
PTMP_2_001	-	Trained Data (Pure Samples)	TMP	0.978258	0.9995	0.001	TMP	0.9995
PTMP_2_002	-	Trained Data (Pure Samples)	TMP	3.644404	0.9994	0.001	TMP	0.9994

PTMP_2_003	-	Trained Data (Pure Samples)	TMP	3.737695	1	0	TMP	1
PTMP_2_004	-	Trained Data (Pure Samples)	TMP	1.365838	1	0	TMP	1
PTMP_2_005	-	Trained Data (Pure Samples)	TMP	2.592063	1	0	TMP	1
PTMP_2_006	-	Trained Data (Pure Samples)	TMP	3.310052	1	0	TMP	1
PTMP_2_007	-	Trained Data (Pure Samples)	TMP	1.132761	1	0	TMP	1
PTMP_2_008	-	Trained Data (Pure Samples)	TMP	2.643836	1	0	TMP	1
PTMP_2_009	-	Trained Data (Pure Samples)	TMP	1.048401	1	0	TMP	1
PTMP_2_010	-	Trained Data (Pure Samples)	TMP	3.426095	0.9963	0.004	TMP	0.9963
PTMP_2_011	-	Trained Data (Pure Samples)	TMP	3.171731	1	0	TMP	1
PTMP_2_012	-	Trained Data (Pure Samples)	TMP	3.172877	0.9906	0.009	TMP	0.9906
PTMP_2_013	-	Trained Data (Pure Samples)	TMP	1.791273	1	0	TMP	1
PTMP_2_014	-	Trained Data (Pure Samples)	TMP	0.64408	0.9999	0	TMP	0.9999
PTMP_2_015	-	Trained Data (Pure Samples)	TMP	1.789468	0.9996	0	TMP	0.9996
PTMP_2_016	-	Trained Data (Pure Samples)	TMP	4.280321	1	0	TMP	1
PTMP_2_017	-	Trained Data (Pure Samples)	TMP	1.723834	1	0	TMP	1
PTMP_2_018	-	Trained Data (Pure Samples)	TMP	2.4064	0.9996	0	TMP	0.9996
PTMP_2_019	-	Trained Data (Pure Samples)	TMP	2.399394	1	0	TMP	1
PTMP_2_020	-	Trained Data (Pure Samples)	TMP	4.198513	1	0	TMP	1
2-L-TMP-001	L	Tested Samples	-	-	-	-	TMP	1
2-L-TMP-002	L	Tested Samples	-	-	-	-	TMP	1
2-L-TMP-003	L	Tested Samples	-	-	-	-	TMP	0.9996
2-L-TMP-004	L	Tested Samples	-	-	-	-	TMP	1
2-L-TMP-005	L	Tested Samples	-	-	-	-	TMP	1
2-L-TMP-006	L	Tested Samples	-	-	-	-	TMP	1
2-L-TMP-007	L	Tested Samples	-	-	-	-	TMP	1
2-L-TMP-008	L	Tested Samples	-	-	-	-	TMP	0.9998
2-L-TMP-009	L	Tested Samples	-	-	-	-	TMP	1

