ORIGINAL RESEARCH

Genotyping by sequencing reveals contrasting patterns of population structure, ecologically mediated divergence, and long-distance dispersal in North American palms

Anastasia Klimova[1,2] (ID) | Alfredo Ortega-Rubio[1] (ID) | David L. J. Vendrami[2] |
Joseph I. Hoffman[2] (ID)

[1]Centro de Investigaciones Biologicas del Noroeste S.C., La Paz, Baja California Sur, Mexico

[2]Department of Animal Behaviour, Bielefeld University, Bielefeld, Germany

Correspondence
Joseph I. Hoffman, Department of Animal Behaviour, Bielefeld University, Bielefeld, Germany.
Email: joseph.hoffman@uni-bielefeld.de
and
Alfredo Ortega-Rubio, Centro de Investigaciones Biologicas del Noroeste S.C., La Paz, Baja California Sur, Mexico.
Email: aortega@cibnor.mx

Abstract

Comparative studies can provide powerful insights into processes that affect population divergence and thereby help to elucidate the mechanisms by which contemporary populations may respond to environmental change. Furthermore, approaches such as genotyping by sequencing (GBS) provide unprecedented power for resolving genetic differences among species and populations. We therefore used GBS to provide a genomewide perspective on the comparative population structure of two palm genera, *Washingtonia* and *Brahea*, on the Baja California peninsula, a region of high landscape and ecological complexity. First, we used phylogenetic analysis to address taxonomic uncertainties among five currently recognized species. We resolved three main clades, the first corresponding to *W. robusta* and *W. filifera*, the second to *B. brandegeei* and *B. armata*, and the third to *B. edulis* from Guadalupe Island. Focusing on the first two clades, we then delved deeper by investigating the underlying population structure. Striking differences were found, with GBS uncovering four distinct *Washingtonia* populations and identifying a suite of loci associated with temperature, consistent with ecologically mediated divergence. By contrast, individual mountain ranges could be resolved in *Brahea* and few loci were associated with environmental variables, implying a more prominent role of neutral divergence. Finally, evidence was found for long-distance dispersal events in *Washingtonia* but not *Brahea*, in line with knowledge of the dispersal mechanisms of these palms including the possibility of human-mediated dispersal. Overall, our study demonstrates the power of GBS together with a comparative approach to elucidate markedly different patterns of genomewide divergence mediated by multiple effectors.

KEYWORDS
Arecaceae, Baja California peninsula, genotyping by sequencing (GBS), human-mediated dispersal, single nucleotide polymorphism (SNP)

1 | INTRODUCTION

Comparative phylogeographic studies of codistributed species are of paramount importance in decoding how shared geological and ecological histories may have affected contemporary species and populations (Papadopoulou & Knowles, 2016). By comparing the spatial genetic structure of multiple codistributed species, comparative studies allow the assessment of phylogeographical congruence, which is a baseline for historical inference (Avise, 2000; Garrick, Rowell, Simmons, Hillis, & Sunnucks, 2008; Hickerson et al., 2010; Hoffman, Clarke, Linse, & Peck, 2011). Moreover, a general understanding of how species and populations responded to past challenges might provide us with insights into how present and future human activities will impact the planets biota (Avise, Bowen, & Ayala, 2016).

It is also becoming increasingly recognized that interactions between organisms and their environment can shape the distribution of spatial genetic variation and lead to local adaptation (Anderson, Willis, & Mitchell-Olds, 2011). This in turn may reduce the amount of gene flow between populations, ultimately putting them on the path to speciation (Rundle & Nosil, 2005; Schluter, 2000). The primary difference between ecologically mediated divergence and neutral divergence is that the former results in a pattern of isolation by environment (IBE) in which genetic and environmental distances are positively correlated independently of geographic distance (Shafer & Wolf, 2013; Wang & Bradburd, 2014). By contrast, the neutral divergence model emphasizes the role of geographic isolation in restricting the exchange of migrants between populations, which results in a pattern of isolation by distance (IBD) that is frequently observed in natural populations (Prunier et al., 2017; Wright, 1943).

Although species that share the same landscape may face similar selective environments, their capacity to adapt to these environments may be species- or even population-specific and can depend on both extrinsic and intrinsic factors, including the strength and nature of selection, the amount of genetic diversity, and the extent of phenotypic plasticity (Sork, Gugger, Chen, & Werth, 2016). Consequently, species inhabiting a common landscape may vary in their capacity to adapt to ecological gradients and changing environments. Assessing adaptive responses among related species within shared heterogeneous landscapes may thus help us to understand patterns of biodiversity. A major step toward this goal is identifying to what extent ecology, gene flow and genomic architecture contribute toward variability in the evolutionary potential of multiple-related species inhabiting the same landscape (Raeymaekers et al., 2017).

