Evolution, 2024, **78(2)**, 300–314
https://doi.org/10.1093/evolut/qpad207
Advance access publication 14 November 2023
Original Article

Environment predicts the maintenance of reproductive isolation in a mosaic hybrid zone of rubber rabbitbrush

Trevor M. Faske[1,2], Alison C. Agneray[1,2,3], Joshua P. Jahner[4], Carolina Osuna-Mascaró[1], Lana M. Sheta[1], Bryce A. Richardson[5], Elizabeth A. Leger[1,2], Thomas L. Parchman[1,2]

[1]Department of Biology, University of Nevada, Reno, Reno, NV 89557, United States
[2]Ecology, Evolution, and Conservation Biology, University of Nevada, Reno, NV 89557, United States
[3]Bureau of Land Management, Nevada State Office, Reno, NV 89502, United States
[4]Department of Botany, University of Wyoming, Laramie, WY, 82071, United States
[5]Rocky Mountain Research Station, USDA Forest Service, Moscow, ID 83843, United States

Corresponding author: Department of Biology, University of Nevada, Reno, 1664 N. Virginia Street, Reno, NV 89557, United States. Email: tfaske@nevada.unr.edu

Abstract

Widely distributed plants of western North America experience divergent selection across environmental gradients, have complex histories shaped by biogeographic barriers and distributional shifts and often illustrate continuums of reproductive isolation. Rubber rabbitbrush (*Ericameria nauseosa*) is a foundational shrub species that occurs across diverse environments of western North America. Its remarkable phenotypic diversity is currently ascribed to two subspecies—*Ericameria nauseosa nauseosa* and *Ericameria nauseosa consimilis*—and 22 named varieties. To understand how genetic variation is partitioned across subspecies, varieties, and environments, we used high throughput sequencing of reduced representation libraries. We found clear evidence for divergence between the two subspecies, despite largely sympatric distributions. Numerous locations exhibiting admixed ancestry were not geographically localized but were widely distributed across a mosaic hybrid zone. The occurrence of hybrid and subspecific ancestries was strongly predicted by environmental variables as well as the proximity to major ecotones between ecoregions. Although this repeatability illustrates the importance of environmental factors in shaping reproductive isolation, variability in the prevalence of hybridization also indicates these factors likely differ across ecological contexts. There was mixed evidence for the evolutionary cohesiveness of varieties, but several genetically distinct and narrow endemic varieties exhibited admixed subspecific ancestries, hinting at the possibility for transgressive hybridization to contribute to phenotypic novelty and the colonization of new environments in *E. nauseosa*.

Keywords: admixture, *Ericameria*, genetic-environment association, landscape genomics, mosaic hybrid zone, reduced representation sequencing

Introduction

Evolutionary biologists have long sought to understand the processes causing phenotypic and genetic divergence between lineages, along with the origin and maintenance of reproductive isolation (Nosil & Feder, 2012; Schluter, 2001; Thompson, 2005). As a result, much work has focused on populations spanning the continuum of divergence and reproductive isolation that may characterize the progression of speciation (Bolnick et al., 2023; Stankowski & Ravinet, 2021). Widely distributed plant taxa of western North America often display extensive variation in phenotype and connectivity due to divergent selection across heterogeneous landscapes and strong environmental gradients (e.g., Baughman et al., 2019), as well as historical periods of isolation owing to biogeographical barriers and distributional shifts associated with Quaternary climate oscillations (Dynesius & Jansson, 2000; Massatti & Knowles, 2020; Massatti et al., 2018; Shafer et al., 2010). Temporal and spatial variation in these processes may lead to a continuum of phenotypic and genetic divergence with variable reproductive isolation among populations and species (i.e., syngameons; Buck et al., 2023; Chhatre et al., 2018; Grant, 1981; Lotsy, 1925). Plant groups composed of complexes of interbreeding and isolated lineages are relatively common in western North America due to topographic and environmental complexity and can be valuable for understanding how ecological variation underlies adaptation, speciation, and hybridization.

The outcomes of hybridization between lineages depend on hybrid fitness and the genetic basis of reproductive isolation (Barton, 2001). Understanding the factors that prevent or promote hybridization can provide insight into the genetic and environmental causes of divergence, isolation, and gene flow. Ecological factors, such as variation in specialization, species interactions, or phenology, can mediate prezygotic barriers to reproduction (e.g., Hood et al., 2019). Postzygotic barriers can also be influenced by ecological factors, such as reduced hybrid fitness in either the parental environment or by genetic interactions (Maheshwari & Barbash, 2011; Orr, 1996). Gene flow occurring through hybridization can obscure species boundaries, blur phylogenetic relationships, and contribute to taxonomic uncertainty (McVay et al., 2015; Zhang et al., 2021). Furthermore, hybridization can contribute to adaptation and evolutionary novelty through both adaptive introgression (e.g., Calfee et al., 2020; Giska et al., 2019) and transgressive variation in hybrids that may facilitate the colonization of novel environments (Gompert et

Received July 31, 2023; revisions received October 23, 2023; accepted November 9, 2023
Associate Editor: Jen-Pan Huang; Handling Editor: Jason Wolf

al., 2006; Hodel et al., 2022; Rieseberg et al., 1999). While many instances of hybridization are structured as geographic clines across major ecotones (Barton & Hewitt, 1985; Endler, 1977), more complex outcomes can arise in the form of mosaic hybrid zones, where hybrids between lineages occur at multiple locations across a mosaic of environmental variation (Rand & Harrison, 1989). Geography and environment are decoupled in mosaic hybrid zones, offering replicate views into the predictability of reproductive isolation and the environmental factors underlying it. Here we evaluate the ecological predictors of divergence and secondary contact in a morphologically and taxonomically diverse foundational shrub species found throughout the geologically complex, arid regions of western North America.

Ericameria (Asteraceae) is a genus of approximately 70 shrub species distributed in western North America. Rubber rabbitbrush (*Ericameria nauseosa*) is broadly distributed and abundant across desert, woodland, and arid montane habitats, where it thrives as a primary successional colonizer of dryland and disturbed sites. It is a foundational species with deep root systems that stabilize soil for dryland plant communities (Donovan & Ehleringer, 1994) and provides important resources for birds, mammals, and insects. It hosts an exceptional number of gall-forming insects (Baine et al., 2023; Fernandes et al., 2000; Floate et al., 1996; Mcarthur et al., 1986) and represents a keystone late-season flowering resource for diverse pollinator communities (Carril et al., 2018; Hansen, 1986; McArthur et al., 1979; Ogle et al., 2007; Rogers, 1979). Across the environments and geographic regions it occupies, *E. nauseosa* displays substantial phenotypic variation in its stature, morphology, and coloration, and is noted for its exceptional phytochemical diversity (Hegerhorst et al., 1987a, b; Weber et al., 1985). Pronounced population differentiation in germination timing has also illustrated local adaptation to variation in temperature across its distribution (Meyer et al., 1989). Given its ecological significance, phenotypic and genetic variation in *E. nauseosa* are likely to have extended community and ecosystem-level consequences (Grady et al., 2011; Whitham et al., 2003) and to have value for the restoration of native plant communities across much of western North America (Faske et al., 2021).

Due to extensive phenotypic differentiation, taxonomists have recognized numerous subspecies and varieties within *E. nauseosa*, and taxonomic revisions have been common. *Ericameria nauseosa* was moved from *Chrysothamnus* to *Ericameria* (Nesom & Baird, 1993) and subspecific classifications have shifted from 22 recognized subspecies (Anderson, 1986a; Hall, 1919) to the current designation of two recognized subspecies and 22 named varieties (Nesom & Baird, 1993). The two recognized subspecies (*Ericameria nauseosa nauseosa* [grey] and *Ericameria nauseosa consimilis* [green]) have overlapping distributions but are phenotypically distinguishable by variation in grey and green pigmentation in the leaves and stems (Anderson, 1986b). The 22 named varieties vary in vegetative and floral morphology and in distribution, with some varieties spanning much of western North America while others are endemic soil specialists with small areas of occupancy (Anderson, 1986a; Nesom & Baird, 1993). There is distinctive phytochemical variation between both subspecies and among varieties (Hegerhorst et al., 1987a; b; Weber et al., 1985) and this variation is associated with distinct insect and herbivore communities (Floate et al., 1996; Mcarthur et al., 1986; McArthur et al., 1979; Wangberg, 1981). The subspecies and certain varieties are sympatric in some locales, and while hybridization has been noted or suspected (Anderson, 1973; Hanks et al., 1975), the persistence of differentiation suggests some degree of reproductive isolation (Anderson, 1986b; Anderson & Reveal, 1966; McArthur et al., 1979). Thus, the phenotypic and taxonomic diversity in *E. nauseosa* could reflect diversification driven by adaptation to different environments that have resulted in a series of independently evolving lineages across a continuum of diversification.

As the few prior population genetic analyses of the group have been limited to single varieties or subspecies (Faske et al., 2021; Gang & Weber, 1995), we lack an understanding of the evolutionary cohesiveness of taxonomic designations and the potential for genetic differentiation and reproductive isolation among such groups. Furthermore, while hybridization among subspecies and varieties has been anecdotally suspected, we have no empirical understanding of its actual extent. High-throughput sequencing approaches have improved our ability to resolve the early phases of diversification (Buck et al., 2023; Chhatre et al., 2018; R. Fu et al., 2022; Wagner et al., 2013), to characterize the influence of environment on genetic variation (Forester et al., 2016, 2018; Rellstab et al., 2015), and quantify ancestry variation among and within admixed populations (Gompert et al., 2017; Moran et al., 2021). Geographic variation in the distribution of hybrid ancestry can provide insights into the ecological and environmental factors mediating reproductive isolation and shaping divergence in diverse plant assemblages.

Here we quantified range-wide genetic variation in *E. nauseosa* using high-throughput reduced representation sequencing for 76 sampling localities spanning taxonomic variation and a substantial portion of the geographic and environmental breadth of its distribution. We first characterized the degree of genetic structure between and within subspecies and asked how it has been shaped by geography, environment, and history. Second, given past observations of suspected or confirmed hybrids, we quantified the extent and predictors of hybridization across genetic lineages. Finally, given that the broad morphological and ecotypic diversity in *E. nauseosa* has been partitioned into 22 named varieties, we evaluated the evolutionary and genetic cohesiveness of putative varietal designations. Our results reveal a continuum of genetic divergence and reproductive isolation spanning lineages, environments, and phenotypes that will guide future population genomic work in this syngameon-like foundational species and provide a useful perspective for restoration in western North America.

Methods

Plant sampling and taxonomic identification

We collected fresh leaf tissue from 76 sampling locations (608 individuals; 2–15 per location; Supplementary Table S1), spanning a substantial portion of the *E. nauseosa* range in the western United States (Figure 1). During field collection, plants were identified to subspecies and variety for most locations (Cronquist, 1994). Our sampling ultimately captured 14 of the 22 named varieties. Subspecific designations of the sampled varieties follow those currently ascribed to *E. n. nauseosa* (varieties: *bigelovii*, *graveolens*, *hololeuca*, *iridis*, *latisquamea*, *nana*, *salicifolia*, and *speciosa*) and *Ericameria nauseosa consimilis* (varieties: *arenaria*, *juncea*, *mohavensis*,