2-L-TMP-010	L	Tested Samples	-	-	-	-	TMP	0.9991
2-L-TMP-011	L	Tested Samples	-	-	-	-	TMP	0.9994
2-L-TMP-012	L	Tested Samples	-	-	-	-	TMP	0.9996
2-L-TMP-013	L	Tested Samples	-	-	-	-	TMP	1
2-L-TMP-014	L	Tested Samples	-	-	-	-	TMP	1
2-L-PP-001	L	Tested Samples	-	-	-	-	Pitch	1
2-L-TMP-015	L	Tested Samples	-	-	-	-	TMP	0.8972
2-L-TMP-016	L	Tested Samples	-	-	-	-	TMP	1
2-L-TMP-017	L	Tested Samples	-	-	-	-	TMP	1
2-L-TMP-018	L	Tested Samples	-	-	-	-	TMP	1
2-L-TMP-019	L	Tested Samples	-	-	-	-	TMP	0.9796
2-L-TMP-020	L	Tested Samples	-	-	-	-	TMP	1
2-H-PP-001	H	Tested Samples	-	-	-	-	Pitch	1
2-H-PP-002	H	Tested Samples	-	-	-	-	Pitch	0.9992
2-H-TMP-001	H	Tested Samples	-	-	-	-	TMP	0.9997
2-H-TMP-002	H	Tested Samples	-	-	-	-	TMP	1
2-H-TMP-003	H	Tested Samples	-	-	-	-	TMP	0.9999
2-H-PP-003	H	Tested Samples	-	-	-	-	Pitch	1
2-H-TMP-004	H	Tested Samples	-	-	-	-	TMP	0.9973
2-H-TMP-005	H	Tested Samples	-	-	-	-	TMP	1
2-H-PP-004	H	Tested Samples	-	-	-	-	Pitch	1
2-H-PP-005	H	Tested Samples	-	-	-	-	Pitch	0.9998
2-H-PP-006	H	Tested Samples	-	-	-	-	TMP	0.9863
2-H-PP-007	H	Tested Samples	-	-	-	-	Pitch	0.9999
2-H-PP-008	H	Tested Samples	-	-	-	-	Pitch	1
2-H-PP-009	H	Tested Samples	-	-	-	-	Pitch	1
2-H-PP-010	H	Tested Samples	-	-	-	-	Pitch	0.9992

2-H-PP-011	H	Tested Samples	-	-	-	-	Pitch	1
2-H-PP-012	H	Tested Samples	-	-	-	-	Pitch	0.9989
2-H-PP-013	H	Tested Samples	-	-	-	-	Pitch	1
2-M-PP-001	M	Tested Samples	-	-	-	-	Pitch	0.7581
2-M-PP-002	M	Tested Samples	-	-	-	-	Pitch	1
2-M-TMP-001	M	Tested Samples	-	-	-	-	TMP	1
2-M-TMP-002	M	Tested Samples	-	-	-	-	TMP	0.8366
2-M-TMP-003	M	Tested Samples	-	-	-	-	TMP	0.9999
2-M-TMP-004	M	Tested Samples	-	-	-	-	Pitch	0.8812
2-M-TMP-005	M	Tested Samples	-	-	-	-	TMP	1
2-M-TMP-006	M	Tested Samples	-	-	-	-	TMP	1
2-M-TMP-007	M	Tested Samples	-	-	-	-	TMP	1
2-M-TMP-008	M	Tested Samples	-	-	-	-	TMP	0.9997
2-M-TMP-009	M	Tested Samples	-	-	-	-	TMP	1
2-M-TMP-010	M	Tested Samples	-	-	-	-	TMP	0.9453
2-M-PP-002	M	Tested Samples	-	-	-	-	Pitch	0.993
2-M-PP-003	M	Tested Samples	-	-	-	-	Pitch	0.9148
2-M-PP-004	M	Tested Samples	-	-	-	-	Pitch	1
2-M-TMP-011	M	Tested Samples	-	-	-	-	TMP	0.9909
2-M-TMP-013	M	Tested Samples	-	-	-	-	Pitch	0.6136

Table S4: Linear discriminant analysis output using Site 1 pure samples as training data and Site 2 pure samples as testing data.

