Intraspecific Niche Models for Ponderosa Pine (Pinus ponderosa) Suggest Potential Variability in Population-Level Response to Climate Change

KAITLIN C. MAGUIRE1,*, DOUGLAS J. SHINNEMAN1, KEVIN M. POTTER2, and VALERIE D. HIPKINS3
1U.S. Geological Survey, Forest and Rangeland Ecosystem Science Center, 970 S. Lusk Street Boise, ID 83706, USA; 2Department of Forestry and Environmental Resources, North Carolina State University, Research Triangle Park, NC 27695, USA; and 3National Forest Genetics Laboratory, USDA Forest Service, 2480 Carson Road, Placerville, CA 95667, USA
*Correspondence to be sent to: U. S. Geological Survey, Forest and Rangeland Ecosystem Science Center, 970 S. Lusk Street Boise, ID 83706, USA; E-mail: kmaguire@usgs.gov.

Abstract — Unique responses to climate change can occur across intraspecific levels, resulting in individualistic adaptation or movement patterns among populations within a given species. Thus, the need to model potential responses among genetically distinct populations within a species is increasingly recognized. However, predictive models of future distributions are regularly fit at the species level, often because intraspecific variation is unknown or is identified only within limited sample locations. In this study, we considered the role of intraspecific variation to shape the geographic distribution of ponderosa pine (Pinus ponderosa), an ecologically and economically important tree species in North America. Morphological and genetic variation across the distribution of ponderosa pine suggest the need to model intraspecific populations: the two varieties (var. ponderosa and var. scopulorum) and several haplotype groups within each variety have been shown to occupy unique climatic niches, suggesting populations have distinct evolutionary lineages adapted to different environmental conditions. We utilized a recently available, geographically widespread dataset of intraspecific variation (haplotypes) for ponderosa pine and a recently devised lineage distance modeling approach to derive additional, likely unique species occurrences. We confirmed the relative uniqueness of each haplotype-climate relationship using a niche-overlap analysis, and developed ecological niche models (ENMs) to project the distribution for two varieties and eight haplotypes under future climate forecasts. Future projections of haplotype niche distributions generally revealed greater potential range loss than predicted for the varieties. This difference may reflect intraspecific responses of distinct evolutionary lineages. However, directional trends are generally consistent across intraspecific levels, and include a loss of climatic distributional area and an upward shift in elevation. Our results demonstrate the utility in modeling intraspecific response to changing climate and they inform management and conservation strategies, by identifying haplotypes and geographic areas that may be most at risk, or most secure, under projected climate change.

© 2018 John Wiley & Sons, Ltd. Syst. Biol. 0(0):1–14, 2018 DOI:10.1093/sysbio/syy017

1–15

A species’ response to environmental change is not always uniform across its distribution as genotypic and phenotypic differences within a species can interact with local environmental conditions and result in local adaptation (Valladares et al. 2014). Even though most biogeographical modeling studies treat a species as a single unit, its response to climate change will typically vary across space, either through plastic gene expression, adaptation, extirpation, or migration (Nerini et al. 2015; Valladares et al. 2014). Therefore, population and phylogeographic structure should not be ignored when modeling species distributions, inferring niche structure, or predicting how climate will affect species distributions (Pfenninger et al. 2007; Pearman et al. 2010). Not including phylogeographic structure may lead to underrepresentation of certain lineages in species models (Pearman et al. 2010), especially when they are clearly defined and occupy distinct niche spaces (Hallfors et al. 2016). Distinct lineages defined by genetic variation are the units of evolutionary change and thus crucial for providing insight into the ability of a species to adapt to environmental change. Predicting intraspecific geographic response to environmental factors is becoming increasingly feasible as research reveals genetic variation across the distribution of individual species (e.g., Pearman et al. 2010; Espindola et al. 2012; D’Amen et al. 2013; Yannic et al. 2014; Hallfors et al. 2016; Prasad and Potter 2017).

Modeling the ecological niches of populations or lineages may be especially important for species with spatial variation in genotypic or phenotypic traits that suggest they are in the process of differentiating into two or more species (Valladares et al. 2014). Ponderosa pine (Pinus ponderosa) is a widespread, well-studied, and economically important species in Western North America, and its substantial morphological variation (Wells 1964) suggests it may be in early stages of differentiation into multiple species (Wang 1977, Jaramillo-Correa et al. 2009). Two varieties within the species are recognized—the Pacific variety, P. ponderosa var. ponderosa Laws., has open, plume-like foliage, and a low proportion of two-needle fascicles as opposed to the Rocky Mountain variety, P. ponderosa var. scopulorum Engelm., which has compact and brush-like foliage and moderate to high proportions of two-needle fascicles. The separation of the two varieties is also established by leaf oil terpene composition (Rudloff and Lapp 1992), mitochondrial DNA (Johansen and Latta 2003; Potter et al. 2013), highly polymorphic nuclear microsatellite and isozyme markers (Potter et al. 2015), and allozymes (Niebling and Conkle 1990). The two varieties have similar temperature limitations but...
occur within different precipitation regimes, with var. ponderosa locations generally corresponding to higher winter moisture and var. scopulorum to higher summer moisture availability (Norris et al. 2006; Shinneman et al. 2016). It is thought that the varieties diverged at least before the last glacial maximum, and more likely around 250,000 years ago (Lascoux et al. 2004), with Pleistocene glacial and interglacial cycles leading to further genetic diversification through isolation in localized refugia, followed by subsequent migration, hybridization, and introgression (Shinneman et al. 2016).