In this regard, the Baja California peninsula, with its heterogeneous array of landscapes and habitats varying from tropical deciduous forests and mesic oases to xeric desert scrub mountains and low altitude arid plains, is of particular interest (Dolby, Bennett, Lira-Noriega, Wilder, & Munguia-Vega, 2015). The peninsula was formed around 5–10 million years ago (mya), when tectonic forces gave rise to the Gulf of California and separated a narrow section of land from the Mexican mainland. Currently, a series of mountain ranges, most

prominently Sierra San Pedro Martir, Sierra Libertad, and Sierra La Giganta, run in succession along the peninsula from north to south. It has been suggested that the uplift of these major sierras probably began around 6–10 mya and that their formation was tied directly to the same tectonic forces that opened the Gulf of California and created the San Andreas Fault system (Martín-Barajas, 2014; Mueller, Kier, Rockwell, & Jones, 2009). Together, the sierras constitute a more or less continuous mountainous backbone that separates ecosystems sloping east into the Gulf of California from those running westwards into the Pacific and also creates a complex landscape gradient along the entire peninsular that is reflected in a very high diversity of environments (Riemann & Ezcurra, 2007). This unique natural setup, characterized by unusually high levels of both landscape and ecological complexity, could conceivably shape the population structure of species inhabiting the peninsula via both neutral and ecologically mediated processes, with their relative importance depending on intrinsic qualities of the species in question (Dolby et al., 2015).

An additional dimension is provided by the more recent history of the Baja California peninsula, which experienced a notable ecological shift involving progressive aridification after the Last Glacial Maximum (Lindell, Ngo, & Murphy, 2006; Riddle, Hafner, Alexander, & Jaeger, 2000). Drastic changes in precipitation across the peninsula are believed to have led to several endemic species becoming either locally extinct or restricted to small isolated sites such as oases and canyons (Bernardi, Ruiz-Campos, & Camarena-Rosales, 2007; Wehncke, López-Medellín, & Ezcurra, 2010). This major ecological transition is thought to have most strongly affected species dependent on freshwater such as amphibians (Grismer, 2002), fishes (Bernardi et al., 2007) and many plants (León de la Luz, 2014; Wehncke & López-Medellín, 2014). However, the sessile nature of plants renders them particularly susceptible to the influence of fine-scale environmental heterogeneity (Sork et al., 2016), which makes the flora of the Baja peninsula ideally suited to exploring the relative contributions of geography and ecology toward species-specific responses to aridification.

Palms (Arecaceae) are a species-rich taxonomic group that has long been considered a model plant family for evolutionary and ecological studies in the tropics (Bacon et al., 2016; Couvreur & Baker, 2013; Kissling et al., 2012; Savolainen et al., 2006). Although palms mainly occur in tropical and subtropical climates (Tregear, Rival, & Pintaud, 2011), a few species can also be found in desert regions (Dransfield et al., 2008; Tomlinson, 2006). Palm populations on the Baja California peninsula are widely considered to be relicts of historically more widespread and continuous populations that are now largely confined to sites where permanent water exists either above or below the ground (Axelrod, 1958; Bacon, Baker, & Simmons, 2012; Cornett, 1985a,b; Cornett, Glen, & Stewart, 1986; Felger & Joyal, 1999; Grismer & McGuire, 1993). This is reflected by the fossil record, which shows that during the late Cretaceous, palms were common across North America and extended much further than their current geographic distribution (Couvreur, Forest, & Baker, 2011; Harley, 2006). Today, the only suitable natural habitats

for tropical species like palms on the Baja California peninsula can be found in small isolated pockets at the bottoms of canyons or oases that are separated by a continuous expanse of mountainous deserts and xeric shrubland (Grismer, 2000; Hafner & Riddle, 2011; Minnich, Franco-Vizcaíno, & Salazar-Ceseña, 2011; Wehncke & López-Medellín, 2014).