Sample ID	Training/Test Data	Actual	SqDist(Actual)	Prob(Actual)	-Log(Prob)	Predicted	Prob(Pred)
PPP_001	Training Data (Pure Site 1 Samples)	Pitch	8.94667	1	0	Pitch	1
PPP_002	Training Data (Pure Site 1 Samples)	Pitch	1.0019	1	0	Pitch	1
PPP_003	Training Data (Pure Site 1 Samples)	Pitch	0.37696	1	0	Pitch	1
PPP_004	Training Data (Pure Site 1 Samples)	Pitch	0.19502	1	0	Pitch	1
PPP_005	Training Data (Pure Site 1 Samples)	Pitch	0.66996	1	0	Pitch	1
PPP_006	Training Data (Pure Site 1 Samples)	Pitch	2.52983	1	0	Pitch	1
PPP_007	Training Data (Pure Site 1 Samples)	Pitch	1.59979	1	0	Pitch	1
PPP_008	Training Data (Pure Site 1 Samples)	Pitch	1.46951	1	0	Pitch	1
PPP_009	Training Data (Pure Site 1 Samples)	Pitch	1.32854	1	0	Pitch	1
PPP_010	Training Data (Pure Site 1 Samples)	Pitch	0.36586	1	0	Pitch	1
PPP_011	Training Data (Pure Site 1 Samples)	Pitch	3.40049	0.9996	0	Pitch	0.9996
PPP_012	Training Data (Pure Site 1 Samples)	Pitch	2.41498	1	0	Pitch	1
PPP_013	Training Data (Pure Site 1 Samples)	Pitch	0.73457	1	0	Pitch	1
PPP_014	Training Data (Pure Site 1 Samples)	Pitch	0.64359	1	0	Pitch	1
PPP_015	Training Data (Pure Site 1 Samples)	Pitch	7.54299	0.9706	0.03	Pitch	0.9706
PPP_016	Training Data (Pure Site 1 Samples)	Pitch	6.71453	1	0	Pitch	1
PPP_017	Training Data (Pure Site 1 Samples)	Pitch	1.98947	1	0	Pitch	1
PPP_018	Training Data (Pure Site 1 Samples)	Pitch	3.44405	1	0	Pitch	1
PPP_019	Training Data (Pure Site 1 Samples)	Pitch	1.12792	1	0	Pitch	1
PPP_020	Training Data (Pure Site 1 Samples)	Pitch	5.58356	1	0	Pitch	1
PPP_021	Training Data (Pure Site 1 Samples)	Pitch	3.06165	1	0	Pitch	1
PPP_022	Training Data (Pure Site 1 Samples)	Pitch	0.86273	1	0	Pitch	1

PPP_023	Training Data (Pure Site 1 Samples)	Pitch	0.48882	1	0	Pitch	1
PPP_024	Training Data (Pure Site 1 Samples)	Pitch	0.99855	1	0	Pitch	1
PPP_025	Training Data (Pure Site 1 Samples)	Pitch	7.16696	1	0	Pitch	1
PTMP_001	Training Data (Pure Site 1 Samples)	TMP	4.08012	1	0	TMP	1
PTMP_002	Training Data (Pure Site 1 Samples)	TMP	16.26512	1	0	TMP	1
PTMP_003	Training Data (Pure Site 1 Samples)	TMP	2.13956	1	0	TMP	1
PTMP_004	Training Data (Pure Site 1 Samples)	TMP	1.48523	1	0	TMP	1
PTMP_005	Training Data (Pure Site 1 Samples)	TMP	8.12807	1	0	TMP	1
PTMP_006	Training Data (Pure Site 1 Samples)	TMP	3.83594	1	0	TMP	1
PTMP_007	Training Data (Pure Site 1 Samples)	TMP	0.6847	1	0	TMP	1
PTMP_008	Training Data (Pure Site 1 Samples)	TMP	1.29675	1	0	TMP	1
PTMP_009	Training Data (Pure Site 1 Samples)	TMP	1.35885	1	0	TMP	1
PTMP_010	Training Data (Pure Site 1 Samples)	TMP	0.54644	1	0	TMP	1
PTMP_011	Training Data (Pure Site 1 Samples)	TMP	0.7141	1	0	TMP	1
PTMP_012	Training Data (Pure Site 1 Samples)	TMP	5.11768	0.9997	0	TMP	0.9997
PTMP_013	Training Data (Pure Site 1 Samples)	TMP	3.46837	1	0	TMP	1
PTMP_014	Training Data (Pure Site 1 Samples)	TMP	5.20676	1	0	TMP	1
PTMP_015	Training Data (Pure Site 1 Samples)	TMP	5.94333	0.9918	0.008	TMP	0.9918
PTMP_016	Training Data (Pure Site 1 Samples)	TMP	1.8482	1	0	TMP	1
PTMP_017	Training Data (Pure Site 1 Samples)	TMP	0.99852	1	0	TMP	1
PTMP_018	Training Data (Pure Site 1 Samples)	TMP	0.81745	1	0	TMP	1
PTMP_019	Training Data (Pure Site 1 Samples)	TMP	0.99483	1	0	TMP	1
PTMP_020	Training Data (Pure Site 1 Samples)	TMP	3.69074	0.9998	0	TMP	0.9998
PTMP_021	Training Data (Pure Site 1 Samples)	TMP	1.23719	1	0	TMP	1
PTMP_022	Training Data (Pure Site 1 Samples)	TMP	3.68237	1	0	TMP	1
PTMP_023	Training Data (Pure Site 1 Samples)	TMP	1.17114	1	0	TMP	1
PTMP_024	Training Data (Pure Site 1 Samples)	TMP	1.68422	1	0	TMP	1