Within each variety, several geographically structured haplotypes have also been identified using mitochondrial DNA (Johansen and Latta 2003; Potter et al. 2013) and plastome sequences (Wofford et al. 2014), and genetic clusters have been identified from nuclear DNA microsatellite markers (Potter et al. 2015). Geographic structure identified from mtDNA suggests each haplotype resulted from unique patterns of migration, isolation, and local adaptation (Potter et al. 2013). Mitochondrial DNA is maternally inherited and dispersed via seeds with limited dispersal compared with pollen-dispersed chloroplast DNA, and geographic structure in genetic differentiation is retained longer (Neale and Sederoff 1989; Petit et al. 1993). Thus, the haplotypes may represent evolutionarily distinct units capable of responding differently to climate change as a result of unique adaptations to refugial climate conditions and long-term isolation during glacial periods (Potter et al. 2013, Shinneman et al. 2016).

Given unique intraspecific lineages and the potential for distinctive climate niches, the need to model distributions of ponderosa pine at the subspecific level has been recognized (Norris et al. 2006; Rehfeldt et al. 2014, Shinneman et al. 2016). Shinneman et al. (2016) modeled the contemporary and historical climate niche distributions of the two varieties, as well as 10 geographically structured mtDNA haplotypes identified by Potter et al. (2013) to further infer the evolutionary and phylogeographic history of these lineages. Recent modeling research has also suggested that ponderosa pine may be threatened by changing climate, but that the response may differ for the two varieties (Rehfeldt et al. 2014). However, the potential for future climate change to shape the distribution of ponderosa pine populations, as defined by haplotypic diversity, has not been explored, and has only been studied for a handful of other species relative to intraspecific variation (Pearman et al. 2010; Bálint et al. 2011; Bentio Garzón et al. 2011; Oney et al. 2014). Geographic structure identified from mtDNA nad1 second intron minisatellite region of foliage samples collected from 3113 trees representing 104 populations of ponderosa pine across its range (Potter et al. 2013). At least 20 individual trees were sampled in each population (except two), with most populations encompassing at least 30 sampled trees. Sampled trees were at least 100 m apart to encompass the entire range of the population's genetic composition. Haplotype 1, 5, and 8 are within the distribution of var. ponderosa and group together in phylogenetic sequences of the same haplotype that fell within the same geographic grid cell (800 m × 800 m) were removed from the dataset.

**Methods**

**Haplotype Occurrence Data**

We examined eight of the ten haplotypes identified in Potter et al. (2013). We did not examine haplotypes 9 and 10, each limited to a small area of California, because of restricted and localized populations less amenable to niche modeling. Haplotype were identified from the mtDNA nad1 second intron minisatellite region of foliage sampled in each population (except two), with most populations encompassing at least 30 sampled trees. Haplotype 1, 5, and 8 are within the distribution of var. ponderosa and group together in phylogenetic sequences of the same haplotype that fell within the same geographic grid cell (800 m × 800 m) were removed from the dataset.

**Climate Data**

Climate variables derived using thin-plate splines to interpolate 1961–1990 monthly temperature and precipitation data at 800 m resolution were obtained from the Research on Forest Climate Change website (charcoal.cnre.vt.edu/climate; Rehfeldt 2006;
TABLE 1. Climate variables considered for niche modeling of ponderosa pine varieties and haplotypes