On the Baja peninsula, the Arecaceae is represented by two native North American palm genera, *Washingtonia* and *Brahea* (Garcillán, Vega, & Martorell, 2012; Minnich et al., 2011). *Washingtonia* is represented by two species, *W. robusta* and *W. filifera*, although in general taxonomic relationships within this genus remain poorly resolved (Felger & Broyles, 2007; Felger & Joyal, 1999; Henderson, Galeano, & Bernal, 1995; McClintock, 1978) as they are mainly based on morphological characters such as size, leaf shape and inflorescence structure. *W. robusta* is more abundant on the southern part of the peninsula, while *W. filifera* mainly occupies the northern peninsula as well as southeastern California (Minnich et al., 2011). *W. robusta* is also found in a few riparian canyons at the southern edge of the Sonoran Desert on the Mexican mainland (Felger & Joyal, 1999).

The *Brahea* complex comprises nine species, two of which are endemic and restricted to the Baja California peninsula (*B. brandegeei* and *B. armata*) and one (*B. edulis*) to Guadalupe Island, which lies 260 km off the Pacific coast of the peninsula (Garcillán et al., 2012; Moran, 1996; Oberbauer, 2005). *Brahea* is arguably the least studied genus of American palms, and consequently, relationships between and within *Brahea* species have not been clearly described (Henderson et al., 1995; Minnich et al., 2011; Quero, 1992). Nevertheless, *B. brandegeei* has been described as occupying the southern half of the Baja California peninsula, from Sierra La Laguna at the southernmost tip of the peninsula to Sierra San Francisco in the northern Baja California Sur (Felger, Johnson, & Wilson, 2001; Minnich et al., 2011), while *B. armata* has been described as extending northwards from the state line of Baja California Sur in the central peninsula to just south of the United States–Mexico border (Franco-Vizcaíno, López-Beltrán, & Salazar-Ceseña, 2007; Wiggins, 1980). However, the exact distributional limits of *B. brandegeei* and *B. armata* remain somewhat unclear due to taxonomic uncertainties (Felger & Joyal, 1999; Felger et al., 2001; Henderson et al., 1995). Finally, *B. edulis* is an endemic species found uniquely on Guadalupe Island (29°N, 118°W), a seven-million-year-old volcanic island located 260 km west of the Baja California peninsula (Batiza, 1977). It is generally believed that Guadalupe Island has never been in contact with either the Mexican mainland or the Baja peninsula, which would imply that *Brahea* palms colonized the island via long-distance dispersal. Currently, these palms mainly occupy a small area of fog oasis in the far northern part of the Island (Garcillán et al., 2012; León de la Luz, Rebman, & Oberbauer, 2003; Oberbauer, 2005).

From an ecological perspective, palm populations of the Baja peninsula generally form small local colonies or elongate galleries, with plants growing almost entirely in areas with supplemental water, that is, in close vicinity to springs or along water courses (Franco-Vizcaíno et al., 2007; Minnich et al., 2011). *Washingtonia* and *Brahea* co-occur along the slopes of most of the sierras, including

Sierra Juarez, Sierra Asamblea, Sierra Mechudo and Sierra La Laguna (Klimova, Hoffman, Gutierrez-Rivera, Leon de la Luz, & Ortega-Rubio, 2017; Minnich et al., 2011). However, they differ somewhat in their ecological requirements, with their estimated niche overlap ranging from around 0.5–0.7 depending on the statistic used (Klimova et al., 2017). In particular, *Washingtonia* palms have stricter ecological requirements than *Brahea* palms and therefore occupy a smaller total area of the Baja California peninsula (Minnich et al., 2011). They are most commonly found at low-elevation oases where water is relatively plentiful and temperatures tend to be warmer and more stable. By contrast, *Brahea* palms are capable of tolerating drier conditions and thus occupy a much wider elevational range. Accordingly, *W. filifera* and *W. robusta* are seldom found above 1,000 m, whereas *B. armata* may grow at elevations as high as 1,400 m in the Sierra San Pedro Martir and *B. brandegeei* occurs at up to around 1,700 m in the Sierra La Laguna (Minnich et al., 2011).

Species-specific modes and patterns of dispersal are another key determinant of genetic structure in the Arecaceae (Eiserhardt, Svenning, Kissling, & Balslev, 2011). Although vertebrates such as birds, bats, coyotes, and foxes have been proposed as possible seed dispersal agents for both *Washingtonia* and *Brahea*, the effect of these agents on the population structure and genetic diversity of palms has not been evaluated (Cornett, 2008; Silverstein, 2005). Furthermore, recent studies of *Brahea* palms suggest that water pulses may be far more significant in terms of seed dispersal than birds and small mammals (Wehncke, López-Medellín, & Ezcurra, 2009; Wehncke et al., 2010). This leads to the prediction that dispersal in *Brahea* palms should occur mainly within watercourses or canyons and hence that connectivity among sierras will be restricted. A further complication is the potential influence of humans, which is predicted to be stronger for *Washingtonia* as this genus is widely considered to have been favored by indigenous people over *Brahea* as an important source of both food and building materials (Felger & Joyal, 1999; Felger & Moser, 1985; Minnich et al., 2011).