PTMP_025	Training Data (Pure Site 1 Samples)	TMP	2.94541	1	0	TMP	1
PPP_2_001	Test Data (Pure Site 2 Samples)	Pitch	-	-	-	Pitch	1
PPP_2_002	Test Data (Pure Site 2 Samples)	Pitch	-	-	-	Pitch	1
PPP_2_003	Test Data (Pure Site 2 Samples)	Pitch	-	-	-	Pitch	1
PPP_2_004	Test Data (Pure Site 2 Samples)	Pitch	-	-	-	Pitch	1
PPP_2_005	Test Data (Pure Site 2 Samples)	Pitch	-	-	-	Pitch	1
PPP_2_006	Test Data (Pure Site 2 Samples)	Pitch	-	-	-	Pitch	1
PPP_2_007	Test Data (Pure Site 2 Samples)	Pitch	-	-	-	Pitch	1
PPP_2_008	Test Data (Pure Site 2 Samples)	Pitch	-	-	-	Pitch	1
PPP_2_009	Test Data (Pure Site 2 Samples)	Pitch	-	-	-	Pitch	1
PPP_2_010	Test Data (Pure Site 2 Samples)	Pitch	-	-	-	Pitch	1
PPP_2_011	Test Data (Pure Site 2 Samples)	Pitch	-	-	-	Pitch	1
PPP_2_012	Test Data (Pure Site 2 Samples)	Pitch	-	-	-	Pitch	1
PPP_2_013	Test Data (Pure Site 2 Samples)	Pitch	-	-	-	Pitch	1
PPP_2_014	Test Data (Pure Site 2 Samples)	Pitch	-	-	-	Pitch	0.9998
PPP_2_015	Test Data (Pure Site 2 Samples)	Pitch	-	-	-	Pitch	1
PPP_2_016	Test Data (Pure Site 2 Samples)	Pitch	-	-	-	Pitch	1
PPP_2_017	Test Data (Pure Site 2 Samples)	Pitch	-	-	-	Pitch	0.7201
PPP_2_018	Test Data (Pure Site 2 Samples)	Pitch	-	-	-	Pitch	1
PPP_2_019	Test Data (Pure Site 2 Samples)	Pitch	-	-	-	Pitch	1
PPP_2_020	Test Data (Pure Site 2 Samples)	Pitch	-	-	-	Pitch	1
PTMP_2_001	Test Data (Pure Site 2 Samples)	TMP	-	-	-	TMP	0.8763
PTMP_2_002	Test Data (Pure Site 2 Samples)	TMP	-	-	-	TMP	1
PTMP_2_003	Test Data (Pure Site 2 Samples)	TMP	-	-	-	Pitch	0.8121
PTMP_2_004	Test Data (Pure Site 2 Samples)	TMP	-	-	-	TMP	1
PTMP_2_005	Test Data (Pure Site 2 Samples)	TMP	-	-	-	TMP	1
PTMP_2_006	Test Data (Pure Site 2 Samples)	TMP	-	-	-	TMP	0.9992