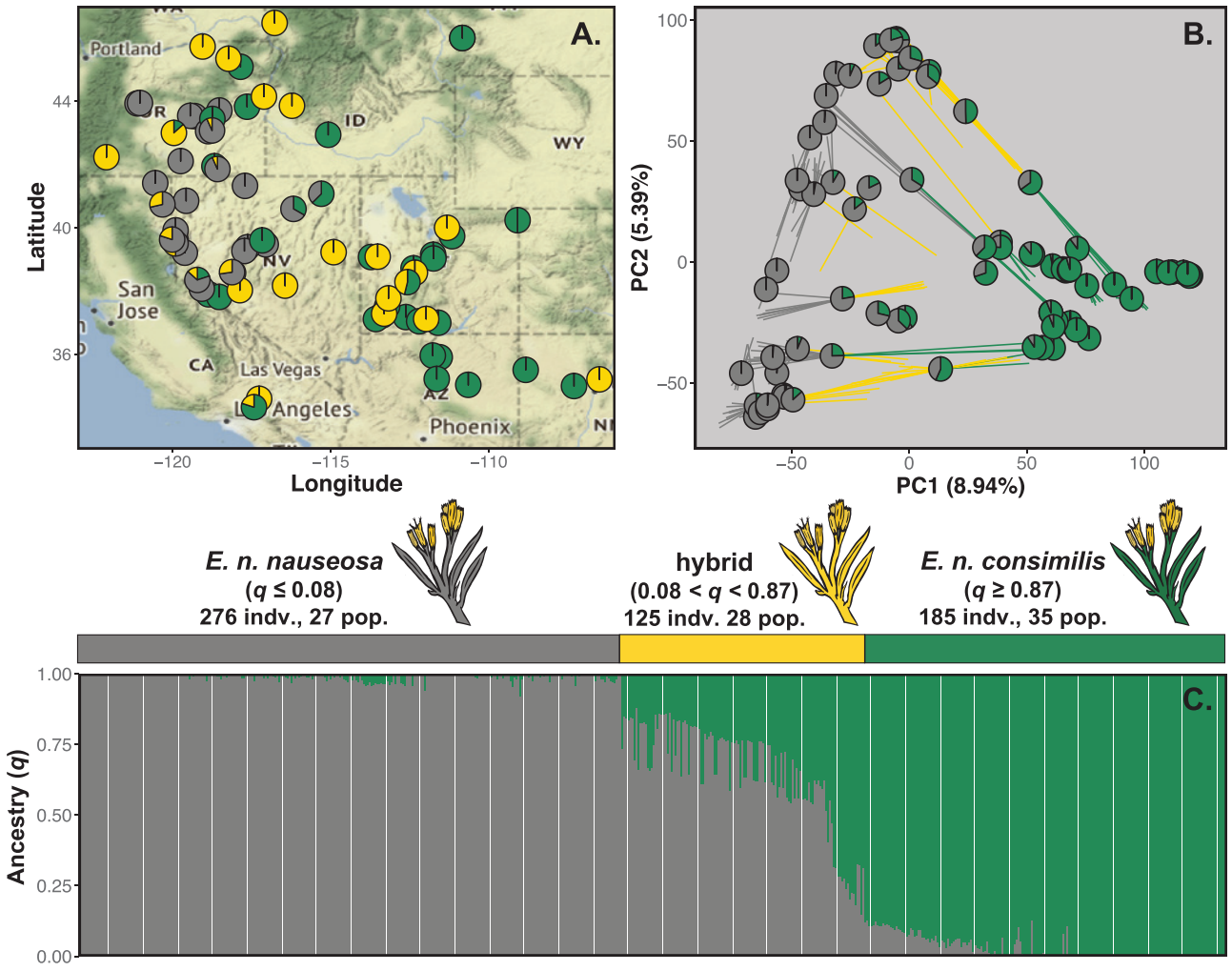

Figure 1. *E. nauseosa* lineages, *E. n. nauseosa* and *E. n. consimilis*, exhibit genetic differentiation and admixture occurs in multiple locations. (A) Map illustrating sampling locations. Pie chart colors correspond to the proportion of individuals at a locality assigned to each lineage or as a hybrid. (B) Principal component analysis (PCA) illustrates genetic differentiation between lineages and the presence of hybrid individuals. Circles represent mean principal component (PC) values for populations, pie charts reflect ancestry, and lines connect PC values for individuals to population means. Yellow lines represent hybrid individuals with ancestry coefficients $0.08 < q < 0.87$. (C) Ancestry coefficients (q) for each individual from the hierarchical Bayesian model of ENTROPY for $k = 2$, ordered by PC1. Color proportion of each bar ancestry proportion. See Supplementary Figure S1 for similar analyses conducted on a downsampled dataset.

nitida, *oreophila*, and *turbinata*). Individuals from locations where phenotypic assignment to variety proved difficult were designated as unknown; this was at times due to the potential existence of more than a single subspecies or variety in a single location and the difficulty of using phenotypic characters in dichotomous keys. Three samples of *Ericameria discoidia*, a sister taxon known to hybridize with *E. nauseosa* (Anderson & Reveal, 1966; Roberts & Urbatsch, 2003), were additionally included as a phylogenetic outgroup.

Reduced representation sequencing, bioinformatic processing, and variant calling

Dried leaf material from each plant was ground into powder form using a Qiagen TissueLyser and DNA was extracted using Qiagen DNeasy Plant kits (Qiagen, Valencia CA). Reduced representation libraries were prepared using a double-digest restriction-site associated DNA sequencing (ddRADseq) method (Parchman et al., 2012; Peterson et al., 2012). The restriction endonucleases *Eco*RI and *Mse*I were used to digest genomic DNA and uniquely barcoded Illumina adaptors were subsequently ligated to the resulting fragments using T4 DNA ligase (New England Biolabs, Ipswich MA). Barcoded fragments were polymerase chain reaction (PCR) amplified with Iproof DNA polymerase (BioRad, Hercules CA), and fragments ranging from 350 to 450 bp were size selected using a Pippin Prep quantitative electrophoresis unit (Sage Science, Inc). Single-end sequencing (100 bp read lengths) was performed using two lanes on an Illumina HiSeq 4000 platform at the University of Wisconsin.

Sequences representing potential contaminants (*Escherichia coli* and PhiX) or various Illumina-associated oligos were detected and discarded using a pipeline of Perl and bash scripts (http://github.com/ncgr/tapioca). A custom Perl script was then used to correct barcode sequencing errors, to trim cut site and barcoded oligo-associated bases, and to demultiplex reads by individual. As a reference genome is not available for *E. nauseosa* or a close relative, we used a de novo assembly of unique reads to create a consensus reference of genomic regions sampled by our reduced representation approach (reference assembly, hereafter). Parameters for this step were

optimized using shell scripts and documentation provided for dDocent (Puritz et al., 2014; cutoffs: individual = 8, coverage = 5; clustering similarity: -c.92), and the reference assembly was produced using CD-HIT-EST v4.8.1 (Fu et al., 2012; Li & Godzik, 2006). Demultiplexed reads for each individual were mapped to the reference assembly using BWA-MEM v0.7.17 (Li, 2013). Sequence variants were identified using SAMTOOLS v1.9 and BCFTOOLS v1.9 (Li et al., 2009), and subsequent filtering was completed using VCFTOOLS v0.1.16 (Danecek et al., 2011). We retained only biallelic loci covered by reads present in at least 70% of individuals, thinned to one locus per contig to reduce the effects of linkage disequilibrium and sequencing error. In addition, individuals missing data for greater than 40% of loci were removed before further analyses. After additional filtering with custom Python scripts, we retained loci with minor allele frequency (MAF) ≥ 0.02, mean read depth across individuals ≥ 3 or < 25, and alternate allele call quality ≥ 750. Filtering cutoffs were decided based on distributions of the metrics in question (see the filtering script on the associated GitHub for summary statistics), with the goal of retaining the highest quality loci without sacrificing genomic coverage. We removed loci with excessive read depth and only retained loci with two alleles present to ameliorate genotyping bias from the potential misassembly of paralogous genomic regions (Hapke & Thiele, 2016; McKinney et al., 2018). Lastly, we retained loci with $F_{IS} < -0.5$, as misassembly of paralogous genomic regions can lead to abnormal heterozygosity (Hohenlohe et al., 2013; McKinney et al., 2017). To detect duplicated, clonal, or highly related individuals, we quantified relatedness among all pairs of individuals within each population using VCFTOOLS, employing the cutoff criteria provided by the KING method (Manichaikul et al., 2010; Turner et al., 2022).

Landscape genetic structure

To estimate genotype probabilities for each individual at each locus, infer the number of ancestral genetic clusters (k), and estimate individual ancestry coefficients (q), we utilized a hierarchical Bayesian model that incorporates genotype uncertainty (ENTROPY v2.0; Gompert et al., 2014; Shastry et al., 2021). The model employed by ENTROPY utilizes an allele frequency prior and incorporates genotype uncertainty arising from sequencing and alignment error, as well as stochastic variation among individuals and loci in sequencing depth during parameter estimation (Gompert et al., 2014). Using genotype likelihoods calculated from SAMTOOLS, we executed k-means clustering of the first five principal components (PCs) to generate starting values of ancestry coefficients (q) to seed the Markov chain Monte Carlo (MCMC) (Gompert et al., 2014). ENTROPY was used to estimate genotype probabilities and ancestry coefficients for models based on a range of ancestral genetic clusters (k). We ran 60,000 MCMC iterations across four chains with a burn-in of 10,000, thinning every tenth step, for models based on each level of k. Genotype probabilities from the top five models based on DIC (k = 2–6) were averaged across all chains and used for all subsequent population genetic analyses (Supplementary Table S2). To verify that ordination and ancestry analyses were not biased by uneven sampling (McVean, 2009; Novembre & Stephens, 2008; Sinclair & Hobbs, 2009), we filtered variants, inferred ancestry coefficients, and conducted population genetic analyses as described below on a reduced data set with even sampling across sampling locations and subspecies (231 individuals, 47 locations). Because patterns of clustering and population differentiation were qualitatively similar for both datasets (see Figure 1; Supplementary Figure S1), irrespective of sampling density, we used the full data set for analyses detailed below. As a complementary model-free approach to quantify genetic variation among all individuals spanning the named subspecies and varieties, we conducted a principal component analysis (PCA) on genotype probabilities, following the standardization proposed by Patterson et al. (2006).

Both PCA and ENTROPY suggested the two lineages were genetically differentiated and sympatric (Figure 1), with hybrids throughout the sampled range, so we additionally conducted population genomic analyses separately for sets of individuals with pure ancestry for each lineage (excluding hybrids). We use the term "lineage" along with the subspecific nomenclature (*E. n. nauseosa* and *E. n. consimilis*) to refer to these genetic groups, overriding field identifications in cases where subspecies designation was incongruent with genetic data. Specifically, individuals were categorized by ancestry coefficients (q) based on cut-offs that made sense both geographically and biologically (Figure 1C; *E. n. nauseosa* [grey] < 0.08 ≤ hybrid ≤ 0.87 < *E. n. consimilis* [green]). Sampling locations containing individuals across multiple ancestral categorizations (i.e., *E. n. nauseosa*; *E. n. consimilis*; or hybrid ancestry n = 13 locations) were treated as separate populations. Variants were sampled independently from the genotype probabilities and re-filtered with a MAF ≥ 0.02 (*E. n. nauseosa*: 18,366 loci, 276 indv., 27 pop.; *E. n. consimilis*: 15,618 loci, 185 indv., 34 pop.).

We visualized population structure within each lineage (*E. n. nauseosa* and *E. n. consimilis*) using PCA. To evaluate the concordance of population genetic variation and geography, we applied Procrustes rotation of the first two PC axes onto population latitude and longitude coordinates (Wang et al., 2010). To further quantify the spatial scale of population differentiation, we used Random Forest models to assess the degree to which individuals could be classified according to their population of origin based on genotypic data alone. We conducted this analysis separately within *E. n. nauseosa* and *E. n. consimilis* and, again, for the full set of individuals. Sampling location was used as the response variable and principal component scores for each individual from PCAs of genotype probabilities were treated as predictors. The number of PC axes was selected using Tracy–Widom tests in R (R Core Team, 2020) package LEA (Gain & François, 2021) with an $\alpha \leq 0.001$ (overall = 59 PCs; *E. n. nauseosa* = 25 PCs, *E. n. consimilis* = 27 PCs). Models were analyzed using the R package RANDOM FOREST (Breiman, 2001), and accuracy was assessed using the out-of-bag error. As a metric of genetic differentiation among populations within and between each lineage, we estimated with multilocus F_{ST} calculated with the HIERFSTAT R package (Goudet & Jombart, 2020). We used Mantel tests (Mantel, 1967) to quantify isolation by distance (IBD; Wright, 1943) by associating genetic distances (Nei's D; Nei, 1972) with geographic distances within and between each lineage. Geographic distances among populations were calculated as haversine distance using the GEOSPHERE (Hijmans, 2017) R package, and the Procrustes rotation and Mantel tests were calculated using the VEGAN (Oksanen et al., 2019) R package.

Because variation in genetic diversity parameters could provide cursory information about the historical demography

of the subspecific lineages, we estimated a suite of population-level metrics. Specifically, we estimated nucleotide diversity (θ_π; Nei & Li, 1979), the Watterson estimator (θ_W; Watterson, 1975), and Tajima's *D* (Tajima, 1989) for each population (min. three individuals) using methods implemented in ANGSD v 0.923 that incorporate genotype uncertainty (Korneliussen et al., 2013, 2014). As estimates of genetic diversity and Tajima's *D* can be influenced by variation in sample size (Subramanian, 2016), we additionally conducted the same analyses after rarefying our data to have the same number of individuals for each population (*E. n. nauseosa*: $n = 6$; *E. n. consimilis*: $n = 5$). In addition, individual inbreeding coefficients (*F*) were calculated while incorporating genotype uncertainty using NGSF (Vieira et al., 2016), as implemented in the ANGSD-WRAPPER v 0.933 (Durvasula et al., 2016; see Supplementary methods).