<table>
<thead>
<tr>
<th>Climate variables</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAP</td>
<td>Mean annual precipitation.</td>
</tr>
<tr>
<td>CSP</td>
<td>Growing season precipitation.</td>
</tr>
<tr>
<td>MTW</td>
<td>Mean temperature of the warmest month.</td>
</tr>
<tr>
<td>MTCM</td>
<td>Mean temperature of the coldest month.</td>
</tr>
<tr>
<td>TDIFF</td>
<td>Degree-days &lt; 0°C accumulating within the frost-free period.</td>
</tr>
<tr>
<td>GSP</td>
<td>Summer precipitation.</td>
</tr>
<tr>
<td>GSDD5</td>
<td>Degree-days &gt; 5°C accumulating within the frost-free period.</td>
</tr>
<tr>
<td>MTCM</td>
<td>Length of frost-free period (days).</td>
</tr>
<tr>
<td>GSP</td>
<td>Growing season precipitation balance.</td>
</tr>
<tr>
<td>MTW</td>
<td>Degree-days &gt; 5°C (based on mean monthly temperature)</td>
</tr>
<tr>
<td>MAP</td>
<td>Degree-days &gt; 5°C (based on mean min. monthly temperature)</td>
</tr>
<tr>
<td>MTW</td>
<td>Julian date of the first freezing date in autumn.</td>
</tr>
<tr>
<td>MTW</td>
<td>Julian date of the last freezing date of spring.</td>
</tr>
<tr>
<td>CSP</td>
<td>Degree-days &gt; 5°C, mean temperature of the warmest month.</td>
</tr>
<tr>
<td>MTCM</td>
<td>Mean temperature of the warmest month.</td>
</tr>
<tr>
<td>TDIFF</td>
<td>Mean temperature of the coldest month.</td>
</tr>
<tr>
<td>GSP</td>
<td>Growing season precipitation balance.</td>
</tr>
<tr>
<td>GSDD5</td>
<td>Degree-days &gt; 5°C, mean temperature of the warmest month.</td>
</tr>
<tr>
<td>MTCM</td>
<td>Growing season precipitation balance.</td>
</tr>
<tr>
<td>TDIFF</td>
<td>Degree-days &gt; 5°C, mean temperature of the warmest month.</td>
</tr>
<tr>
<td>GSP</td>
<td>Growing season precipitation balance.</td>
</tr>
<tr>
<td>GSDD5</td>
<td>Degree-days &gt; 5°C, mean temperature of the warmest month.</td>
</tr>
</tbody>
</table>

---

Crookston and Rehfeldt (2008), as well as an ensemble of 17 global climate models (Supplementary Appendix S1 available on Dryad at https://dx.doi.org/10.5061/dryad.jd64) for the RCP60 (medium-high greenhouse gas concentration pathway) scenario for 2060 and 2090. These climate layers were restricted to the western United States. Additional climate variables were derived from these climate variables (Table 1). We removed one of any two variables that was highly correlated (Pearson correlation > 0.7), keeping the more biologically relevant variable of a correlated pair (Rehfeldt et al. 2014; Shinneman et al. 2016). This resulted in seven variables used to fit the final models (except for random forest models of the varieties—see below): growing season precipitation, summer–spring precipitation balance, degree days below 0°C, mean temperature of the warmest month, summer–winter temperature differential, annual moisture index, and precipitation ratio. Variables were log or square root transformed if needed to normalize data.

Lineage Distance Modeling for Haplotypes

Although predicting the niches of intraspecific populations is more feasible given the rise in research that uncovers genetic variation of species across their distributions, often there are not enough sampled occurrences of each population to define their climate niche space and fit robust ENMs, limiting the ability to hindcast or forecast their distributions (Rosauer et al. 2015). In addition, it may not be appropriate to fit an ENM directly to only known haplotype occurrences because these occurrences may be governed by other factors than environmental conditions, such as competition. In the case of ponderosa pine, there may be too few sampled occurrences among intraspecific populations to produce robust ENMs, and sampled occurrences alone most likely do not represent the entire climatic niche of each haplotype. Thus, to overcome low sample sizes of known haplotype occurrences, we employed a recently developed method, lineage distance modeling (Rosauer et al. 2015), that probabilistically estimates the likely distribution of intraspecific lineages, therefore, increasing the number of occurrence points to fit ENMs. This is accomplished by first fitting an initial ENM for a species, and then assigning pixels within the species distribution to haplotypes based on the distance to known haplotype occurrences, taking into account barriers to dispersal using a measured cost distance (Fig. 1). The cost-distance analysis is calculated as −log(habitat suitability) where habitat suitability is the probability of occurrence from the species ENM. The output is a probability map for each lineage, such that all estimated probabilities for a cell sum to the total probability of the species occupying the cell from the original ENM. For this analysis, ENMs were fit for the two varieties separately (as described below), rather than the entire species, because they have distinct climate niches (Norris et al. 2006). The resulting distribution maps for each haplotype are effectively defined by the niche of the variety and geographically constrained by theoretical dispersal limitations. Custom python code from Rosauer et al. (2015) was used in ArcMap 10.4.1 to develop the haplotype probability maps. Estimated probability maps for each haplotype were compared with known haplotype locations using the Brier Score (Brier 1950), a measure of the accuracy of probabilistic predictions to binary (presence/absence) observations calculated as the mean squared difference between the two. The lower the Brier Score, the better the predictions.