PTMP_2_007	Test Data (Pure Site 2 Samples)	TMP	-	-	-	TMP	0.9995
PTMP_2_008	Test Data (Pure Site 2 Samples)	TMP	-	-	-	TMP	0.9858
PTMP_2_009	Test Data (Pure Site 2 Samples)	TMP	-	-	-	TMP	0.9989
PTMP_2_010	Test Data (Pure Site 2 Samples)	TMP	-	-	-	TMP	0.7192
PTMP_2_011	Test Data (Pure Site 2 Samples)	TMP	-	-	-	TMP	1
PTMP_2_012	Test Data (Pure Site 2 Samples)	TMP	-	-	-	Pitch	0.9971
PTMP_2_013	Test Data (Pure Site 2 Samples)	TMP	-	-	-	TMP	0.6055
PTMP_2_014	Test Data (Pure Site 2 Samples)	TMP	-	-	-	TMP	0.9916
PTMP_2_015	Test Data (Pure Site 2 Samples)	TMP	-	-	-	Pitch	0.9558
PTMP_2_016	Test Data (Pure Site 2 Samples)	TMP	-	-	-	TMP	1
PTMP_2_017	Test Data (Pure Site 2 Samples)	TMP	-	-	-	TMP	1
PTMP_2_018	Test Data (Pure Site 2 Samples)	TMP	-	-	-	TMP	1
PTMP_2_019	Test Data (Pure Site 2 Samples)	TMP	-	-	-	TMP	1
PTMP_2_020	Test Data (Pure Site 2 Samples)	TMP	-	-	-	TMP	1

Table S5: Linear discriminant analysis output using Site 2 pure samples as training data and Site 1 pure samples as testing data.

Sample ID	Training/Test Data	Actual	SqDist(Actual)	Prob(Actual)	-Log(Prob)	Predicted	Prob(Pred)
PPP_2_001	Training Data (Pure Site 2 Samples)	Pitch	0.296919	1	0	Pitch	1
PPP_2_002	Training Data (Pure Site 2 Samples)	Pitch	0.245529	0.9999	0	Pitch	0.9999
PPP_2_003	Training Data (Pure Site 2 Samples)	Pitch	4.656726	0.9617	0.039	Pitch	0.9617
PPP_2_004	Training Data (Pure Site 2 Samples)	Pitch	4.468687	1	0	Pitch	1
PPP_2_005	Training Data (Pure Site 2 Samples)	Pitch	9.433665	1	0	Pitch	1
PPP_2_006	Training Data (Pure Site 2 Samples)	Pitch	5.054485	0.9672	0.033	Pitch	0.9672
PPP_2_007	Training Data (Pure Site 2 Samples)	Pitch	7.270038	0.9999	0	Pitch	0.9999
PPP_2_008	Training Data (Pure Site 2 Samples)	Pitch	5.553804	1	0	Pitch	1
PPP_2_009	Training Data (Pure Site 2 Samples)	Pitch	5.218662	1	0	Pitch	1
PPP_2_010	Training Data (Pure Site 2 Samples)	Pitch	0.556584	1	0	Pitch	1
PPP_2_011	Training Data (Pure Site 2 Samples)	Pitch	1.302749	1	0	Pitch	1
PPP_2_012	Training Data (Pure Site 2 Samples)	Pitch	1.551719	0.9996	0	Pitch	0.9996
PPP_2_013	Training Data (Pure Site 2 Samples)	Pitch	0.006481	1	0	Pitch	1
PPP_2_014	Training Data (Pure Site 2 Samples)	Pitch	1.558321	0.9999	0	Pitch	0.9999
PPP_2_015	Training Data (Pure Site 2 Samples)	Pitch	5.126154	1	0	Pitch	1
PPP_2_016	Training Data (Pure Site 2 Samples)	Pitch	0.755371	1	0	Pitch	1
PPP_2_017	Training Data (Pure Site 2 Samples)	Pitch	5.833862	0.5466	0.604	Pitch	0.5466
PPP_2_018	Training Data (Pure Site 2 Samples)	Pitch	4.684069	1	0	Pitch	1
PPP_2_019	Training Data (Pure Site 2 Samples)	Pitch	0.073452	1	0	Pitch	1
PPP_2_020	Training Data (Pure Site 2 Samples)	Pitch	0.895429	0.9995	0.001	Pitch	0.9995
PTMP_2_001	Training Data (Pure Site 2 Samples)	TMP	0.978258	0.9995	0.001	TMP	0.9995
PTMP_2_002	Training Data (Pure Site 2 Samples)	TMP	3.644404	0.9994	0.001	TMP	0.9994