Environmental and ecological influences of lineage differentiation and hybridization

In some locations, individuals of pure *E. n. nauseosa* and *E. n. consimilis* ancestry co-occurred, but hybrid ancestry was also widely dispersed across the sampled region, as revealed by both ENTROPY and PCA. Thus, to consider the extent to which extrinsic factors might shape reproductive isolation, or the lack thereof, we analyzed the extent to which environmental variation predicted the spatial distribution of lineage and hybrid ancestry. We characterized environmental variation across sampling sites with 30 climate variables of the Climatic Water Deficit Toolbox (Dilts et al., 2015). This dataset provides estimates of potential evapotranspiration, actual evapotranspiration, climatic water deficit, and 30-year PRISM normals (1981–2010) at 500 m × 500 m resolution and has been shown to predict aspects of spatial and distributional variation across plant communities (Lutz et al., 2010; Stephenson, 1998). Variables were reduced from 30 to 10 (based on Pearson's $|r| < 0.75$) to reduce the potential influence of multicollinearity. Full descriptions of these variables and their values for each sample location can be found in Supplementary Tables S3 and S4.

As both ecological and environmental factors can influence differentiation between lineages and reproductive isolation after secondary contact, we explored multiple facets of genotype-environment associations. To interpret environmental associations appropriately, the confounding correlation of geography and environment (Supplementary Figure S2) was taken into consideration throughout the following analyses. We first quantified the association between genetic and environmental distances, or isolation by environment (Wang & Bradburd, 2014) within and between each lineage using Mantel tests. We then estimated the magnitude and directionality of each environmental variable's influence on spatial genetic structure through redundancy analyses (RDAs). We conducted three additional analyses aiming to: (a) decouple the combined and independent influences of environment and geography on lineage divergence (i.e., ancestry coefficients) using a variance partitioning approach; (b) assess the predictability of lineage and hybrid assignment from environmental data using Random Forest; and (c) examine the association of individual environmental and geographic predictors with both *q* and lineage assignment of each population. Finally, as hybridization may occur more commonly across ecotones, we measured the average minimum geographic distance to the ecoregion boundary between the sampling locations within each lineage and hybrids based on the Environmental Protection Agency (EPA)-guided North American ecoregion level 2 (Omernik & Griffith, 2014; see Supplementary methods).

Assessing the evolutionary cohesiveness of varieties

Phylogenetic relationships were inferred for locations covering the two lineages and hybrids, as well as 14 of the named varieties. We included samples of *E. discoidea* for outgroup purposes. This analysis was used as an additional measure of divergence between the two lineages, and to evaluate the evolutionary cohesiveness of the putative varieties. For this analysis, we used field-designated subspecies and varieties and did not adjust them based on previous results. To ease computational investment and facilitate clear graphical visualization, we subsetted the data by randomly selecting up to six individuals from a reduced number of locations (231 individuals across 47 populations with even sampling [four to six individuals] across locations and subspecies). As the phylogenetic method used does not account for admixture, an additional tree was generated excluding hybrid individuals to complement inference. Variants were refiltered using a MAF ≥ 0.02 cutoff and resulting VCF files were converted into PHYLIP format using VCF2PHYLIP v 2.0 (Ortiz, 2019). We then reconstructed a maximum likelihood phylogenetic tree using IQ-TREE v 2.0.3 (Minh et al., 2020; Nguyen et al., 2015) with the PHYLIP file as input. We used the "Model Finder Plus" parameter (-m MFP) to determine the model that minimized the Bayesian information criterion (BIC) score, implemented the ascertainment bias correction method (Lewis, 2001), and applied the ultrafast bootstrap option with 1,000 bootstrap replicates (-bb 1000) to measure branch support. The resulting tree was visualized and annotated using GGTREE (Yu et al., 2017) in R. In addition, to assess the degree of range sympatry among the lineages and varieties, range maps from all sampled varieties were digitized from Anderson (1986b) with area and pairwise percent overlap calculated. Digitizing and georeferencing were performed manually in Inkscape v 0.92.5 (Inkscape project, 2020) with the shapefile generation, area calculations, and graphics were created using the R packages RASTER, RGEOS, SP, and SF (Bivand & Rundel, 2020; Hijmans, 2022; Pebesma, 2018; Pebesma & Bivand, 2005).

Results

Landscape genetic structure

After contaminant cleaning, mapping, variant calling, and filtering, we retained 22,917 loci in 586 individuals spanning 76 sampling locations (2–15 indv. per locale). There was no evidence of highly related individuals in our sampled locations (mean kinship coefficient = 0.039; only one pairing above the full-sibling estimate), consistent with outcrossing, and suggesting that subsequent population genetic analyses were not confounded by the presence of clones or familial relationships. Genetic differentiation between two lineages largely corresponding with subspecific designations (*E. n. nauseosa* and *E. n consimilis*) was evident in both the PCA and ancestry coefficients for the $k = 2$ ENTROPY model (Figure 1B, C). The first PC axis ($PVE = 8.94\%$) partitioned the two lineages and was highly correlated to ancestry ($|r| = 0.968$), while the second PC axis ($PVE = 5.39\%$) was associated with population structure within each lineage (i.e., geography; PC2—latitude: $|r| = 0.425$,

PC2—longitude: $|r| = 0.350$). Thus, unless specified otherwise, for subsequent analyses we categorized individuals based on ancestry coefficients as hybrid or belonging to one of the two lineages (*E. n. nauseosa*: $q \leq 0.08$; *E. n. consimilis*: $q \geq 0.87$; hybrid: $0.08 < q < 0.87$, 125 indv., 28 pop.). Despite substantial geographic overlap (Figure 1A), the two lineages exhibited clear genetic differentiation ($F_{ST} = 0.111$, 95% bootstrap CI = 0.107–0.114).

When the two lineages were analyzed separately, the population structure within each was strongly predicted by geographic variation and was often clear at fine spatial scales (<30 km, Figure 2). Individuals from the same populations generally formed tight clusters in PCA space and Random Forest analyses correctly categorized individuals according to their sampling origin at extremely high rates. Specifically, individual assignment accuracy was 94.37% with 69/76 sampling locations having at most one individual incorrectly classified for the full set of samples. Individual assignment accuracy within each lineage was similarly high (95.29% and 92.43%), with 26/27 (*E. n. nauseosa*) and 31/35 (*E. n. consimilis*) sampling locations having at most one individual incorrectly classified to location. Procrustes rotation correlations of geography onto the first two PCs were high, indicating that genetic structure was strongly predicted by geography (*E. n. nauseosa*: Procrustes $r = 0.761$; *E. n. consimlis*: Procrustes $r = 0.572$). In addition, Mantel tests of geographic distance and genetic distance (IBD) within lineages were strong (*E. n. nauseosa*: Mantel $r = 0.406$, $p = .001$; *E. n. consimilis*: Mantel $r = 0.239$, $p = .006$) but absent when IBD was assessed comparing populations spanning the two lineages (Mantel $r = -0.083$, $p = .790$; Figure 2C); consistent with geographically overlapping distributions of two differentiated lineages (see Supplementary Figure S2 for full associations). *Ericameria n. nauseosa* had the greatest genetic diversity and lowest inbreeding, on average, with the hybrids having intermediate values between *E. n. nauseosa* and *E. n. consimilis* (Supplementary Figure S3). All population estimates of Tajima's *D* were negative, but Tajima's *D* was associated

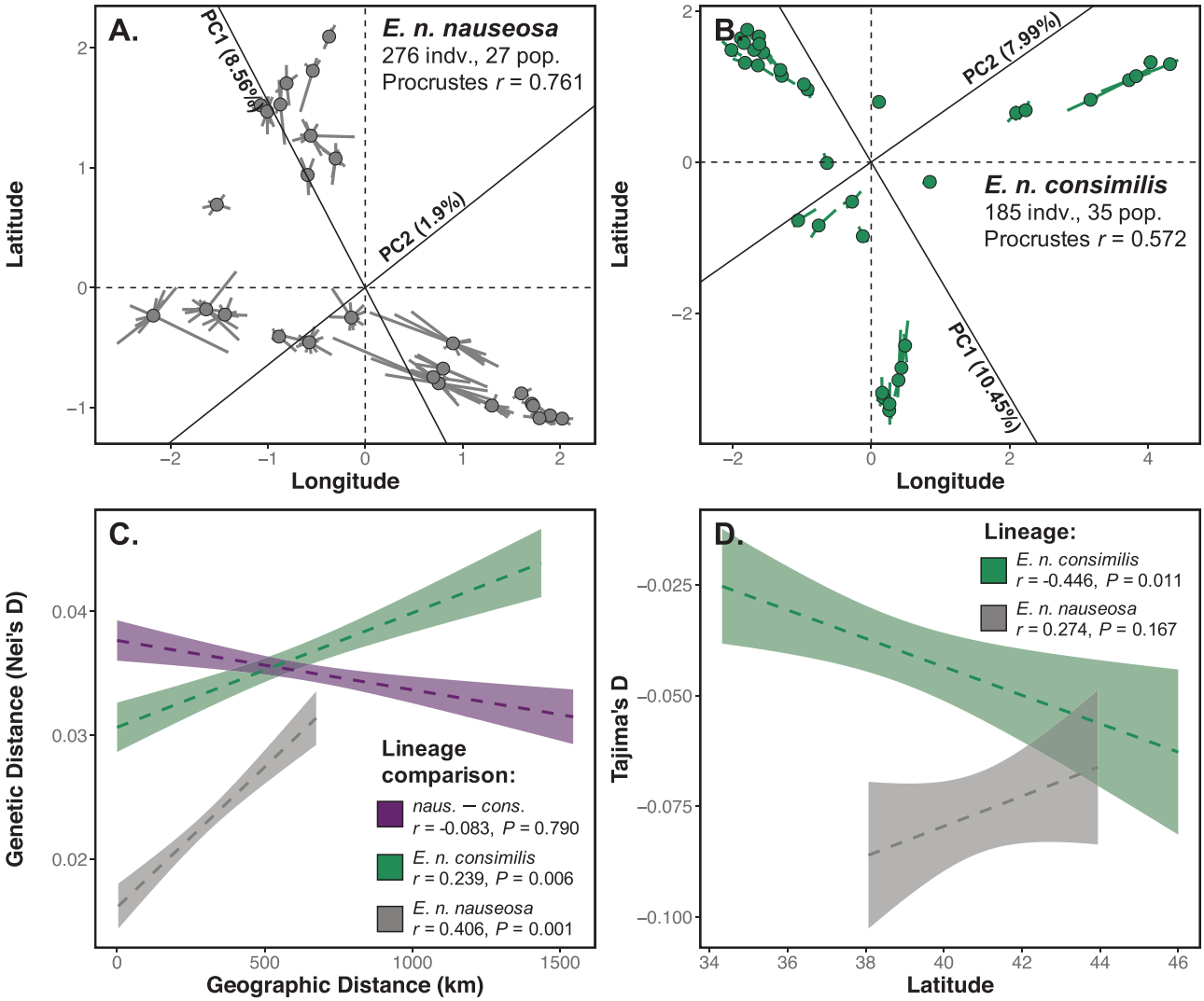

Figure 2. *E. n. nauseosa* and *E. n. consimilis* form independently evolving lineages, exhibiting strong population structure, patterns of isolation by distance (IBD), and variation in demographic history. Population genetic structure within both (A) *E. n. nauseosa* and (B) *E. n. consimilis* is demonstrated using a Procrustes rotation of the first two principal component (PC) axes onto latitude/longitude. Circles represent population means connected to lines representing individual PC scores from each population. (C) Strong patterns of IBD were present within the two lineages but not when comparing populations among and within both lineages (shown in purple). (D) Tajima's *D* is negative in both lineages and associated with latitude, but in opposite directions. These results potentially highlight differences in post-glaciation expansion routes between the two lineages.