For the reasons described above, *Washingtonia* and *Brahea* palms provide an attractive system for investigating the contributions of neutral, non-neutral and human-mediated effects toward population structure in a comparative context. However, despite the emblematic status of these desert palms, many aspects of their taxonomy remain ambiguous (Minnich et al., 2011). This is partly because previous genetic studies do not always support currently recognized species. For example, Bacon et al. (2012) did not find any evidence in support of the designation of *W. robusta* and *W. filifera* as separate species based on three plastid and three nuclear genes. This ambiguity is to some extent reflected by a more recent study of *Washingtonia* and *Brahea* palms from the Baja California peninsula and adjacent areas, again based on nuclear and chloroplast sequence data (Klimova et al., 2017). Here, both genera were found to exhibit low genetic diversity and minimal structuring within the peninsula, similarly to previously reported for *W. filifera* in a small region of the Californian mainland using allozymes (McClenaghan & Beauchamp, 1986). However, *W. filifera* could be distinguished from *W. robusta* based on chloroplast but not nuclear DNA, whereas *B. edulis* was

divergent from its peninsular sister species based on nuclear, but not chloroplast DNA. Such incongruences together with the generally low genetic resolution provided by the markers used in this study precluded comprehensive hypothesis testing, and thus, more detailed inferences could not be made. Consequently, at the present time, even the geographic limits of the *Washingtonia* and *Brahea* species present on the Baja California peninsula remain unclear (Henderson et al., 1995; Klimova et al., 2017; Minnich et al., 2011), while virtually nothing is known about relationships among populations from different sierras.

Studies based on one or handful of genes like those described above may also suffer from a number of biases related to stochastic processes (Heath, Hedtke, & Hillis, 2008; Moore, 1995; Rokas & Carroll, 2005). However, approaches capable of genotyping thousands of single nucleotide polymorphisms (SNPs) such as genotyping by sequencing (GBS) are capable of providing much greater coverage of the genome (De Donato, Peters, Mitchell, Hussain, & Imumorin, 2013; Elshire et al., 2011). Recent simulation and empirical studies suggest that these approaches should be superior to a handful of markers at capturing variation in drift, selection, recombination, and mutation (Morin, Luikart, & Wayne, 2004) and thereby provide a more accurate depiction of population differentiation (Spinks, Thomson, & Shaffer, 2014; Vendrami et al., 2017) and genetic diversity (Fischer et al., 2017; Hoffman et al., 2014).

To address the issues described above, we combined GBS with near-exhaustive sampling of all five native *Washingtonia* and *Brahea* palm species present on the Baja California peninsula and Guadalupe Island. The resulting data were then analyzed on two levels. First, we attempted to resolve taxonomic relationships and from there to delimit the geographic boundaries of each species on the peninsula. Second, we focused within the major clades identified by the former analysis and conducted population genetic analyses to uncover patterns of population structure on the Baja peninsula and investigate

the potential underlying drivers. Our main working hypotheses were as follows: (1) We expected to find genomic support for most if not all of the currently recognized species, with the possible exception of *W. robusta* and *W. filifera*; (2) as water pulses appear to be an important mediator of dispersal in *Brahea* palms, we hypothesized for this genus that dispersal would be mainly restricted within watersheds, which should be reflected in stronger population structure than in *Washingtonia* and potentially manifested in a pattern whereby each sierra is genetically distinct; (3) we hypothesized that ecologically mediated selection should be comparably more important in *Washingtonia* palms due to their stricter ecological requirements, which could potentially lead to an IBE pattern. By contrast, we expected *Brahea* palms to be more influenced by neutral processes and thus to exhibit an IBD pattern; (4) finally, due to the long historical association between *Washingtonia* palms and humans, we hypothesized that population structure in *Washingtonia* could potentially also have been influenced by human-mediated translocation events.