PTMP_2_003	Training Data (Pure Site 2 Samples)	TMP	3.737695	1	0	TMP	1
PTMP_2_004	Training Data (Pure Site 2 Samples)	TMP	1.365838	1	0	TMP	1
PTMP_2_005	Training Data (Pure Site 2 Samples)	TMP	2.592063	1	0	TMP	1
PTMP_2_006	Training Data (Pure Site 2 Samples)	TMP	3.310052	1	0	TMP	1
PTMP_2_007	Training Data (Pure Site 2 Samples)	TMP	1.132761	1	0	TMP	1
PTMP_2_008	Training Data (Pure Site 2 Samples)	TMP	2.643836	1	0	TMP	1
PTMP_2_009	Training Data (Pure Site 2 Samples)	TMP	1.048401	1	0	TMP	1
PTMP_2_010	Training Data (Pure Site 2 Samples)	TMP	3.426095	0.9963	0.004	TMP	0.9963
PTMP_2_011	Training Data (Pure Site 2 Samples)	TMP	3.171731	1	0	TMP	1
PTMP_2_012	Training Data (Pure Site 2 Samples)	TMP	3.172877	0.9906	0.009	TMP	0.9906
PTMP_2_013	Training Data (Pure Site 2 Samples)	TMP	1.791273	1	0	TMP	1
PTMP_2_014	Training Data (Pure Site 2 Samples)	TMP	0.64408	0.9999	0	TMP	0.9999
PTMP_2_015	Training Data (Pure Site 2 Samples)	TMP	1.789468	0.9996	0	TMP	0.9996
PTMP_2_016	Training Data (Pure Site 2 Samples)	TMP	4.280321	1	0	TMP	1
PTMP_2_017	Training Data (Pure Site 2 Samples)	TMP	1.723834	1	0	TMP	1
PTMP_2_018	Training Data (Pure Site 2 Samples)	TMP	2.4064	0.9996	0	TMP	0.9996
PTMP_2_019	Training Data (Pure Site 2 Samples)	TMP	2.399394	1	0	TMP	1
PTMP_2_020	Training Data (Pure Site 2 Samples)	TMP	4.198513	1	0	TMP	1
PPP_001	Test Data (Pure Site 1 Samples)	Pitch	-	-	-	Pitch	0.9998
PPP_002	Test Data (Pure Site 1 Samples)	Pitch	-	-	-	Pitch	1
PPP_003	Test Data (Pure Site 1 Samples)	Pitch	-	-	-	Pitch	1
PPP_004	Test Data (Pure Site 1 Samples)	Pitch	-	-	-	Pitch	0.9999
PPP_005	Test Data (Pure Site 1 Samples)	Pitch	-	-	-	Pitch	0.9999
PPP_006	Test Data (Pure Site 1 Samples)	Pitch	-	-	-	Pitch	1
PPP_007	Test Data (Pure Site 1 Samples)	Pitch	-	-	-	Pitch	0.8571
PPP_008	Test Data (Pure Site 1 Samples)	Pitch	-	-	-	Pitch	0.9986
PPP_009	Test Data (Pure Site 1 Samples)	Pitch	-	-	-	Pitch	1