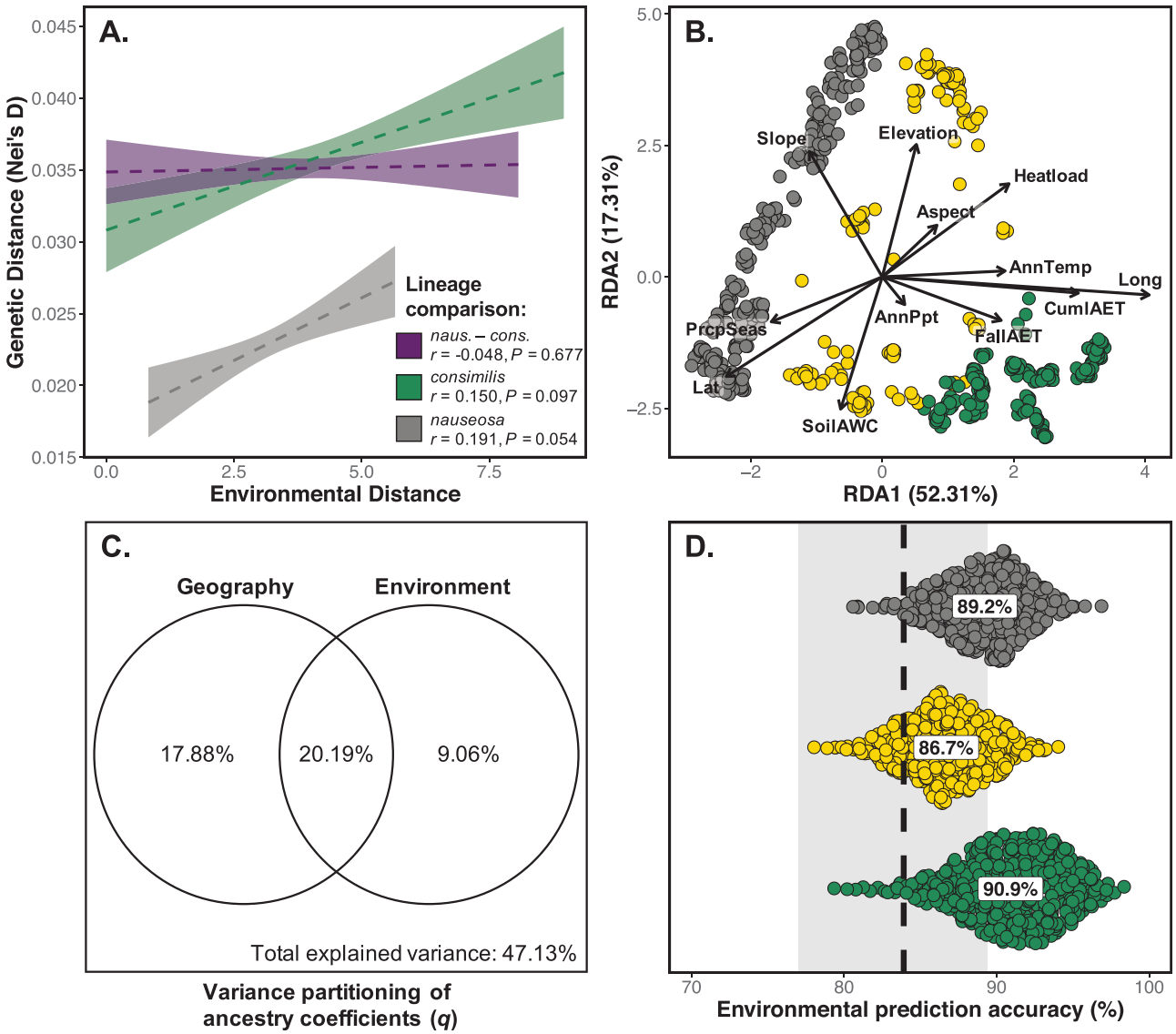

Figure 3. Environmental variation predicts differentiation and occurrence of hybridization between *E. n. nauseosa* and *E. n. consimilis* lineages. (A) Strong patterns of isolation by environment (IBE) were present within the two lineages but not when comparing populations among and within both lineages (shown in purple). (B) Environmental variation associated with genetic structure between the lineages illustrated by redundancy analysis (RDA). Axes were rotated in order to match that of the principal component analysis (PCA) in Figure 1 and environmental loadings were scaled (4.36×) for graphical representation. (C) Venn diagram illustrating partial variance of environment and geographic effects on ancestry coefficients in terms of adjusted r^2. (D) Environment alone predicts an individuals' classification to either parental lineage or hybrid using Random Forest. Points represent permuted accuracy within each lineage with means for each shown in text. Dashed line and shaded bar represent the overall mean predictive accuracy (83.9%) and 95% confidence interval.

with latitude in opposite directions in *E. n. nauseosa* and *E. n. consimilis* (Figure 2D). Tajima's *D* estimates from the rarefied dataset with equal sample sizes for all populations yielded very similar population-level estimates and the same geographic patterns (Pearson's $r = 0.932$).

Environmental and ecological influences of lineage differentiation and hybridization

Multiple analyses indicated the influence of both environment and geography on spatial patterns of genetic variation. Isolation by environment was weakly evident within each lineage (*E. n. nauseosa*: Mantel $r = 0.191$, $p = .054$; *E. n. consimilis*: Mantel $r = 0.150$, $p = .097$) but not when analyzed across both lineages (Mantel $r = -0.048$, $p = .677$; Figures 3A, Supplementary Figure S2), as expected given evidence for genetic divergence among lineages despite substantially overlapping distributions. Geographic and environmental variables explained genetic variation between the two lineages and were heavily loaded on the first RDA axis (*PVE* = 52.31%; Figure 3B). The geographic locality was unsurprisingly among the largest contributors to genetic differentiation between the lineages with increasing longitude (i.e., the more eastern part of the range) associated with the distribution of *E. n. consimilis* ancestry. Environmental variables also explained ancestry variation with increased cumulative and fall actual evapotranspiration (CumlAET and FallAET), heatload, and annual mean temperature (AnnTemp) associated with *E. n. consimilis* ancestry, and increased values of slope and precipitation seasonality (PrcpSeas) associated with *E. n. nauseosa* ancestry. Individuals with hybrid ancestry were associated with

intermediate values of these variables (Figure 3B). In addition, the second RDA axis splits the hybrids into two clusters with one associated with increasing soil water capacity (SoilAWC) and decreasing elevation and slope (Figure 3B).

In variance partitioning analyses, environment and geography explained a combined 47.1% of the total variance in q, but when taken individually, environment explained 29.3% and geography explained 38.1%. When surveying the partial contributions of environment and geography independent of correlative effects of each other, the combined partial effects contributed the most (adj. r^2 = 20.2%), followed by geography (adj. r^2 = 0.18), and then environment (adj. r^2 = 0.09; Figure 3C). Environmental variation was strongly predictive of ancestry in Random Forest models, which had an overall accuracy of 83.9% (range: 77.0–89.4%) in predicting an individual's lineage classification (mean r^2 = 0.860, RMSE = 0.160; Figure 4). Classification accuracies for *E. n. nauseosa* and *E. n. consimilis* were 89.2% and 90.9%, respectively, while the classification accuracy for hybrids was 86.7% (Figure 3D). In the Random Forest models, cumulative AET was overwhelmingly the strongest predictor of q, followed by annual precipitation, slope, and FallAET.

Univariate analyses additionally illustrated how environmental and geographic variables were associated with q and varied between the lineage categories (Supplementary Figures S4 and S5). Longitude was the most strongly variable associated with q, the latter of which exhibited a clear west-to-east pattern for *E. n. nauseosa* and *E. n. consimilis* (β = 0.274, r^2 = 0.327). Cumulative AET was also strongly associated with q and lineage distinction, increasing dramatically from *E. n. nauseosa* to *E. n. consimilis*, with hybrids being more intermediate (β = 0.015, r^2 = 0.162). Slope, AnnTemp, and FallAET were all additionally associated with lineage differentiation, albeit, to a much lesser degree.

Finally, the proximity of the sampling location to the nearest ecotone (designated by ecoregion level 2 boundaries; Omernik & Griffith, 2014) was predictive of hybrid ancestry (hybrid distance: mean = 39.65 km, median = 22.42 km; pure distance: mean = 64.10 km, median = 28.53 km; permuted p < .001; Figure 4), suggesting a role for ecotone transitions in hybridization. The same environmental association analyses applied to the subsetted dataset with equal sampling across locations and lineages illustrated similar patterns and parameter estimates (Supplementary Figure S6), confirming that our results were not confounded by uneven sampling.

Assessing the evolutionary cohesiveness of varieties

Phylogenetic reconstruction further illuminated divergence between the two lineages with patterns of substructure coinciding with some varietal designations (Figure 5B; see Supplementary Figure S7 for tree uninfluenced by admixture). High bootstrap support (>75%) characterized many clades within the tree, including divergence between *E. n. nauseosa* and *E. n. consimilis*, deeper divergences characterizing several southwestern edaphic varieties within *E. n. consimilis*, and monophyly of many populations and varieties. The varieties *bigelovii*, *graveolens*, *juncea*, and *turbinata* were each represented by samples from multiple locations that clustered together monophyletically with respect to each variety (Figure 5B). In contrast, individuals from other varieties appeared in multiple clusters throughout the tree and were represented across both lineages and hybrids (e.g., *oreophila*, *hololeuca*, *salicifolia*, and *speciosa*; Figure 5B). While individuals of hybrid ancestry with respect to *E. n. nauseosa* and *E. n. consimilis* are spread across many varieties, several varieties appear to be of entirely hybrid origin (*mohavensis*, *nana*, and *latisquamea*).

The digitization of the historical range maps from Anderson (1986b) revealed a large degree of sympatry across the lineages and varieties (Figures 1A and 5A), as well as variation in range size. While the error margins of manually georeferencing hand-drawn range maps to digital form should be considered, some inferences of range size and overlap can be useful for this highly complex species. The *E. nauseosa* range was estimated to be ~2.4 mil km^2, with a variety of

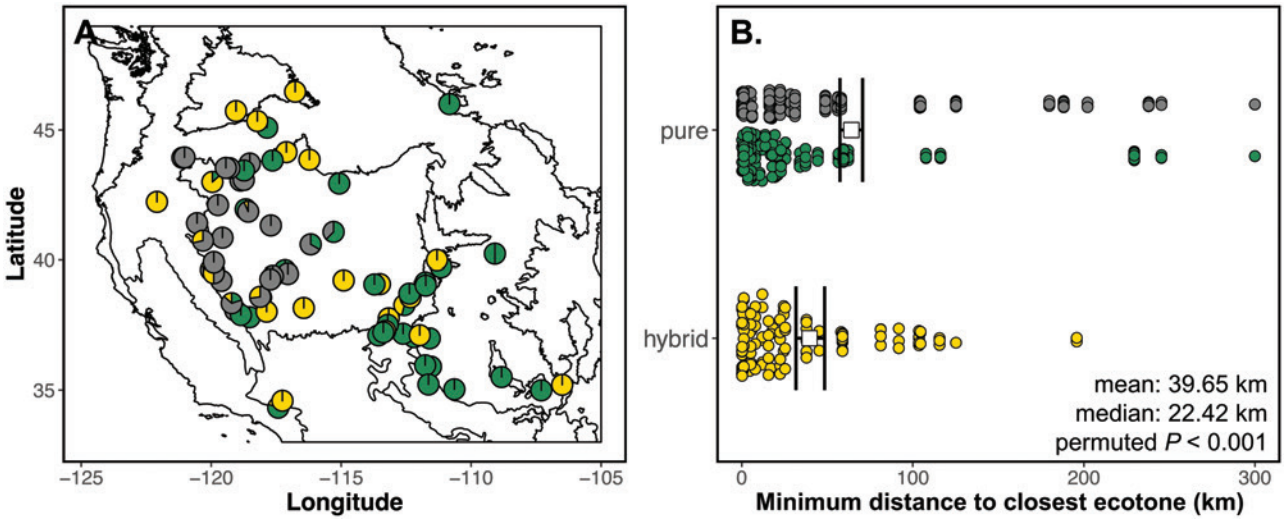

Figure 4. Proximity to ecotone predicts differentiation and hybridization between *E. n. nauseosa* and *E. n. consimilis* lineages. (A) Map displaying the proportion of individuals from each population with parental or hybrid ancestry. Solid lines indicate major ecotones between ecoregions (EPA ecoregion level 2). (B) Proximity to ecotone predicts likelihood of hybridization. A rarified, permutation test confirmed the estimated mean distance to ecotone for a hybrid was closer, on average, than an individual of pure ancestry. Colored points designate assignment to either *E. n. nauseosa, E. n. consimilis*, or hybrid.

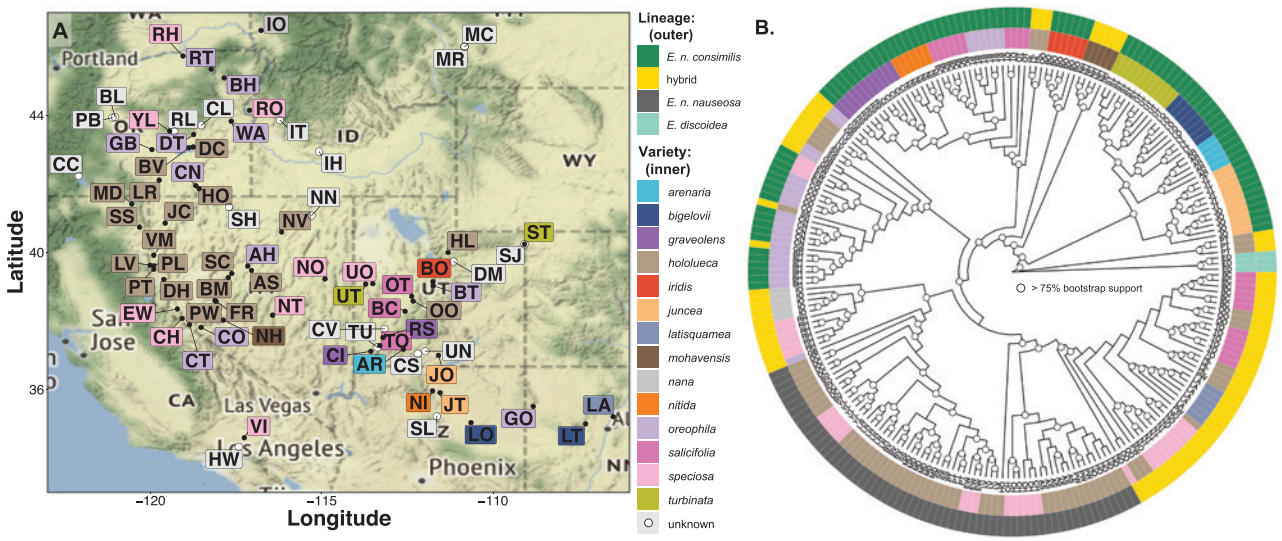

Figure 5. Varietal designation lacks clear evolutionary cohesiveness, likely due to range sympatry. (A) Map illustrating every sampling location within the study colored by the varietal designation. (B) Phylogenetic reconstruction for 231 individuals (47 populations, 6 ≤ individuals per population) including *E. discoidea* as the outgroup, built using IQ-TREE. The tip labels are the populations and circles at node represent > 75% bootstrap support for the node. The two-colored rings indicate ancestry assignment of lineage or hybrid (outer) and varietal designation (inner).

sizes ranging from just two single locations (*iridis*) to ~1 mil km^2 (*graveolens*; Supplementary Figure S8). There is also a substantial range overlap of the varieties, specifically concentrated among the edaphic varieties in UT/northern AZ and among varieties in southern CA (Supplementary Figure S9).