Variety and Haplotype Niche Modeling

To fit ENMs for each variety, 20,000 cells were subsampled from ponderosa pine distribution maps of the USDA Forest Service’s National Individual Tree Species Atlas (https://www.fs.fed.us/foreshare/applied_sciences/mapping-reporting/remote-sensing/index.shtml) that were also occupied by vegetation types that contain ponderosa pine in the U.S. Geological Survey National Gap Analysis Program (https://gapanalysis.usgs.gov) map (Fig. 1). Because both of these mapped sources were derived from classifications of satellite imagery, we used agreement between the two datasets for presence locations to minimize errors of commission (similar to Shinneman et al. 2016). Each of the occupied cells was assigned to one of the two varieties based on its geographic
FIGURE 1. Workflow of lineage distance modeling and ENM shown for var. ponderosa and haplotypes 1, 5, and 8. First, an ENM is fit for the variety using RF (Step 1). Then, lineage distance modeling is used to produce distribution maps for each haplotype (Step 2). These distribution maps are sampled, and ENMs are fit for each haplotype using RF, MARS, and NPMR (Step 3). Lastly, ENMs are fit for the variety using RF, MARS, and NPMR (Step 4). An ensemble distribution is created for each haplotype and the variety by averaging the RF, MARS, and NPMR models.

Moreover, we also explored the possibility that using raw probability of occurrence values from the lineage distance modeling (analogous to using abundance) would produce less-restrictive ENMs for haplotypes, but the projected distributions were similar (or even more restricted, see Supplementary Appendix S3 available on Dryad). We also found little difference in Brier Scores using raw probability of occurrence values compared with the threshold sampling procedure (Supplementary Appendix S3 available on Dryad).

Absences for each haplotype consisted of original absences from Shinneman et al. (2016) and absences generated in the previous step for the variety to which it belongs. We used the absences of each variety rather than of the entire species because the delineation of haplotypes between the varieties is robust based on mtDNA and nuclear DNA analyses (Potter et al. 2013; Potter et al. 2015).

For all haplotype and variety niche models, data were split into 10 random partitions of training (80%) and testing (20%) sets. To minimize uncertainty in model algorithm choice, three different ENM modeling approaches were used to fit each partition (i.e., resulting in 30 models per haplotype/variety): random forest (RF), multivariate adaptive regression spline (MARS), and non-parametric multiplicative regression (NPMR). RF and MARS models were fit and projected in R v 3.3.1 (R Core Team 2014) using the randomForest (Liaw and Wiener 2002) and earth (Milborrow 2016) packages, respectively. NPMR models were run in...
HyperNiche 2.28 (McCune and Mefford 2009). All 25-climate variables were input into the initial RF models built for the varieties that provided inputs for the lineage distance modeling to explore all climate variables and minimize under-prediction of variety distributions (RF randomly chooses predictor variables and eliminates those that are not improving model performance; therefore, removal of correlated variables a priori is not needed.) However, to facilitate compatibility of results among different modeling approaches, final variety and haplotype RF models were fit with the same seven uncorrelated climate variables used in the MARS and NPMR models (Table 1). Each individual haplotype and variety model was projected onto current and future climatic scenarios, and an ensemble model was created by first averaging the 10 model runs of ecological niche maps estimated from each modeling approach, and then averaging the three resulting maps per model approach for each climate scenario. We also combined the distributions of haplotypes within each variety (H1, H5, and H8 for var. ponderosa; H2, H3, H4, H6, and H7 for var. scopulorum) at each time period in order to observe how the distribution of the variety produced from ENM haplotypes compared with the distribution produced by an ENM fit with variety occurrences. These combined distribution maps represent the probability of at least one of the haplotypes occurring in a particular cell (i.e., by calculating \( 1 - X \), where X is the probability that all haplotypes are absent).

Model performance for each ENM (haplotypes and varieties) was tested using the area under the receiver operating curve (AUC), a metric used when comparing continuous outputs (ENM occurrence probabilities) to binary observations (presence/absence dataset sampled from the lineage distance method), ranging from 0 to 1 (value of 0.5 indicating model performance is no better or worse than random). We also evaluated model performance by comparing the known haplotype locations to the predicted occurrence probabilities using the Brier Score.

### Niche-Overlap Analyses

The amount of niche overlap and niche similarity between haplotypes was calculated following Broennimann et al. (2012), using the “PCA-env option”, in which a Kernel smoother is applied to densities of species occurrence in a gridded environmental space calibrated on the available environmental space. The D metric (Schoener 1970) measures the amount of niche overlap between two haplotypes in the gridded environmental space (D = 1: complete overlap; D = 0: no overlap):

\[
D = 1 - 0.5 \left( \sum_{ij} |z_{1ij} - z_{2ij}| \right)
\]

where \( z_{1ij} \) and \( z_{2ij} \) are the occupancies of entity 1 and 2, respectively, and i and j refer to the cell corresponding to the i\(^{th}\) and j\(^{th}\) bins of the environmental variables.