2 | MATERIALS AND METHODS

2.1 | Sample collection

We collected a total of 190 leaf samples from all five palm species native to the Baja California peninsula and Guadalupe Island (Figure 1; Table S1). Our sample size reflects the difficulty of collecting samples from endemics that are locally rare and can only be found at small and isolated stands that can often only be reached by foot. We collected specimens from virtually all accessible oases during two consecutive field seasons, each lasting ~4 weeks. Whenever possible, we avoided sampling immediately adjacent individuals as Migliore et al. (2013) showed that around a third of adjacent sampled individuals of a relict shrub species were clones. For comparison, we also included two populations of *W. robusta*
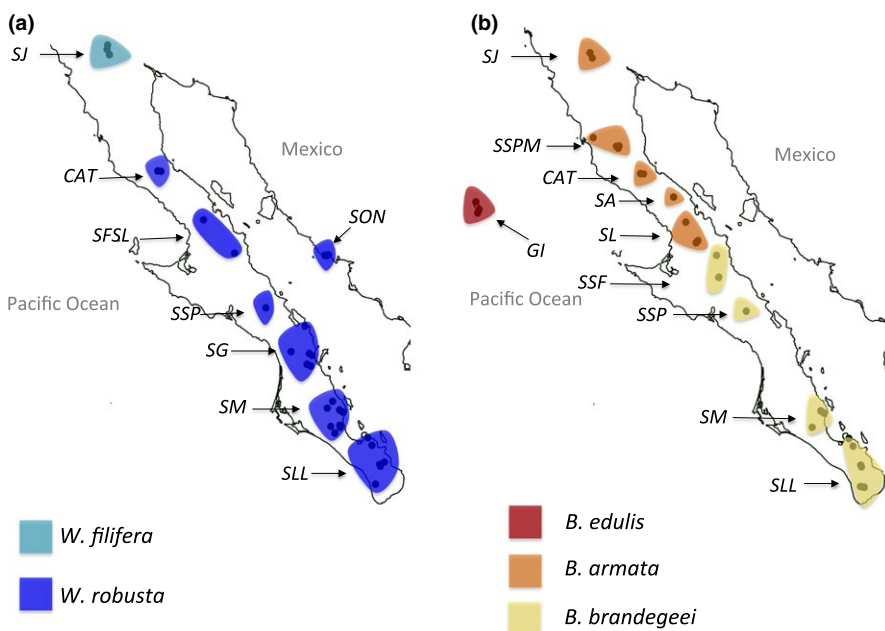
FIGURE 1 Maps showing the oases (points) and the sierras (shape files) from which (a) *Washingtonia* and (b) *Brahea* palms were sampled. The full names of the oases and sample sizes are given in Table S1. The colors of the shapefiles correspond to species (as defined in Minnich et al., 2011): (a) *W. robusta* on the Baja peninsula and Mexican mainland and *W. filifera* at *SJ*; (b) *B. edulis* at the Guadalupe Island (*GI*), *B. armata* at *SL*, *SA*, *CAT*, *SSPM*, and *SJ* and *B. brandegeei* at *SLL*, *SM*, *SPP*, and *SSF*

from Sonora on the Mexican mainland. We were therefore able to cover the full distributional ranges of three *Brahea* species (*B. edulis*, *B. armata*, and *B. brandegeei*) and one *Washingtonia* species (*W. robusta*), while *W. filifera* could only be sampled from its southern distributional limit in Sierra Juarez, Mexico (Figure 1; Table S1). Specimens were assigned to taxa based on species distributions given by Minnich et al. (2011). Within species, populations were defined based on the sierra from which the samples were collected. For *Washingtonia*, we specified eight populations corresponding to palms from (1) Sierra La Laguna (SLL), (2) Sierra Mechudo (SM), (3) Sierra Giganta (SG), (4) Sierra San Pedro (SSP), (5) Sierra San Francisco combined with Sierra Libertad (SFSL), (6) Cataviña (CAT), (7) Sierra Juarez (SJ), and (8) Mexican mainland, state of Sonora (SON). For *Brahea*, we specified nine populations on the Baja peninsula corresponding to (1) Sierra La Laguna (SLL), (2) Sierra Mechudo (SM), (3) Sierra San Pedro (SSP), (4) Sierra San Francisco (SSF), (5) Sierra Libertad (SL), (6) Sierra Asamblea (SA), (7) Cataviña (CAT), (8) Sierra San Pedro Martir (SSPM), and (9) Sierra Juarez (SJ) (Figure 1; Table S1).

2.2 | Genetic analysis

Total genomic DNA was extracted from silica-dried leaves using a modified CTAB protocol (Gutierrez-Rivera in preparation), and 50 µl of DNA from each sample was sent to the Cornell Institute of Genomic Diversity for library preparation and GBS (Elshire et al., 2011). Each DNA extract was digested using the restriction enzyme PstI, and subsequently, a sample-specific barcoded adapter and a common adapter were ligated to the sticky ends of fragments to allow for sample discrimination after pooling. A unique barcoded adapter was used for each sample plus two negative controls, giving a total of 192 barcodes employed. Next, samples from the two different genera were pooled together into two separate libraries that were each 100 bp single-end sequenced on one lane of an Illumina HiSeq 2000.