Genetic diversity estimates and inbreeding coefficients (F) varied considerably among the lineages and varieties, with the general pattern that the edaphic and narrow endemic varieties in the *E. n. consimilis* lineage had both decreased genetic diversity and increased inbreeding (avg. endemics: $F = 0.094$, $\theta_\pi = 0.0185$, $\theta_W = 0.0172$; avg. all others: mean $F = 0.017$, $\theta_\pi = 0.0196$, $\theta_W = 0.0202$; Supplementary Figure S10).

Discussion

Due to complex biogeographic histories and heterogeneous landscapes, numerous groups of plants in western North America consist of phenotypically and genetically diverged groups with variable reproductive isolation among populations and species with partial sympatry (e.g., Buck et al., 2023; Chhatre et al., 2018; Grant, 1981). Our results indicated that *E. nauseosa* has been strongly affected by the geographic and environmental variation in this region, resulting in two sympatric, independently evolving lineages that largely correspond with the *E. n. nauseosa* and *E. n. consimillis* subspecific designations. Within this distribution, environmental variation predicted the location of subspecific ancestry as well as the occurrence of numerous hybrid populations that occurred in environmentally intermediate locations relative to the two subspecies. Hybrid populations exhibited a continuum of admixed ancestry across a geographically dispersed mosaic hybrid zone, with hybridization more likely in populations closer to ecotone boundaries, which illustrated both replicated instances of hybridization and context-dependent variation in hybrid genetic variation. Within each subspecies, populations were surprisingly differentiated given the outcrossing mating system and the disturbance-oriented nature of this species. While many of the named varieties were phylogenetically dispersed and did not represent evolutionary cohesive groups, a group of southwestern endemic edaphic varieties represented genetically cohesive and differentiated groups on independent evolutionary trajectories. Our study begins to resolve complex patterns of phenotypic and genetic divergence, with ecologically based variation in hybridization, across the distribution of a phenotypically diverse foundational species.

Genetic differentiation between two subspecific lineages despite sympatric overlap

The changing taxonomy of the *E. nauseosa* complex suggests a complex evolutionary history, but historical classifications have all recognized the existence of multiple subspecies, even in sympatry. In support of these observations, the most pronounced feature of genetic variation we observed was a divergence between two lineages that largely coincided with the two recognized subspecies (*E. n. nauseosa* [green] and *E. n. consimillis* [grey]). The geography of divergence is notable—the differentiated lineages occur sympatrically over much of the sampled region and hybrids were found in many geographically distant locations (Figures 1A and 5A; Supplementary Figures S8 and S9; see below)—consistent with past descriptions of the subspecies having largely overlapping distributions (Anderson, 1986b; Supplementary Figure S7). The geography of differentiation documented here suggests the evolution of reproductive isolation between the subspecies. Consistent with this pattern, reports of the subspecies, and even pairs of specific varieties, occurring sympatrically with no sign of trait intergradation or hybridization have been common (Anderson, 1986b; Hanks et al., 1975; Hegerhorst et al., 1987b). In numerous cases of sympatry, the subspecies have been reported to have distinct phytochemical variation and unique communities of gall-forming insects (Baine et al., 2023; Floate et al., 1996; Hegerhorst et al., 1987b), suggesting extended ecological consequences of genome divergence among the subspecies.

Population differentiation within each lineage was surprisingly pronounced across spatial and environmental gradients, given outcrossing, insect pollination, and wind dispersal of

seed. When sample locations with pure ancestry were analyzed separately for each lineage, pronounced patterns of IBD and fine-scale population differentiation were evident (Figure 2; Supplementary Figure S2). Individuals were almost fully identifiable to sampling location based on genotypic variation alone (Figure 2A, B), as illustrated by Random Forest models which correctly inferred sample location from genotypic data with very little error (individual accuracy = 92.43–95.29%). This evidence for differentiation across small distances is consistent with recent analyses of similar data for a smaller set of *E. n. nauseosa* var. *hololeuca* populations (Faske et al., 2021) and suggests that insect pollination and wind-dispersed seed in *E. nauseosa* (which have sepals modified into a pappus, a dispersal mechanism common to many members of the Asteraceae), do not generate substantial gene flow that would otherwise limit local adaptation across environmental gradients (reviewed in Lenormand, 2002). Strong selection and local adaptation to diverse environments could be responsible for maintaining genetic differentiation across small spatial scales, and previous work detected congruent evidence for local adaptation to environmental variation through analyses of genetic environment associations and common garden phenotypic data (Faske et al., 2021). Despite *E. nauseosa* being widespread, strongly outcrossing, and disturbance-oriented, genetic differentiation across small spatial scales and local adaptation to diverse environments appear capable of influencing fine-scale diversification in this taxonomically diverse species complex.

Many plant groups of western North America have divergence histories associated with Quaternary glacial cycles, including expansions out of refugia and distributional shifts in response to changing climate (Dynesius & Jansson, 2000; Massatti & Knowles, 2020; Massatti et al., 2018; Shafer et al., 2010). Geographic variation of both nucleotide diversity and Tajima's *D* (Figure 2D; Supplementary Figure S3) suggest the potential for different histories in the two lineages which could have involved periods of allopatry followed by a progression of increased sympatry. Tajima's *D* was positively associated with latitude in *E. n. nauseosa* ($r = 0.274$), suggesting a southern expansion, and negatively associated with latitude in *E. n. consimilis* ($r = -0.494$), suggesting a northern expansion (Figure 2D). While negative Tajima's *D* estimates can arise from selective sweeps, it is unlikely to find this pattern across the whole genome; thus, the patterns here are more likely to reflect demographic changes (Stajich & Hahn, 2005) and could help identify the biogeographic history of these subspecies. One possible scenario is that *E. n. consimilis* expanded from the southern Rocky Mountain region, while *E. n. nauseosa* expanded south from a more northwestern distribution with sympatry occurring after periods of independent evolution. Differences in environmental variation associated with the subspecies (discussed below) may also reflect histories of local adaptation to different conditions. Although outside the scope of work presented here, future analyses spanning a range of demographic models based on more continuous sampling across the distribution would improve our understanding of this history.

Ecological predictors of hybridization between lineages

While sympatric distributions and the occurrence of individuals with pure *E. n. nauseosa* and *E. n. consimillis* ancestries within the same sampling locations suggest the evolution of reproductive isolation, the occurrence of admixed ancestry in numerous widely dispersed locations illustrates hybridization in a context-dependent pattern (i.e., a mosaic hybrid zone; Rand & Harrison, 1989). The general pattern of widely dispersed locations with admixed ancestry across the overlapping distributions of *E. n. nauseosa* and *E. n. consimillis* (Figures 1A and 4A) is consistent with expectations for a mosaic hybrid zone, where reproductive isolation is predicted by ecological and environmental variation rather than a geographic cline of secondary contact (Harrison & Larson, 2016; Rand & Harrison, 1989).

Evidence for environmental variation predicting ancestry (Figure 3A, B) suggests that local adaptation across environmental gradients is likely to mediate reproductive isolation between *E. n. nauseosa* and *E. n. consimilis* or that hybrid fitness varies with the environment. The environmental variables most strongly contributing to these results (e.g., CumlAET, FallAET, Slope, and PrcpSeas; Figure 3B) are not surprising given similar variables were implicated in the recent inference of local adaptation across *E. n. nauseosa* of the western Great Basin (Faske et al., 2021), as well as in other plant taxa of western North America (Massatti & Knowles, 2020; Shryock et al., 2017). Past observations have noted that the timing of flowering and seed set varies across the range due to environmental conditions (Meyer et al., 1989), and it has been anecdotally observed that these traits can differ between *E. n. nauseosa* and *E. n. consimillis* when in sympatry, suggesting that local adaptation may also underlie prezygotic isolation.

Admixed ancestries were also more likely to occur near ecological transitions than non-admixed ancestries (Figure 4). This is not surprising, as environment and climate transition across the North American ecoregion level 2 boundaries on which we based analyses (Omernik & Griffith, 2014), but suggests that ecotones could additionally predict the distribution of hybrids. Many of the ecoregion 2 boundaries cover biogeographic breaks that have been described as suture zones, where numerous taxa hybridize at overlapping secondary contact zones (Remington, 1968; Swenson & Howard, 2004, 2005). Our results suggest that the multiple suture zones spanned by the large, biogeographically complex distribution of *E. nauseosa* contribute to the folding of ancestry across intermediate environments.

The decoupling of environment and geography is a feature of our analyses (Supplementary Figure S2) that is shared with other studies of mosaic hybrid zones (e.g., *Bombina*, Vines et al., 2003; *Populus*, Lindtke et al., 2013; *Gryllus*, Ross & Harrison, 2002; *Mytilus*, Bierne et al., 2003). Such systems are valuable for identifying the mechanisms that underlie the repeated and independent formation of hybrids, as they often illustrate context dependence and stochasticity in admixture among populations and across genomes (Mandeville et al., 2017; Teeter et al., 2010). Given our ability to predict the location of admixed populations based on both environmental variation (Figure 3D) and proximity to ecological transitions (Figure 4), one might expect a similar genomic composition for hybrids if the environment serves as a strong filter for hybrids relative to non-admixed individuals. We found a contrasting pattern in which hybrids were found throughout RDA space, indicating substantial variation in hybrid genetic variation and its environmental associations (Figure 3B). This stochasticity points to the challenge of predicting hybridization outcomes, even when there is some understanding of how environment and hybrid fitness vary (e.g., Lexer et al., 2004; Mandeville et al., 2017; McFarlane et al., 2022). As hybrids can have phenotypic variation distinct from either

parent species (Kagawa & Takimoto, 2018; Rieseberg et al., 1999, 2007), this could also suggest that hybridization may generate genomic variation that can effectively explore the environmental fitness landscapes occupied and unoccupied by the parental lineages.

The evolutionary cohesiveness of varieties

Reconstruction of the evolutionary relationships among samples spanning different named varieties revealed a spectrum of outcomes, ranging from monophyly, consistent with a variety being an evolutionarily cohesive lineage, to polyphyly, where individuals from the same putative variety were found throughout the phylogeny. Several varieties, including *speciosa*, *salicifolia*, *oreophila,* and *hololeuca*, showed instances of polyphyly (Figures 1A and 5A). Varietal designation in such cases could be a result of phenotypic similarity of distantly related populations, which could arise through phenotypic plasticity and/or convergent evolution to similar ecological conditions. Both processes can challenge taxonomy and have been widely documented in plant taxa of western North America (e.g., Love et al., 2023; Wilson et al., 2007). Furthermore, difficulties in identifying field collected *E. nauseosa* to variety could obfuscate our analyses and challenge our ability to evaluate the evolutionary cohesiveness of some varieties. For example, Anderson (1986b) suggested that the *hololeucus* and *oreophila* varieties maintained phenotypic differentiation in sympatry to the extent they could be recognized as different species, which contrasts with the polyphyly of these varieties observed here. However, *hololeuca* and *speciosa* both had polyphyletic populations with admixed subspecific ancestry, suggesting the mosaic distribution of hybridization described above could also have influenced these analyses. Nonetheless, the numerous occurrences of polyphyly reported here suggest much future work will be needed to resolve the varietal-level taxonomy of *E. nauseosa*.