### Results

#### Lineage Distance Modeling for Haplotypes

Distribution maps of each haplotype from the lineage distance modeling are generally in agreement with known locations of each haplotype (Brier Scores < 0.2; Table 2; Supplementary Appendix S2 available on Dryad). However, there were limitations to this approach for certain haplotypes. Specifically, because the distributions of H2, H4, and H7 are geographically restricted (Fig. 3 in Supplementary Appendix S2 available on Dryad), fewer than 1000 occurrence points from the lineage distance maps for these haplotypes were available for ENM construction (Table 2; H4: \( n = 102 \), H2: \( n = 296 \), and H7: \( n = 253 \)). Furthermore, because the ENM for var. scopulorum predicted only low probabilities of occurrence (< 0.50) for the small and isolated populations of H2 in southern California/Nevada and H7 in Nevada, in those specific areas only the original haplotype locations from (Potter et al. 2013) were used to construct ENMs.
Variety and Haplotype Niche Models

ENMs for varieties and haplotypes were able to discriminate well between occupied and unoccupied cells (AUC scores > 0.9; Table 2), and predicted the occurrence of known haplotype occurrences well (Brier Score <= 2.0, Table 2). Variable importance differed for each variety and haplotype as well as among model type (Supplementary Appendix S4 available on Dryad).

Var. ponderosa.—The ensemble ENM for the variety matches its known distribution, although it slightly over-predicts in portions of the Great Basin (in Nevada, Southern Idaho, and Eastern Oregon) and the Wasatch Range in Northern Utah (Fig. 2a). Projections of the ENM for the variety to 2060 and 2090 (Fig. 2b and c) suggest a range contraction will occur, with a projected 24% loss in total area within high occurrence probability areas (> 0.5 occurrence probability) by 2090 (compared with the recent period) (Fig. 3), especially at lower elevations in central Oregon, northeastern Washington, and northeastern California. This trend is compensated slightly with potential range extension into higher elevations, especially in central Idaho and the southern
Sierra Nevada of California. Indeed the mean elevation for cells with high probability (> 0.5) of occurrence is projected to increase by 132 m on average by 2090 (compared with the recent period) (Fig. 3). Moreover, the mean probability of all occupied cells with > 0.5 probability in the recent period is generally projected to decrease by 0.02 and then by 0.53 in 2060 and 2090, respectively (Fig. 4). The combined distribution for the three haplotypes (H1, H5, and H8) within the variety is not as extensive compared with the variety model for the recent climate period (Fig. 2d), but similar distributional trends, such as loss in central Oregon, are projected through 2090 (Fig. 2e–f). However, for the combined haplotype distribution there is no loss in total area and only a 57 m increase in mean elevation for cells with high probability (> 0.5) of occurrence (Fig. 3), although mean probability values for all occupied cells in the recent period with > 0.5 probability occurrence are projected to decline for both 2060 and 2090, by 0.17 and 0.53, respectively (Fig. 4). Individual haplotype models generally project similar trends to the variety (Supplementary Appendix S5 available on Dryad), as highlighted here for H1 (Fig. 2g–i), which is projected to have a decrease in area (41%), an increase in mean elevation (132 m), and a decrease in mean probability (by 0.63, of occupied cells in the recent period with > 0.5 probability) for 2090 (Figs. 3 and 4).

**Var. scopulorum.**—The modeled niche distribution of var. scopulorum matches the known distribution of the variety, although there is some over-prediction of the distribution in Montana and under prediction in the Great Plains (Fig. 5a). In general, the ensemble ENM of the variety and the individual haplotypes (H3 and H6) could not accurately model the northern portion of the variety’s distribution. The easternmost populations of the variety along the Niobrara River in Nebraska are not predicted by the ensemble ENMs most likely because these populations exist in microclimates along the river that are not available in surrounding areas, making it difficult for coarse scale climate data (800 m x 800 m resolution) to capture their presence.
Similar to var. *ponderosa*, ENM projections for var. *scopulorum* in the future suggest a range contraction, with a 59% area loss and a 670 m increase in mean elevation by 2090 compared with the recent time period for areas with > 0.5 occurrence probability (Fig. 3). There is some predicted expansion northward into central Utah and western Colorado, and some expansion into southwest Montana even though there is predicted loss for most Montana populations. Occupied cells with probabilities > 0.5 in the recent period decline by 0.17 in 2060 and by 0.57 in 2090 (Fig. 4). The combined predicted niche distributions of the individual haplotypes within the variety (H2, H3, H4, H6, and H7) are more delimited than the variety (similar to the pattern observed with var. *ponderosa* haplotypes), but they have a similar pattern of loss and upward range shift as the variety, with a 23% loss in area and a mean elevational increase of 363 m for high probability (> 0.5) occurrences by 2090 (Fig. 3). The probability of the combined haplotypes occurring in predicted occupied sites in the recent time period with > 0.5 probability in the recent time period decreases by 0.61 in 2090 (Fig. 4). Individual haplotypes within var. *scopulorum* follow a similar trend (Supplementary Appendix S5 available on Dryad), as highlighted by H3 (Fig. 5g–i), which is projected to have a decrease in area (40%), an increase in mean elevation (422 m), and a decrease in mean probability (by 0.63, of occupied cells in the recent period with > 0.5 probability) for 2090 (Figs. 3 and 4).