The resulting raw reads were processed using the TASSEL 3.0 pipeline (Bradbury et al., 2007) which implements the customized workflow specifically designed for GBS data described by Glaubitz et al. (2014). Specifically, all identical reads were first collapsed into tags and the number of reads used for the generation of each tag was reported. Before this step, in order to ensure the usage of exclusively high-quality reads, only reads containing a barcode, the correct restriction enzyme cutting site and with no Ns were retained. After removing the barcode sequences, the remaining reads were then trimmed to a final length of 64 bp, with any reads containing a second restriction site being truncated. Then, a "master" list of tags was created containing only tags built from at least three reads (i.e., the minimum depth of coverage of a tag had to be three reads). This number was chosen because lower values will result in more sequencing errors being included, while higher values will result in the exclusion of rare alleles. We also chose a value of three because this allowed tags containing up to three sequencing errors to be assembled, which increased the depth of coverage of non-error positions.

Sequencing errors were subsequently removed by filtering out SNPs with low minor allele frequencies (MAF) as described below.

Finally, the master list of tags was aligned to the date palm (*Phoenix dactylifera*) reference genome (date palm, downloaded from https://www.ncbi.nlm.nih.gov/genome/ in June 2016) using the software BWA (Li & Durbin, 2009) to produce a SAM file. After conversion of this file with SAMConvertor, the tbt2vcf plug-in within TASSEL 3.0 was used to call SNPs from tags that aligned to unique locations in the reference genome, which were then exported in VCF format. SNPs with MAF below than 0.01 were then removed from the dataset in order to filter out false SNPs originating from sequencing errors while retaining genuine variants including moderately rare alleles. InDels were then removed from the variant dataset, which was further filtered to discard chloroplast and mitochondrial variants, SNPs that were not bi-allelic and loci carrying only heterozygote genotypes, which represent putatively paralogous loci. Next, in order to avoid linkage between loci, we pruned the SNPs for linkage disequilibrium using the thinning option in VCFTOOLS (Danecek et al., 2011) with a 5 k sliding window. Finally, we removed loci with more than 20% missing genotypes. This conservative measure was taken to minimize the frequency of gaps and thereby to maximize the robustness of our downstream analyses. The final dataset was used for phylogenetic analysis and to compare levels of genetic diversity between the two genera. After that, we generated two separate datasets for *Washingtonia* and *Brahea*, respectively, to allow finer-scale population genetic analyses. These datasets were generated for each genus using the same filtering steps described above. Additionally, five individuals with more than 30% missing data (four *Washingtonia* and one *Brahea*) were excluded from further analyses. Manipulations of the VCF files were carried out using VCFTOOLS. Inbreeding coefficients and observed and expected heterozygosities were calculated for each palm genus separately using PLINK 1.9 (Chang et al., 2015).

2.3 | Comparative phylogenetic analyses

In order to resolve taxonomic relationships among the morphologically defined species and to delimit the geographic boundaries of each species, we performed Bayesian phylogenetic reconstruction using the SNAPP package (Bryant, Bouckaert, Felsenstein, Rosenberg, & Choudhury, 2012), within the program BEAST2 (Bouckaert et al., 2014). Due to computational limitations, we were unable to analyze the full dataset and therefore restricted our analysis to a random selection of three individuals per population, resulting in a total of 54 individuals representing two palm genera. We used the default prior and model parameters, including the defaults for u and v (the backward and forward mutation rates, respectively) and ran a single Markov chain Monte Carlo (MCMC) chain of 2,000,000 iterations with sampling every 1,000 steps. After running the full dataset, we then analyzed each genus separately using same parameters as before. Acceptable mixing (requiring effective sample size values to be at least 200) and convergence were checked by visual inspection of the posterior samples using TRACER (Rambaut, Suchard, Xie, &

Drummond, 2014). We used a burn-in of 10% and visualized the distribution of trees using DENSITREE 2.1 (Bouckaert, 2010).