In contrast, some groups illustrated genetic variation consistent with varietal designations. For instance, varieties *bigelovii*, *graveolens*, and *turbinata* all formed cohesive, monophyletic groupings of samples spanning multiple locations. The other seemingly monophyletic variety from multiple locations is *juncea*, which along with *arenaria*, *bigelovii*, and *turbinata* formed a distinct, substructured group within the *E. n. consimillis* lineage. These varieties, many of which had reduced genetic diversity (Supplementary Figure S10), are largely thought to be edaphic endemics (Cronquist, 1994) and are primarily found within the southeastern-most part of the range (Utah and northern Arizona). This region coincidentally sees the most amount of varietal overlap (Figure 5A; Supplementary Figures S8 and S9) and contains a large portion of the sampling locations with admixed ancestry with respect to subspecies. Several endemic varieties that formed monophyletic groups (*mohavensis*, *nana*, and *latisquamea*) consisted entirely of admixed ancestry from the two subspecies (Figures 1 and 5). As hybrids often have distinct, or transgressive, phenotypic variation with respect to parent species (Kagawa & Takimoto, 2018; Rieseberg et al., 1999, 2007), this result suggests that hybridization could facilitate the colonization of, and local adaptation to, novel environments in these narrow endemic and phenotypically distinct varieties. While this result hints at a role for admixture in varietal diversification, our sampling of individuals and locations is limited for these varieties and future work with more comprehensive sampling will support a more thorough understanding of the potential for hybridization to provide a creative role in the evolution of morphological novelty.

Conclusions

Our analyses illustrate a continuum of divergence across lineages and environments occupied by *E. nauseosa*, a foundational shrub species of western North America known for its marked phenotypic variation. Our analyses indicate that *E. nauseosa* comprises two genetically distinct lineages that exhibit environmentally dependent reproductive isolation and maintain differentiation despite widely overlapping distributions. We found admixed ancestries in a subset of locations that were widely distributed in a mosaic hybrid zone where ancestry was strongly predicted by both environmental variation and the proximity to ecotonal transitions. While most studies of hybridization focus on one instance of secondary contact, those that have looked at multiple zones of contact have found mixed evidence of predictability in the occurrence and outcome of hybridization. Our results suggest that environmental variation is a key factor in mediating reproductive isolation in these young lineages, but also illustrate considerable variation in the association of genetic and environmental variation across multiple instances of hybridization. This is the first geographically extensive survey of genetic variation in *E. nauseosa*, yet we lacked sampling in the southeastern and northeastern portions of the distribution, and our sampling of putative-named varieties was limited in scope and space. Inference of evolutionary history for many of the named varieties we sampled may be confounded by the potential for misidentification of field-collected samples as well as our sampling of only one or two localities across small geographic distances in a species with substantial population structure. Future analyses with a more comprehensive sampling of the range, including the sampling of many locations within specific named varieties, stand to further shape our understanding of how environment, history, and admixture underlie the extensive phenotypic variation the group is noted for.

Supplementary material

Supplementary material is available online at *Evolution*.

Data availability

The data that supports the finding of this study are available from the Dryad Digital Repository: doi:10.5061/dryad.5hqbzkhbg. This includes population-level summaries of environmental data, information of sampling locality, filtered VCF file, and a matrix of genotype probabilities. All scripts from pre-processing, variant calling, to statistical analyses and data visualization are available on either Dyrad or GitHub: https://github.com/trevorfaske/ERNA_analyses.

Author contributions

E.A.L., B.A.R., and T.L.P. developed the original idea for this research and secured funding. A.C.A. and B.A.R. (with associated field crews) collected samples and identified them to subspecies/variety. L.M.S. extracted the DNA and prepped the libraries for Illumina sequencing. T.M.F. analyzed the dataset with C.O.M. conducting the phylogenetic analyses. T.M.F., J.P.J., and T.L.P. wrote the initial draft and all authors contributed to revising the manuscript.

Funding

This project was funded by grant 2017-67019-26336 from the USDA National Institute of Food and Agriculture Sustainable Agroecosystems program, Cooperative Agreements L16AC00318 and L19AC00013 from the Bureau of Land Management through the Great Basin CESU to E.A.L. and T.L.P., and National Fish and Wildlife Foundation program funding L17PG00234 through the Bureau of Land Management, Colorado Plateau Native Plant Program to B.A.R. J.P.J. was supported by the Modelscape Consortium with funding from NSF (OIA-2019528). Illumina sequencing was performed by the Genomic Sequencing and Analysis Facility at UT Austin, Center for Biomedical Research Support (RRID#: SCR_021713).

Conflict of interest: None declared.

Acknowledgments

We would like to thank Tanner Tobiasson for assistance with field collections. We thank Eliza Grames and Vince Martinson for their assistance in digitizing and georeferencing of range maps. We would also like to thank Josh Hallas, Chris Halsch, Josh Harrison, Lanie Galland, and Katie Uckele for conceptual support and Kelly Loria for the graphical illustrations. We thank Eryn McFarlane for helpful comments on an early version of the manuscript. The findings and conclusions in this publication are those of the authors and should not be construed to represent any official USDA or U.S. Government determination or policy. Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

References

Anderson, L. C. (1973). Unique *Chrysothamnus* hybridizations in Ash Meadows, Nevada. *Bulletin of the Torrey Botanical Club*, *100*(3), 171. https://doi.org/10.2307/2484629

Anderson, L. C. (1986a). An overview of the genus *Chrysothamnus* (Asteraceae). In E. D. Mcarthur, & B. L. Welsh (Eds.), *Proceedings, Symposium on the Biology of Artemisia and Chrysothamnus* (pp. 29–35). USDA, Forest Service, Intermountain Research Station.

Anderson, L. C. (1986b). Sympatric subspecies in *Chrysothamnus nauseosa*. In E. D. Mcarthur, & B. L. Welsh (Eds.), *Proceedings, Symposium on the Biology of Artemisia and Chrysothamnus* (pp. 98–103). USDA, Forest Service.

Anderson, L. C., & Reveal, J. L. (1966). *Chrysothamnus bolanderi*, an intergeneric hybrid. *Madroño*, *18*(8), 225–233.

Baine, Q., Casares, E. E., Carabotta, E., Martinson, V. G., & Martinson, E. O. (2023). Galls on galls: A hypergall-inducing midge and its parasitoid community. *Ecology*, *104*(5), e4018. https://doi.org/10.1002/ecy.4018

Barton, N. H. (2001). The role of hybridization in evolution. *Molecular Ecology*, *10*(3), 551–568. https://doi.org/10.1046/j.1365-294x.2001.01216.x

Barton, N. H., & Hewitt, G. M. (1985). Analysis of hybrid zones. *Annual Review of Ecology and Systematics*, *16*(1), 113–148. https://doi.org/10.1146/annurev.es.16.110185.000553

Baughman, O. W., Agneray, A. C., Forister, M. L., Kilkenny, F. F., Espeland, E. K., Fiegener, R., Horning, M. E., Johnson, R. C., Kaye, T. N., Ott, J., St. Clair, J. B., & Leger, E. A. (2019). Strong patterns of intraspecific variation and local adaptation in Great Basin plants revealed through a review of 75 years of experiments. *Ecology and Evolution*, *9*(11), 6259–6275. https://doi.org/10.1002/ece3.5200

Bierne, N., Borsa, P., Daguin, C., Jollivet, D., Viard, F., Bonhomme, F., & David, P. (2003). Introgression patterns in the mosaic hybrid zone between *Mytilus edulis* and *M galloprovincialis*. *Molecular Ecology*, *12*(2), 447–461. https://doi.org/10.1046/j.1365-294x.2003.01730.x

Bivand, R., & Rundel, C. (2020). *rgeos: Interface to Geometry Engine—Open Source ("GEOS"). R package version 0.5-5*. https://CRAN.R-project.org/package=rgeos

Bolnick, D. I., Hund, A. K., Nosil, P., Peng, F., Ravinet, M., Stankowski, S., Subramanian, S., Wolf, J. B. W., & Yukilevich, R. (2023). A multivariate view of the speciation continuum. *Evolution*, *77*(1), 318–328. https://doi.org/10.1093/evolut/qpac004

Breiman, L. (2001). Random forests. *Machine Learning*, *45*, 5–32. https://doi.org/10.1023/A:1010933404324

Buck, R., Ortega-Del Vecchyo, D., Gehring, C., Michelson, R., Flores-Rentería, D., Klein, B., Whipple, A. V., & Flores-Rentería, L. (2023). Sequential hybridization may have facilitated ecological transitions in the Southwestern pinyon pine syngameon. *New Phytologist*, *237*(6), 2435–2449. https://doi.org/10.1111/nph.18543

Calfee, E., Agra, M. N., Palacio, M. A., Ramírez, S. R., & Coop, G. (2020). Selection and hybridization shaped the rapid spread of African honey bee ancestry in the Americas. *PLoS Genetics*, *16*(10), e1009038. https://doi.org/10.1371/journal.pgen.1009038

Carril, O. M., Griswold, T., Haefner, J., & Wilson, J. S. (2018). Wild bees of grand staircase-escalante national monument: Richness, abundance, and spatio-temporal beta-diversity. *PeerJ*, *6*(11), e5867–e5829. https://doi.org/10.7717/peerj.5867

Chhatre, V. E., Evans, L. M., DiFazio, S. P., & Keller, S. R. (2018). Adaptive introgression and maintenance of a trispecies hybrid complex in range-edge populations of *Populus*. *Molecular Ecology*, *27*(23), 4820–4838. https://doi.org/10.1111/mec.14820

Cronquist, A. (1994). *Intermountain flora, vascular plants of the Intermountain West, U.S.A* (Vol. *5*). New York Botanical Garden.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., & Durbin, R.; 1000 Genomes Project Analysis Group. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

Dilts, T. E., Weisberg, P. J., Dencker, C. M., & Chambers, J. C. (2015). Functionally relevant climate variables for arid lands: A climatic water deficit approach for modelling desert shrub distributions. *Journal of Biogeography*, *42*(10), 1986–1997. https://doi.org/10.1111/jbi.12561

Donovan, L. A., & Ehleringer, J. R. (1994). Water stress and use of summer precipitation in a great basin shrub community. *Functional Ecology*, *8*(3), 289–297. https://doi.org/10.2307/2389821

Durvasula, A., Hoffman, P. J., Kent, T. V., Liu, C., Kono, T. J. Y., Morrell, P. L., & Ross-Ibarra, J. (2016). Angsd-Wrapper: Utilities for analysing next-generation sequencing data. *Molecular Ecology Resources*, *16*(6), 1449–1454. https://doi.org/10.1111/1755-0998.12578

Dynesius, M., & Jansson, R. (2000). Evolutionary consequences of changes in species' geographical distributions driven by Milankovitch climate oscillations. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(16), 9115–9120. https://doi.org/10.1073/pnas.97.16.9115

Endler, J. A. (1977). *Geographic variation, speciation, and clines* (10th ed.). Princeton University Press.

Faske, T. M., Agneray, A. C., Jahner, J. P., Sheta, L. M., Leger, E. A., & Parchman, T. L. (2021). Genomic and common garden approaches yield complementary results for quantifying environmental drivers of local adaptation in rubber rabbitbrush, a foundational Great Basin shrub. *Evolutionary Applications*, *14*(12), 2881–2900. https://doi.org/10.1111/eva.13323

Fernandes, G. W., Saraivia, C., Cornelissen, T. G., & Price, P. W. (2000). Diversity and morphology of insect in Northern galls on *Chrysothamnus nauseosus* (Asteraceae) in Arizona. *BIOS*, *8*(8), 39–48.

Floate, K. D., Fernandes, G. W., & Nilsson, J. A. (1996). Distinguishing intrapopulational categories of plants by their insect faunas: Galls on rabbitbrush. *Oecologia*, *105*(2), 221–229. https://doi.org/10.1007/BF00328550

Forester, B. R., Jones, M. R., Joost, S., Landguth, E. L., & Lasky, J. R. (2016). Detecting spatial genetic signatures of local adaptation

in heterogeneous landscapes. *Molecular Ecology*, *25*(1), 104–120. https://doi.org/10.1111/mec.13476

Forester, B. R., Lasky, J. R., Wagner, H. H., & Urban, D. L. (2018). Comparing methods for detecting multilocus adaptation with multivariate genotype–environment associations. *Molecular Ecology*, *27*(9), 2215–2233. https://doi.org/10.1111/mec.14584

Fu, L., Niu, B., Zhu, Z., Wu, S., & Li, W. (2012). CD-HIT: Accelerated for clustering the next-generation sequencing data. *Bioinformatics*, *28*(23), 3150–3152. https://doi.org/10.1093/bioinformatics/bts565

Fu, R., Zhu, Y., Liu, Y., Feng, Y., Lu, R., Li, Y., Li, P., Kremer, A., Lascoux, M., & Chen, J. (2022). Genome-wide analyses of introgression between two sympatric Asian oak species. *Nature Ecology & Evolution*, *6*(7), 924–935. https://doi.org/10.1038/s41559-022-01754-7

Gain, C., & François, O. (2021). LEA 3: Factor models in population genetics and ecological genomics with R. *Molecular Ecology Resources*, *21*(8), 2738–2748. https://doi.org/10.1111/1755-0998.13366

Gang, D. R., & Weber, D. J. (1995). Genetic variability and relationships among ten populations of rubber rabbitbrush (*Chrysothamnus nauseosus* ssp. *hololeucus*) determined by RAPD analysis of bulked genomic DNA samples. *Botanical Bulletin of Academic Sinica*, *36*, 1–8.