**Niche Overlap**

Haplotypes of ponderosa pine occupy distinct niche spaces in western North America, as no two haplotypes had statistically similar niche spaces (Similarity Test $P < 0.01$; Supplementary Appendix S6 available on
Dryad). Thus, even though the niche spaces of some haplotypes overlap (e.g., H3 and H7), they are not significantly similar (Fig. 6 and Supplementary Appendix S6 available on Dryad). Haplotypes with larger distributional ranges tend to occupy more niche space (e.g., H1 vs. H4; Fig. 6; Supplementary Appendix S6 available on Dryad). Those haplotypes within var. *ponderosa* occupy environmental space that contains higher winter precipitation than haplotypes within var. *scopulorum*, in agreement with previous analyses (Norris et al. 2006; Shinneman et al. 2016; Fig. 6).

Within var. *ponderosa*, H1 and H8 have the most overlap (D = 0.497), whereas H8 and H5 have very little overlap (D = 0.014). Within var. *scopulorum*, the most overlap is between H3 and H7 (D = 0.324) and H3 and H2 (D = 0.244) whereas H4 is the most isolated (D < 0.05 with all haplotypes within the variety). The niche space of H4, however, overlaps with the niche space of H1 (D = 0.126) and H8 (D = 0.260) of var. *ponderosa* (Potter et al. 2013) and the niche-overlap results support this (indeed, H2 overlaps with H3 [D = 0.244] more than other haplotypes). However, the low niche overlap between H2 and geographically proximate haplotypes in California associated with var. *ponderosa*, H1 (D = 0.023) and H5 (D = 0.014), could also be influenced by the lack of H2 presences predicted for southern California by the lineage distance modeling.

**DISCUSSION**

**Comparisons Among Projected ENM Distributions for Varieties Versus Haplotypes**

Phylogeographic structures defined by genetic variation may represent clearly defined and distinct niche spaces among lineages within a species that reflect the evolutionary capacity of a species to adapt to environmental change (Prasad and Potter 2017). Thus, modeling the ecological niches of intraspecific lineages may project distributions of each lineage that differ in important ways from projections for the species (Hällfors et al. 2016). For instance, using the fundamental niche of simulated conceptual species and subpopulations, Valladares et al. (2014) found more restricted distributions when modeling distinct subpopulations compared with modeling the entire species. In contrast, Pearman et al. (2010), Oney et al. (2013), and Valladares et al. (2014) found the opposite when modeling the realized niches of real species and populations within them. Valladares et al. (2014) attributed such differences to modeling the fundamental versus realized niches. However, our work suggests an alternative explanation is in order, as our realized niches of the haplotypes were generally more limiting than those of the varieties. Estimates derived from the collective representation of the haplotype probabilities for the recent period are also generally more geographically restricted and have higher average occurrence probabilities than those projected for the varieties (Figs. 2, 4, and 5), and these more restricted ranges are further amplified for most individual haplotypes under future climate change compared with the varieties (Supplementary Appendix S5 available on Dryad). Although this could be partly caused by the influence of haplotype presence locations that restrict estimates of the peripheral climate niche, we suggest that another plausible explanation is that the haplotype populations correspond to generally well defined and often minimally overlapping climate niches (Fig. 6) that are uniquely projected to decrease in the future. In contrast, models for the varieties capture a broader niche representing more potential combinations of climate conditions, a larger portion of which may...
persist into the future. It has been suggested that by not including intraspecific variation in ENMs, fitness-environment curves may be smoothed due to averaging across the broader niche, leading to more pessimistic forecasts of future distributions (Valladares et al. 2014). However, our variety models did not result in more unfavorable distributions (i.e., better delineated with higher probabilities) relative to the haplotype models.

Regardless of intraspecific emphasis, by examining relative changes across models (based on probability estimates > 0.5), we were able to reveal potential trends that generally hold for both the variety and most corresponding haplotype models under climate change. For the Pacific variety (var. *ponderosa*), there is a general loss of area projected under climate change (by ~25%) and a modest increase in elevation (~200 m). Although individual haplotype responses were more variable, they generally supported trends of decreasing range size and increasing elevation projected by the variety model under future climates; however, the combined haplotype distributions did not (Figs. 2 and 3). Projected trends for the Rocky Mountain variety (var. *scopulorum*) were similar but more extreme (~50% decrease in area, and >600 m elevation gain) than the Pacific variety, and were also generally supported by the more highly variable individual haplotype models, and to a lesser degree by the combined haplotype models (Figs. 3 and 5). These trends suggest that the individual haplotype models effectively captured inherent sub-specific variability in response, but that the combined haplotypes were less likely to decline overall because their ranges were already more restricted for the recent period (relative to ranges for the varieties), or because they were buoyed by a few favorable individual haplotype responses to climate change (e.g., projected distributions for H5, see Discussion below).