2.4 | Population structure analyses

Based on the results of phylogenetic analyses (see Section 3), we decided to exclude *B. edulis* from population-level analyses or where appropriate to use it as an outgroup. Downstream population genetic analyses were conducted after defining samples from each of the sierras as *a priori* populations (Figure 1; Table S1). First, pairwise F_{st} values among sierras were calculated within the program GENODIVE (Meirmans & Van Tienderen, 2004) with statistical significance determined on the basis of 10,000 permutations. Then, we used sNMF 1.2 (Frichot, Mathieu, Trouillon, Bouchard, & François, 2014) to estimate individual admixture coefficients and to determine the most probable number of genetic clusters (k) present within each genus. sNMF was chosen in preference to more computer-intensive approaches such as STRUCTURE (Pritchard, Stephens, & Donnelly, 2000) as it uses fast and efficient sparse non-negative matrix factorization algorithms that considerably reduce the computational burden without any appreciable loss of accuracy (Frichot et al., 2014; Popescu, Harper, Trick, Bancroft, & Huber, 2014; Wollstein & Lao, 2015). The best k value was inferred by calculating cross-entropy values from multiple runs with k set between one and ten. Robustness of the results was assessed by running five replicates for the best value of k using an alpha regularization parameter of 100. We conducted model averaging of individual ancestry coefficients across replicates and calculated the average pairwise similarity of individual assignments across runs using CLUMPP (Jakobsson & Rosenberg, 2007). Finally, we converted the SNP data into a matrix of individual pairwise genetic distances using the R package STAMPP (Pembleton, Cogan, & Forster, 2013) and generated a phylogenetic network using the NEIGHBORNET algorithm (Bryant & Moulton, 2004) within SPLITSTREE 4.14.4 (Huson & Bryant, 2006).

2.5 | Population splits and migration modeling

We used the approach of Pickrell and Pritchard (2012) to infer the population history of the palm taxa using genome-wide allele frequency data as implemented in TREEMIX 1.12 (Pickrell & Pritchard, 2012). TREEMIX infers gene flow between populations by simultaneously analyzing population divergence and admixture. On the resulting maximum-likelihood (ML) tree, migration events are represented by edges that connect populations via admixture. SNP data for *Washingtonia* and *Brahea* were converted from a diploid genotype format into population-level allele counts using the python script plink2treemix.py (available with TREEMIX). Each population was represented by individuals from a given sierra as described previously. For *Washingtonia*, we used the most divergent sierra (SJ) as an outgroup, while for *Brahea*, we used *B. edulis* as an outgroup. We first generated a maximum-likelihood graph with no migration events based on 1,000 bootstrap replicates. Then, we tested for between one and ten migration events per taxon (m1–m10) and

performed likelihood ratio tests to allow stepwise comparison of log-likelihood values between each pair of migration events. After that we formally tested for admixture using the "three-population test" (Patterson et al., 2012) implemented in TREEMIX. This is a formal test that can provide evidence of admixture even in the presence of past migration events (Patterson et al., 2012). It allows detection of the presence of admixture in population X from other two populations, A and B; if the value of f3 (X; A, B) is negative, then the deviation from "treeness" is detected and X appears to be a mixture of A and B.

2.6 | Isolation by distance and ecological divergence

Isolation by distance (IBD) and dispersal barriers are known to contribute toward the geographic structuring of genetic variation in many organisms. We therefore used the Isolation By Distance Web Service 3.23 (Jensen, Bohona, & Kelley, 2005) to perform reduced major axis regression and Mantel tests based on 10,000 randomizations of the datasets. Geographic great-circle distance and pairwise genetic distances between individuals were calculated using the Geographic Distance Matrix Generator 1.2.3 (Ersts, 2016) and the R package STAMPP (Pembleton et al., 2013) respectively. Local adaptation can also be manifested in correlations between genetic and environmental distances (Frankham, Ballou, & Briscoe, 2002). We therefore used Mantel and partial Mantel tests as implemented in the R package VEGAN 2.4-0 (Oksanen et al., 2013) to test for correlations between genetic and environmental distances, the latter being generated using the "dist" function in R. As *Washingtonia* and *Brahea* are affiliated to humid tropical climates and are frost sensitive, we expected climatic variables such as the mean temperature of the coldest quarter, precipitation of the driest quarter, and the aridity index to have the greatest influence on these genera. The above-mentioned ecological information was therefore downloaded from WorldClim with the resolution of 30 arc-seconds (~1 km) (Hijmans, Cameron, Parra, Jones, & Jarvis, 2005) and from the Global Aridity and PET database (Zomer, Trabucco, Bossio, & Verchot, 2008) as a set of raster layers. Mantel tests were then performed between each genetic and environmental distance matrix, and these analyses were also repeated as partial Mantel tests controlling for geographic distance. Statistical significance was determined using Pearson tests based on 10,000 permutations of the data.