Giska, I., Farelo, L., Pimenta, J., Seixas, F. A., Ferreira, M. S., Marques, J. P., Miranda, I., Letty, J., Jenny, H., Hackländer, K., Magnussen, E., & Melo-Ferreira, J. (2019). Introgression drives repeated evolution of winter coat color polymorphism in hares. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(48), 24150–24156. https://doi.org/10.1073/pnas.1910471116

Gompert, Z., Fordyce, J. A., Forister, M. L., Shapiro, A. M., & Nice, C. C. (2006). Homoploid hybrid speciation in an extreme habitat. *Science*, *314*(5807), 1923–1925. https://doi.org/10.1126/science.1135875

Gompert, Z., Lucas, L. K., Buerkle, C. A., Forister, M. L., Fordyce, J. A., & Nice, C. C. (2014). Admixture and the organization of genetic diversity in a butterfly species complex revealed through common and rare genetic variants. *Molecular Ecology*, *23*(18), 4555–4573. https://doi.org/10.1111/mec.12811

Gompert, Z., Mandeville, E. G., & Buerkle, C. A. (2017). Analysis of population genomic data from hybrid zones. *Annual Review of Ecology, Evolution, and Systematics*, *48*(1), 207–229. https://doi.org/10.1146/annurev-ecolsys-110316-022652

Goudet, J., & Jombart, T. (2020). *hierfstat: Estimation and tests of hierarchical F-statistics*. https://cran.r-project.org/package=hierfstat

Grady, K. C., Ferrier, S. M., Kolb, T. E., Hart, S. C., Allan, G. J., & Whitham, T. G. (2011). Genetic variation in productivity of foundation riparian species at the edge of their distribution: Implications for restoration and assisted migration in a warming climate. *Global Change Biology*, *17*(12), 3724–3735. https://doi.org/10.1111/j.1365-2486.2011.02524.x

Grant, V. (1981). *Plant speciation*. Columbia University Press. https://doi.org/10.7312/gran92318

Hall, M. H. (1919). *Chrysothamnus nauseosus* and its varieties. *University of California Publications in Botany*, *7*, 159–181.

Hanks, D. L., McArthur, E. D., Plummer, A. P., Giunta, B. C., & Blauer, A. C. (1975). Chromatographic recognition of some palatable and unpalatable subspecies of rubber Rabbitbrush in and around Utah. *Journal of Range Management*, *28*(2), 144. https://doi.org/10.2307/3897448

Hansen, J. D. (1986). Comparison of insects from burned and unburned areas after a range fire. *Great Basin Naturalist*, *46*(4), 721–727.

Hapke, A., & Thiele, D. (2016). GIbPSs: A toolkit for fast and accurate analyses of genotyping-by-sequencing data without a reference genome. *Molecular Ecology Resources*, *16*(4), 979–990. https://doi.org/10.1111/1755-0998.12510

Harrison, R. G., & Larson, E. L. (2016). Heterogeneous genome divergence, differential introgression, and the origin and structure of hybrid zones. *Molecular Ecology*, *25*(11), 2454–2466. https://doi.org/10.1111/mec.13582

Hegerhorst, D., Weber, D. W., & McArthur, E. D. (1987a). Resin and rubber content in *Chrysothamnus*. *The Southwestern Naturalist*, *32*(4), 475–482. https://doi.org/10.2307/3671481

Hegerhorst, D., Weber, D. J., McArthur, E. D., & Khan, A. J. (1987b). Chemical analysis and comparison of subspecies of *Chrysothamnus nauseosus* and other related species. *Biochemical Systematics and Ecology*, *15*(2), 201–208. https://doi.org/10.1016/0305-1978(87)90020-2

Hijmans, R. J. (2017). *geosphere: Spherical trigonometry*. https://cran.r-project.org/package=geosphere

Hijmans, R. J. (2022). *raster: Geographic data analysis and modeling. R package version 3.5-15*. https://CRAN.R-project.org/package=raster

Hodel, R. G. J., Massatti, R., & Knowles, L. L. (2022). Hybrid enrichment of adaptive variation revealed by genotype–environment associations in montane sedges. *Molecular Ecology*, *31*(13), 3722–3737. https://doi.org/10.1111/mec.16502

Hohenlohe, P. A., Day, M. D., Amish, S. J., Miller, M. R., Kamps-Hughes, N., Boyer, M. C., Muhlfeld, C. C., Allendorf, F. W., Johnson, E. A., & Luikart, G. (2013). Genomic patterns of introgression in rainbow and westslope cutthroat trout illuminated by overlapping paired-end RAD sequencing. *Molecular Ecology*, *22*(11), 3002–3013. https://doi.org/10.1111/mec.12239

Hood, G. R., Zhang, L., Hu, E. G., Ott, J. R., & Egan, S. P. (2019). Cascading reproductive isolation: Plant phenology drives temporal isolation among populations of a host-specific herbivore. *Evolution*, *73*(3), 554–568. https://doi.org/10.1111/evo.13683

Inkscape Project. (2020). *Inkscape*. https://inkscape.org

Kagawa, K., & Takimoto, G. (2018). Hybridization can promote adaptive radiation by means of transgressive segregation. *Ecology Letters*, *21*(2), 264–274. https://doi.org/10.1111/ele.12891

Korneliussen, T. S., Albrechtsen, A., & Nielsen, R. (2014). ANGSD: Analysis of next generation sequencing data. *BMC Bioinformatics*, *15*(1), 1–13. https://doi.org/10.1186/s12859-014-0356-4

Korneliussen, T. S., Moltke, I., Albrechtsen, A., & Nielsen, R. (2013). Calculation of Tajima's D and other neutrality test statistics from low depth next-generation sequencing data. *BMC Bioinformatics*, *14*(1), 289. https://doi.org/10.1186/1471-2105-14-289

Lenormand, T. (2002). Gene flow and the limits to natural selection. *Trends in Ecology & Evolution*, *17*(4), 183–189. https://doi.org/10.1016/s0169-5347(02)02497-7

Lewis, P. O. (2001). A likelihood approach to estimating phylogeny from discrete morphological character data. *Systematic Biology*, *50*(6), 913–925. https://doi.org/10.1080/106351501753462876

Lexer, C., Lai, Z., & Rieseberg, L. H. (2004). Candidate gene polymorphisms associated with salt tolerance in wild sunflower hybrids: Implications for the origin of *Helianthus paradoxus*, a diploid hybrid species. *The New Phytologist*, *161*(1), 225–233. https://doi.org/10.1046/j.1469-8137.2003.00925.x

Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM, arXiv. http://arxiv.org/abs/1303.3997, preprint: not peer reviewed.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., & Durbin, R.; 1000 Genome Project Data Processing Subgroup. (2009). The sequence Alignment/Map format and SAMtools. *Bioinformatics*, *25*(16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Li, W., & Godzik, A. (2006). Cd-hit: A fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics*, *22*(13), 1658–1659. https://doi.org/10.1093/bioinformatics/btl158

Lindtke, D., González-Martínez, S. C., Macaya-Sanz, D., & Lexer, C. (2013). Admixture mapping of quantitative traits in *Populus* hybrid zones: Power and limitations. *Heredity*, *111*(6), 474–485. https://doi.org/10.1038/hdy.2013.69

Lotsy, J. P. (1925). Species or linneon? *Genetica*, *7*(5-6), 487–506. https://doi.org/10.1007/bf01676287

Love, S. J., Schweitzer, J. A., & Bailey, J. K. (2023). Climate-driven convergent evolution in riparian ecosystems on sky islands. *Scientific Reports*, *13*(1), 1–9. https://doi.org/10.1038/s41598-023-29564-2

Lutz, J. A., van Wagtendonk, J. W., & Franklin, J. F. (2010). Climatic water deficit, tree species ranges, and climate change in Yosemite National Park. *Journal of Biogeography*, *37*(5), 936–950. https://doi.org/10.1111/j.1365-2699.2009.02268.x

Maheshwari, S., & Barbash, D. A. (2011). The genetics of hybrid incompatibilities. *Annual Review of Genetics*, *45*, 331–355. https://doi.org/10.1146/annurev-genet-110410-132514

Mandeville, E. G., Parchman, T. L., Thompson, K. G., Compton, R. I., Gelwicks, K. R., Song, S. J., & Buerkle, C. A. (2017). Inconsistent reproductive isolation revealed by interactions between *Catostomus* fish species. *Evolution Letters*, *1*(5), 255–268. https://doi.org/10.1002/evl3.29

Manichaikul, A., Mychaleckyj, J. C., Rich, S. S., Daly, K., Sale, M., & Chen, W. M. (2010). Robust relationship inference in genome-wide association studies. *Bioinformatics*, *26*(22), 2867–2873. https://doi.org/10.1093/bioinformatics/btq559

Mantel, N. (1967). The detection of disease clustering and a generalized regression approach. *Cancer Research*, *27*(2), 209–220. https://doi.org/10.1038/240498a0

Massatti, R., & Knowles, L. L. (2020). The historical context of contemporary climatic adaptation: A case study in the climatically dynamic and environmentally complex southwestern United States. *Ecography*, *43*(5), 735–746. https://doi.org/10.1111/ecog.04840

Massatti, R., Prendeville, H. R., Larson, S., Richardson, B. A., Waldron, B., & Kilkenny, F. F. (2018). Population history provides foundational knowledge for utilizing and developing native plant restoration materials. *Evolutionary Applications*, *11*(10), 2025–2039. https://doi.org/10.1111/eva.12704

Mcarthur, E. D., Tierman, C. F., & Welch, B. J. (1986). Specificity of galls on *Chrysothamnus nauseosus* subspecies. *Great Basin Naturalist*, *39*(1), 81–87.

McArthur, E. D., Blauer, A. C., Plummer, A. P., & Stevens, R. (1979). *Characteristics and hybridization of important intermountain shrubs III. Sunflower family. Research Paper INT-220*. USDA Forest Service.

Mcfarlane, S. E., Jahner, J. P., Lindtke, D., Buerkle, C. A., & Mandeville, E. G. (2022). Selection leads to remarkable variability in the outcomes of hybridization across replicate hybrid zones. bioRxiv 509250. https://www.biorxiv.org/content/10.1101/2022.09.23.509250v2.full, preprint: not peer reviewed.

McKinney, G. J., Waples, R. K., Pascal, C. E., Seeb, L. W., & Seeb, J. E. (2018). Resolving allele dosage in duplicated loci using genotyping-by-sequencing data: A path forward for population genetic analysis. *Molecular Ecology Resources*, *18*(3), 570–579. https://doi.org/10.1111/1755-0998.12763

McKinney, G. J., Waples, R. K., Seeb, L. W., & Seeb, J. E. (2017). Paralogs are revealed by proportion of heterozygotes and deviations in read ratios in genotyping-by-sequencing data from natural populations. *Molecular Ecology Resources*, *17*(4), 656–669. https://doi.org/10.1111/1755-0998.12613

McVay, J. D., Flores-Villela, O., & Carstens, B. (2015). Diversification of North American natricine snakes. *Biological Journal of the Linnean Society*, *116*(1), 1–12. https://doi.org/10.1111/bij.12558

McVean, G. (2009). A genealogical interpretation of principal components analysis. *PLoS Genetics*, *5*(10), e1000686. https://doi.org/10.1371/journal.pgen.1000686

Meyer, S. E., McArthur, E. D., & Jorgensen, G. L. (1989). Variation in germination response to temperature in rubber rabbitbrush (*Chrysotamnus nauseosus*: Asteraceae) and its ecological implications. *American Journal of Botany*, *76*(7), 981–991. https://doi.org/10.1002/j.1537-2197.1989.tb15078.x

Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., Von Haeseler, A., Lanfear, R., & Teeling, E. (2020). IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. *Molecular Biology and Evolution*, *37*(5), 1530–1534. https://doi.org/10.1093/molbev/msaa015

Moran, B. M., Payne, C., Langdon, Q., Powell, D. L., Brandvain, Y., & Schumer, M. (2021). The genomic consequences of hybridization. *ELife*, *10*, 1–33. https://doi.org/10.7554/ELIFE.69016

Nei, M. (1972). Genetic distance between populations. *American Naturalist*, *106*(949), 283–292. https://doi.org/10.1086/282771

Nei, M., & Li, W. H. (1979). Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences of the United States of America*, *76*(10), 5269–5273. https://doi.org/10.1073/pnas.76.10.5269

Nesom, G. L., & Baird, G. I. (1993). Completion of *Ericameria* (Asteraceae: Astereae) diminution of Chrysothamnus. *Phytologia*, *75*(1), 74–93.