We expected some predicted distribution loss in the future, especially given that extrapolation of ENMs fit with present-day climates and projected to novel climates tend to under-predict distributions (Maguire et al. 2016). This factor could be further enhanced when ENMs are fit with a narrower subset of occurrences, restricting potential future distributions. However, climate novelty is likely minimal in our models (<0.0003 following methods in Maguire et al. 2016), and the varying degrees of predicted area loss for haplotypes and varieties suggest potentially true signals in their projected responses to climate change.

Indeed, the potentially unique responses of individual haplotypes to future climate within each ponderosa pine variety can be instructive, especially in terms of patterns of area loss and elevational shifts. Given likely long-term isolation and local evolution of each of the haplotypes within glacial periods, each may be expected to exhibit isolation and local evolution of each of the haplotypes across the broader niche, leading to more pessimistic forecasts of future distributions (Valladares et al. 2014). The variety already occupies higher elevational areas in much of this part of its range, and there are large distances to the nearest analogous climate in the future, effectively restricting both elevational and latitudinal range shifts (Rehfeldt et al. 2014, Roberts and Hamann 2016). Haplotype 6 may also be vulnerable because it occupies a unique climate niche space (e.g., occupies higher values of growing season precipitation, precipitation ratio, and growing degree days below 0°C other than haplotypes) that is on the edge of the total niche space for var. *scopulorum* (Fig. 6 and Supplementary Appendix S6 available on Dryad).

Relevance to Adaptation and Phenotypic Plasticity

To successfully adapt to climate change, plant species must either migrate to track their ecological niches or increase their tolerance to new climate conditions in situ. This latter strategy requires either expression of phenotypic plasticity, or selection on existing phenotypes defined by existing genetic material or on novel phenotypes resulting from favorable mutations (Williams et al. 2008; Chevin et al. 2010). Our haplotype ENMs operated under an assumption of adaptation to specific, realized climate niches; they did not directly incorporate plasticity or genetic adaptation under changing climate, and thus did not capture the ability of each haplotype to expand into or occupy other parts of the species' fundamental niche. Moreover, we caution that the haplotypes, which are classified based on identification of neutral genetic markers, are not necessarily effective surrogates for uniquely adapted populations (e.g., Holderegger et al. 2006). However, although these mtDNA defined haplotypes do not necessarily have adaptive significance, they correspond to populations that have been influenced by long-term biogeographical processes.

Previous studies of ponderosa pine have shown the existence of adaptive phenotypic variation within populations (e.g. Conkle 1973; Rehfeldt 1986a, 1986b, 1990; Namkoong and Conkle 1976; Sorensen and Weber 1994; Sorensen et al. 2001; Keller et al. 2004), and
populations with greater phenotypic plasticity may be more tolerant of changing climates (Chevin et al. 2010). There are areas within the distribution of ponderosa pine with high genetic variation (based on nuclear DNA, see Potter et al. 2015), and greater genetic variation can allow populations to adapt to climate change via natural selection (Sork et al. 2010). In addition, high gene flow has been detected within each ponderosa pine variety, occurring across complex structures of genetic variation (Potter et al. 2015). Gene flow between even the most genetically structured populations can reduce population differentiation and relax migrational constraints (Roberts and Hamann 2016).

Given the potential for plasticity or genetically adaptive responses of ponderosa pine to changing climate, as well as shifts in biotic interactions, population-level responses may be best captured by developing ENMs for the varieties. Indeed, when reduced to two dimensions via the niche-overlap analysis, the collective ponderosa pine haplotype climate niche space does not cover the entire niche space of each variety (Fig. 6), potentially suggesting that the modeled niches of the varieties could indirectly represent more plasticity or adaptive potential for the species than the individual haplotypes combined. The ENMs of varieties may also be indirectly capturing intraspecific variation, and therefore, the plasticity of the variety, reflected in the projected broader geographic distributions and higher probabilities under climate change compared with most haplotypes. This highlights the complementary value of modeling at both the variety and haplotype levels, in relation to phenotypic plasticity and adaptation respectively, and providing a better understanding of how variability in intraspecific response might affect the future niche distribution of the species as a whole (Pauls et al. 2013; Marcer et al. 2016).