PPP_010	Test Data (Pure Site 1 Samples)	Pitch	-	-	-	Pitch	0.9999
PPP_011	Test Data (Pure Site 1 Samples)	Pitch	-	-	-	Pitch	0.7788
PPP_012	Test Data (Pure Site 1 Samples)	Pitch	-	-	-	Pitch	0.999
PPP_013	Test Data (Pure Site 1 Samples)	Pitch	-	-	-	Pitch	0.9999
PPP_014	Test Data (Pure Site 1 Samples)	Pitch	-	-	-	Pitch	1
PPP_015	Test Data (Pure Site 1 Samples)	Pitch	-	-	-	Pitch	0.5856
PPP_016	Test Data (Pure Site 1 Samples)	Pitch	-	-	-	Pitch	0.99
PPP_017	Test Data (Pure Site 1 Samples)	Pitch	-	-	-	Pitch	1
PPP_018	Test Data (Pure Site 1 Samples)	Pitch	-	-	-	Pitch	1
PPP_019	Test Data (Pure Site 1 Samples)	Pitch	-	-	-	Pitch	1
PPP_020	Test Data (Pure Site 1 Samples)	Pitch	-	-	-	Pitch	1
PPP_021	Test Data (Pure Site 1 Samples)	Pitch	-	-	-	Pitch	0.995
PPP_022	Test Data (Pure Site 1 Samples)	Pitch	-	-	-	Pitch	0.9966
PPP_023	Test Data (Pure Site 1 Samples)	Pitch	-	-	-	Pitch	0.9996
PPP_024	Test Data (Pure Site 1 Samples)	Pitch	-	-	-	Pitch	1
PPP_025	Test Data (Pure Site 1 Samples)	Pitch	-	-	-	Pitch	1
PTMP_001	Test Data (Pure Site 1 Samples)	TMP	-	-	-	TMP	0.9999
PTMP_002	Test Data (Pure Site 1 Samples)	TMP	-	-	-	TMP	1
PTMP_003	Test Data (Pure Site 1 Samples)	TMP	-	-	-	TMP	1
PTMP_004	Test Data (Pure Site 1 Samples)	TMP	-	-	-	TMP	1
PTMP_005	Test Data (Pure Site 1 Samples)	TMP	-	-	-	TMP	1
PTMP_006	Test Data (Pure Site 1 Samples)	TMP	-	-	-	TMP	1
PTMP_007	Test Data (Pure Site 1 Samples)	TMP	-	-	-	TMP	1
PTMP_008	Test Data (Pure Site 1 Samples)	TMP	-	-	-	TMP	1
PTMP_009	Test Data (Pure Site 1 Samples)	TMP	-	-	-	TMP	1
PTMP_010	Test Data (Pure Site 1 Samples)	TMP	-	-	-	TMP	1
PTMP_011	Test Data (Pure Site 1 Samples)	TMP	-	-	-	TMP	1

PTMP_012	Test Data (Pure Site 1 Samples)	TMP	-	-	-	TMP	0.9999
PTMP_013	Test Data (Pure Site 1 Samples)	TMP	-	-	-	TMP	1
PTMP_014	Test Data (Pure Site 1 Samples)	TMP	-	-	-	TMP	1
PTMP_015	Test Data (Pure Site 1 Samples)	TMP	-	-	-	TMP	0.9906
PTMP_016	Test Data (Pure Site 1 Samples)	TMP	-	-	-	TMP	1
PTMP_017	Test Data (Pure Site 1 Samples)	TMP	-	-	-	TMP	0.9999
PTMP_018	Test Data (Pure Site 1 Samples)	TMP	-	-	-	TMP	1
PTMP_019	Test Data (Pure Site 1 Samples)	TMP	-	-	-	TMP	1
PTMP_020	Test Data (Pure Site 1 Samples)	TMP	-	-	-	TMP	0.9883
PTMP_021	Test Data (Pure Site 1 Samples)	TMP	-	-	-	TMP	1
PTMP_022	Test Data (Pure Site 1 Samples)	TMP	-	-	-	TMP	1
PTMP_023	Test Data (Pure Site 1 Samples)	TMP	-	-	-	TMP	1
PTMP_024	Test Data (Pure Site 1 Samples)	TMP	-	-	-	TMP	0.9992
PTMP_025	Test Data (Pure Site 1 Samples)	TMP	-	-	-	TMP	0.9956

Vita

Alexander Louis Brown was born on the 29th of July 1994 in Horseheads, New York, where he would go on to graduate from Horseheads High School in 2012. He received his Associate of Science degree in Environmental Science from Corning Community College (Corning, New York) in 2013. He then attended the State University of New York College of Environmental Science and Forestry (Syracuse, New York) where he received his Bachelor of Science degree in Environmental Biology in 2016. Following his academic career, he was employed as an Environmental Health Specialist for Wake County (Raleigh, North Carolina) up until his return to graduate school in the Master of Science in Biology program at Virginia Commonwealth University (Richmond, Virginia) in 2019.