Nguyen, L. T., Schmidt, H. A., Von Haeseler, A., & Minh, B. Q. (2015). IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution*, *32*(1), 268–274. https://doi.org/10.1093/molbev/msu300

Nosil, P., & Feder, J. L. (2012). Genomic divergence during speciation: Causes and consequences. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences*, *367*(1587), 332–342. https://doi.org/10.1098/rstb.2011.0263

Novembre, J., & Stephens, M. (2008). Interpreting principal component analyses of spatial population genetic variation. *Nature Genetics*, *40*(5), 646–649. https://doi.org/10.1038/ng.139

Ogle, D., Cane, J., Fink, F., St. John, L., Stannard, M., & Dring, T. (2007). *Plants for pollinators in the Intermountain West Natural Resources Conservation Service (Technical Note 2A)* (pp 1–44). USDA—Natural Resources Conservation Service.

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., Stevens, M. H. H., Szoecs, E., & Wagner, H. (2019). *vegan: Community ecology package*. https://cran.r-project.org/package=vegan

Omernik, J. M., & Griffith, G. E. (2014). Ecoregions of the conterminous United States: Evolution of a hierarchical spatial framework. *Environmental Management*, *54*(6), 1249–1266. https://doi.org/10.1007/s00267-014-0364-1

Orr, H. A. (1996). Dobzhansky, Bateson, and the genetics of speciation. *Genetics*, *144*(4), 1331–1335. https://doi.org/10.1093/genetics/144.4.1331

Ortiz, E. M. (2019). *vcf2phylip v2.0: Convert a VCF matrix into several matrix formats for phylogenetic analysis*. https://doi.org/10.5281/zenodo.1257057

Parchman, T. L., Gompert, Z., Mudge, J., Schilkey, F. D., Benkman, C. W., & Buerkle, C. A. (2012). Genome-wide association genetics of an adaptive trait in lodgepole pine. *Molecular Ecology*, *21*(12), 2991–3005. https://doi.org/10.1111/j.1365-294X.2012.05513.x

Patterson, N., Price, A. L., & Reich, D. (2006). Population structure and Eigenanalysis. *PLoS Genetics*, *2*(12), e190. https://doi.org/10.1371/journal.pgen.0020190

Pebesma, E. (2018). Simple features for R: Standardized support for spatial vector data. *The R Journal*, *10*(1), 439–446. https://doi.org/10.32614/rj-2018-009

Pebesma, E. J., & Bivand, R. S. (2005). Classes and methods for spatial data in R. *R News*, *5*(2). https://cran.r-project.org/doc/Rnews/

Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., & Hoekstra, H. E. (2012). Double digest RADseq: An inexpensive method for *de novo* SNP discovery and genotyping in model and non-model species. *PLoS One*, *7*(5), e37135. https://doi.org/10.1371/journal.pone.0037135

Puritz, J. B., Hollenbeck, C. M., & Gold, J. R. (2014). dDocent: A RADseq, variant-calling pipeline designed for population genomics of non-model organisms. *PeerJ*, *2*, e431. https://doi.org/10.7717/peerj.431

R Core Team. (2020). *R: A language and environment for statistical computing*. https://www.r-project.org/

Rand, D. M., & Harrison, R. G. (1989). Ecological genetics of a mosaic hybrid zone: Mitochondrial, nuclear, and reproductive differentiation of crickets by soil type. *Evolution*, *43*(2), 432–449. https://doi.org/10.1111/j.1558-5646.1989.tb04238.x

Rellstab, C., Gugerli, F., Eckert, A. J., Hancock, A. M., & Holderegger, R. (2015). A practical guide to environmental association analysis in landscape genomics. *Molecular Ecology*, *24*(17), 4348–4370. https://doi.org/10.1111/mec.13322

Remington, C. L. (1968). Suture-zones of hybrid interaction between recently joined biotas. In T. Dobzhansky, M. K. Hecht, & W. C. Steere (Eds.), *Evolutionary biology* (pp. 321–428). Plenum Press.

Rieseberg, L. H., Archer, M. A., & Wayne, R. K. (1999). Transgressive segregation, adaptation and speciation. *Heredity*, *83*(Pt 4), 363–372. https://doi.org/10.1038/sj.hdy.6886170

Rieseberg, L. H., Kim, S. C., Randell, R. A., Whitney, K. D., Gross, B. L., Lexer, C., & Clay, K. (2007). Hybridization and the colonization of novel habitats by annual sunflowers. *Genetica*, *129*(2), 149–165. https://doi.org/10.1007/s10709-006-9011-y

Roberts, R. P., & Urbatsch, L. E. (2003). Molecular phylogeny of *Ericameria* (Asteraceae, Astereae) based on nuclear ribosomal 3' ETS and ITS sequence data. *Taxon*, *52*(2), 209–228. https://doi.org/10.2307/3647390

Rogers, L. E. (1979). *Shrub-inhabiting insects of the 200 area plateau, southcentral Washington*. U.S. Department of Energy. https://doi.org/10.2172/5918132

Ross, C. L., & Harrison, R. G. (2002). A fine-scale spatial analysis of the mosaid hybrid zone between *Gyllus firmus* and *Gryllus pennsylvanicus*. *Evolution*, *56*(11), 2296–2312. https://doi.org/10.1111/j.0014-3820.2002.tb00153.x

Schluter, D. (2001). Ecology and the origin of species. *Trends in Ecology & Evolution*, *16*(7), 372–380. https://doi.org/10.1016/s0169-5347(01)02198-x

Shafer, A. B. A., Cullingham, C. I., Cote, S. D., & Coltman, D. W. (2010). Of glaciers and refugia: A decade of study sheds new light on the phylogeography of northwestern North America. *Molecular Ecology*, *19*(21), 4589–4621.

Shastry, V., Adams, P. E., Lindtke, D., Mandeville, E. G., Parchman, T. L., Gompert, Z., & Buerkle, C. A. (2021). Model-based genotype and ancestry estimation for potential hybrids with mixed-ploidy. *Molecular Ecology Resources*, *21*(5), 1434–1451. https://doi.org/10.1111/1755-0998.13330

Shryock, D. F., Havrilla, C. A., DeFalco, L. A., Esque, T. C., Custer, N. A., & Wood, T. E. (2017). Landscape genetic approaches to guide native plant restoration in the Mojave Desert. *Ecological Applications*, *27*(2), 429–445. https://doi.org/10.1002/eap.1447

Sinclair, E. A., & Hobbs, R. J. (2009). Sample size effects on estimates of population genetic structure: Implications for ecological restoration. *Restoration Ecology*, *17*(6), 837–844. https://doi.org/10.1111/j.1526-100x.2008.00420.x

Stajich, J. E., & Hahn, M. W. (2005). Disentangling the effects of demography and selection in human history. *Molecular Biology and Evolution*, *22*(1), 63–73. https://doi.org/10.1093/molbev/msh252

Stankowski, S., & Ravinet, M. (2021). Defining the speciation continuum. *Evolution*, *75*(6), 1256–1273. https://doi.org/10.1111/evo.14215

Stephenson, N. (1998). Actual evapotranspiration and deficit: Biologically meaningful correlates of vegetation distribution across spatial scales. *Journal of Biogeography*, *25*(5), 855–870. https://doi.org/10.1046/j.1365-2699.1998.00233.x

Subramanian, S. (2016). The effects of sample size on population genomic analyses—Implications for the tests of neutrality. *BMC Genomics*, *17*(1), 1–13. https://doi.org/10.1186/s12864-016-2441-8

Swenson, N. G., & Howard, D. J. (2004). Do suture zones exist? *Evolution*, *58*(11), 2391–2397. https://doi.org/10.1111/j.0014-3820.2004.tb00869.x

Swenson, N. G., & Howard, D. J. (2005). Clustering of contact zones, hybrid zones, and phylogeographic breaks in North America. *The American Naturalist*, *166*(5), 581–591. https://doi.org/10.1086/491688

Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, *123*(3), 585–595. https://doi.org/10.1093/genetics/123.3.585. http://www.ncbi.nlm.nih.gov/pubmed/2513255

Teeter, K. C., Thibodeau, L. M., Gompert, Z., Buerkle, C. A., Nachman, M. W., & Tucker, P. K. (2010). The variable genomic architecture of isolation between hybridizing species of house mice. *Evolution*, *64*(2), 472–485. https://doi.org/10.1111/j.1558-5646.2009.00846.x

Thompson, J. N. (2005). *The geographic mosaic of coevolution*. University of Chicago Press. https://doi.org/10.7208/chicago/9780226118697.001.0001

Turner, S. D., Nagraj, V. P., Scholz, M., Jessa, S., Acevedo, C., Ge, J., Woerner, A. E., & Budowle, B. (2022). Evaluating the impact of dropout and genotyping error on SNP-based kinship analysis with forensic samples. *Frontiers in Genetics*, *13*(June), 1–9. https://doi.org/10.3389/fgene.2022.882268

Vieira, F. G., Albrechtsen, A., & Nielsen, R. (2016). Estimating IBD tracts from low coverage NGS data. *Bioinformatics*, *32*(14), 2096–2102. https://doi.org/10.1093/bioinformatics/btw212

Vines, T. H., Köhler, S. C., Thiel, M., Ghira, I., Sands, T. R., MacCallum, C. J., Barton, N. H., & Nürnberger, B. (2003). The maintenance of reproductive isolation in a mosaic hybrid zone between the fire-bellied toads *Bombina bombina* and *B. variegata*. *Evolution*, *57*(8), 1876–1888. https://doi.org/10.1111/j.0014-3820.2003.tb00595.x

Wagner, C. E., Keller, I., Wittwer, S., Selz, O. M., Mwaiko, S., Greuter, L., Sivasundar, A., & Seehausen, O. (2013). Genome-wide RAD sequence data provide unprecedented resolution of species boundaries and relationships in the Lake Victoria cichlid adaptive radiation. *Molecular Ecology*, *22*(3), 787–798. https://doi.org/10.1111/mec.12023

Wang, C., Szpiech, Z. A., Degnan, J. H., Jakobsson, M., Pemberton, T. J., Hardy, J. A., Singleton, A. B., & Rosenberg, N. A. (2010). Comparing spatial maps of human population-genetic variation using Procrustes analysis. *Statistical Applications in Genetics and Molecular Biology*, *9*(1), 1–20. https://doi.org/10.2202/1544-6115.1493

Wang, I. J., & Bradburd, G. S. (2014). Isolation by environment. *Molecular Ecology*, *23*(23), 5649–5662. https://doi.org/10.1111/mec.12938

Wangberg, J. K. (1981). Gall-forming habits of *Aciurina* species (Diptera: Tephritidae) on Rabbitbrush (Compositae: *Chrysothamnus* spp) in Idaho. *Journal of the Kansas Entomological Society*, *54*(4), 711–732.

Watterson, G. A. (1975). On the number of segregating sites in genetical models without recombination. *Theoretical Population Biology*, *7*(2), 256–276. https://doi.org/10.1016/0040-5809(75)90020-9

Weber, D. J., Davis, T. D., McArthur, E. D., & Sankhla, N. (1985). *Chrysothamnus nauseosus* (Rubber Rabbitbrush): Multiple-use shrub of the desert. *Desert Plants*, *7*(4), 172–210.

Whitham, T. G., Young, W. P., Martinsen, G. D., Gehring, C. A., Schweitzer, J. A., Shuster, S. M., Wimp, G. M., Fischer, D. G., Bailey, J. K., Lindroth, R. L., Woolbright, S., & Kuske, C. R. (2003). Community and ecosystem genetics: A consequence of the extended phenotype. *Ecology*, *84*(3), 559–573. https://doi.org/10.1890/0012-9658(2003)084[0559:caegac]2.0.co;2

Wilson, P., Wolfe, A. D., Armbruster, W. S., & Thomson, J. D. (2007). Constrained lability in floral evolution: Counting convergent origins of hummingbird pollination in *Penstemon* and *Keckiella*. *The New Phytologist*, *176*(4), 883–890. https://doi.org/10.1111/j.1469-8137.2007.02219.x

Wright, S. (1943). Isolation by distance. *Genetics*, *28*(2), 114–138. https://doi.org/10.1093/genetics/28.2.114

Yu, G., Smith, D. K., Zhu, H., Guan, Y., & Lam, T. T. Y. (2017). ggtree: An R Package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods in Ecology and Evolution*, *8*(1), 28–36. https://doi.org/10.1111/2041-210X.12628

Zhang, L., Hood, G. R., Carroo, I., Ott, J. R., & Egan, S. P. (2021). Context-dependent reproductive isolation: Host plant variability drives fitness of hybrid herbivores. *The American Naturalist*, *197*(6), 732–739. https://doi.org/10.1086/714139