Dispersal, Migration, and Competition

Even though optimal climate for ponderosa pine is generally predicted to shift to higher elevations in the future, existing populations may not fill these areas because of dispersal limitations and migration lags (Davis 1989). Thus, rates of migration on leading edges may be slower than rates of habitat loss on trailing edges of existing populations (Campbell and Shinneman 2017). Yet, although ponderosa pine has limited optimal seed dispersal distance for regeneration (15–30 m; Oliver and Ryker 1990; Boyden et al. 2005; Latta et al. 1998), it is capable of low frequency, long-distance dispersal of 75–200 km (Lesser and Jackson 2012, 2013). Indeed, this likely explains its rapid expansion and northward migration from refugial areas during the Last Glacial Maximum, as supported by fossil evidence (Norris et al. 2016) and by a pattern of decreasing unique and rare alleles with increasing latitude found in contemporary populations (Potter et al. 2015). These paleogeographic range shifts may suggest potential for a similar response to future climate change. However, requirements of future migrations may be different from the past due to novel climate conditions, different refugia locations, and new migrational constraints imposed by human activity (Norris et al. 2016; Roberts and Hamann 2016). In addition, because realized climate niches of ENMs (from which the ENMs were built) may be truncated due to biotic interactions; shifting interspecific competition under climate change could also affect future distributions (Wisz et al. 2013).

Potential Genetic Variation and Diversity Under Future Climate

Genetic variation and diversity are likely to be affected differently for ponderosa pine populations under future and potentially novel climate conditions compared with paleoclimates. For instance, based on fossil and genetic evidence (Potter et al. 2013), the contact zone between the two varieties in central Montana is relatively recent, established within the previous 1500 years, and likely controlled by a geographic shift from summer- to winter-dominant precipitation (Norris et al. 2006). Chloroplast and nuclear DNA patterns suggest introgression from west to east in this region (Latta and Mitton 1999; Potter et al. 2015), but there was no exchange of mtDNA (Latta and Mitton 1999; Johansen and Latta 2003; Potter et al. 2013). Potential genetic exchange between haplotypes and varieties in this location could be further restricted under future climate, based on the lack of overlap projected in our models for either the varieties or haplotypes. In addition, some haplotypes could become increasingly isolated under warmer and drier climates (e.g., H4 and H7). Thus, some genetic diversity may be lost, especially if isolated lineages have less suitable habitat and experience even greater geographic isolation under future climates. This underscores the need to identify potential future refugia across intraspecific levels that may offer the best chance for maintenance of genetic diversity and survival of the species under climate change (e.g., Keppel et al. 2012).

Conclusions

Our results highlight the potential value of modeling intraspecific response to climate change and, more specifically, are informative for future management and conservation of ponderosa pine, an economically, and ecologically important tree species. Ponderosa pine haplotype groups may represent evolutionary lineages adapted to different environmental conditions. Thus, by identifying lineages most vulnerable to climate change and detecting potential climatically suitable areas in the future, our results suggest at least three important considerations for successful conservation. First, it is significant that each haplotype occupies a statistically different environmental space within the range of ponderosa pine, as this could be considered evidence for differential adaptation to diverse environmental
conditions and, thus, that the haplotypes may be evolutionary units worthy of distinct conservation attention. However, there is need for further study on adaptive genetic variation in ponderosa pine that could help to identify evolutionarily significant units for conservation, including defining genetic groups differently (i.e., based on nuclear DNA variation; Potter et al. 2015). Second, the results indicate that some haplotype groups, specifically H8 in var. scopulorum and H6 in var. ponderosa, could lose well more than half of their predicted distribution before 2090. These haplotypes may require special efforts to conserve their existing genetic variation, using both ex situ and in situ strategies (Potter et al. 2017). Peripheral populations of isolated lineages, such as H4 and H7, may also be at high risk under climate change, and due to their lower genetic diversity but greater genetic differentiation compared with populations in the core range of each ponderosa pine variety (Potter et al. 2015), may be ideal candidates for conservation action. Finally, the results of this study could help guide assisted species and population migration efforts (Dumroese et al. 2015; Aitken and Bemmels 2016), by identifying locations in which vulnerable ponderosa pine lineages may be best adapted in the future.

SUPPLEMENTARY MATERIAL

Data available from the Dryad Digital Repository: http://dx.doi.org/10.5061/dryad.jj264.

FUNDING

This work was supported in part by Cost Shares Agreement [14-CS-1130110-042 and 15-CS-1130110-067 to K.M.F.] between the Southern Research Station of the USDA Forest Service and North Carolina State University. Funding for developing the climate niche models and niche-overlap analysis presented in this article was provided by an Interagency Agreement [I.L2PG00219 to D.J.S.] between the U.S. Bureau of Land Management and the U.S. Geological Survey.

ACKNOWLEDGMENTS

Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government. We would like to thank Dan Rosauer, Nicholas Matzke, and Brittany Barker for their very helpful reviews of this article. We would also like to acknowledge Robert Means, whose efforts helped to initiate this research, and who passed away before the work was completed.

REFERENCES


Downloaded from https://academic.oup.com/sysbio/advance-article-abstract/doi/10.1093/sysbio/syy017/4934308 by U S Dept of Agriculture user on 29 May 2018