2.7 | Detection of outlier loci associated with environmental variables

Environmental variables showing significant associations with genetic distance in the above analyses (specifically, mean temperature of the coldest quarter in *Washingtonia* and precipitation in the driest quarter for *Brahea*, see Section 3) were further investigated by testing for signatures of local adaptation using the R package LEA (Landscape Genomics and Ecological Association Test, Frichot, Schoville, Bouchard, & François, 2013; Frichot & François, 2015) and SAMβADA (Stucki et al., 2016). The first of these programs uses

latent factor mixed models LFMMs to detect loci exhibiting unusual associations with environmental variables compared to the genomic background. We chose to use this program as it can account for the underlying population structure by introducing "latent factors" while simultaneously estimating random effects due to population history and isolation by distance. We ran 10,000 iterations of the Gibbs sampling algorithm with the first 5,000 iterations discarded as burn-in. Z scores from five independent replicate runs were then combined and the resulting p-values were adjusted for the false discovery rate (FDR) as described in Benjamini and Hochberg (1995) with an alpha level of 0.05.

The second package, SAMβADA, implements logistic regressions to model the probability of observing a particular genotype at each marker given the environmental conditions at the sampling locations (Joost et al., 2007). We chose the multivariate option as this allows a combination of predictor variables to be simultaneously assessed, thereby reducing the occurrence of spurious genotype by environment associations (Stucki et al., 2016). The two predictor variables for both Washingtonia and Brahea were the genetic groups identified with SPLITSTREE and the respective ecological variable identified using partial Mantel tests. Any SNPs associated with the main genetic groups were then discarded and only those SNPs associated with ecological variables were retained. Statistical significance was determined using both log-likelihood ratio and Wald tests (Joost et al., 2007) and FDR was applied to the resulting p-values with an alpha level of .05.

3 | RESULTS

3.1 | Genomic data

We subjected 190 palm samples to GBS, generating a total of 296,358,035 high-quality barcoded reads, which were assembled into 549,976 tags that aligned uniquely to the P. dactylifera reference genome. From these data, we called a total of 26,565 SNPs. After quality filtering, removing InDels, retaining only bi-allelic nuclear SNPs, LD filtering and removing loci with more than 20% missing data, this was reduced to 2,063 SNPs distributed over 724 scaffolds (median = one SNP per scaffold, range = 1–30, see Figure S1). The final dataset comprised 514 diagnostic SNPs (i.e., loci that were fixed for different alleles in the two genera) plus 1,549 polymorphic SNPs, of which 183 were polymorphic in both genera, 312 were polymorphic only in Washingtonia and 1,054 were polymorphic only in Brahea (Figure S2). Additionally, we generated separate datasets for each palm genus comprising 85 Washingtonia palms genotyped at 1,462 polymorphic SNPs and 79 Brahea palms (excluding B. edulis) genotyped at 2,050 polymorphic SNPs. Observed heterozygosity was lower in Washingtonia (0.098, 97.5% CI = 0.089–0.108) than in Brahea (0.139, 97.5% CI = 0.131–0.146), while the opposite was found for expected heterozygosity (Washingtonia: 0.209, 97.5% CI = 0.200–0.217; Brahea: 0.175, 97.5% CI = 0.168–0.182). Consistent with differences in observed heterozygosity, the genomic inbreeding coefficient F_{hat3} was higher in Washingtonia (0.47, CI: 0.37–0.57) than in Brahea (0.25, CI: 0.21–0.29).

3.2 | Phylogenetic relationships

To elucidate taxonomic relationships, we constructed a Bayesian phylogenetic tree based on a subset of 54 Washingtonia and Brahea individuals (Figure 2). As expected, the two genera were clearly resolved as distinct and deeply divergent monophyletic clades (Figure 2a). However, when each genus was analyzed separately, we found little evidence in support of the majority of morphologically defined species. Specifically, W. filifera did not form a monophyletic group, but instead grouped together with W. robusta palms from the northern Baja California peninsula (SSP, SFSL, and CAT) as shown in

FIGURE 2 Bayesian reconstruction of the phylogenetic relationships (a) between *Washingtonia* and *Brahea*; (b) within *Washingtonia*; (c) within *Brahea*. Colored lines delimit morphologically defined species according to Minich et al. (2011), and colored dots correspond to the sampled populations. Major nodes with high posterior probability support (>0.95) are indicated in asterisks (*)
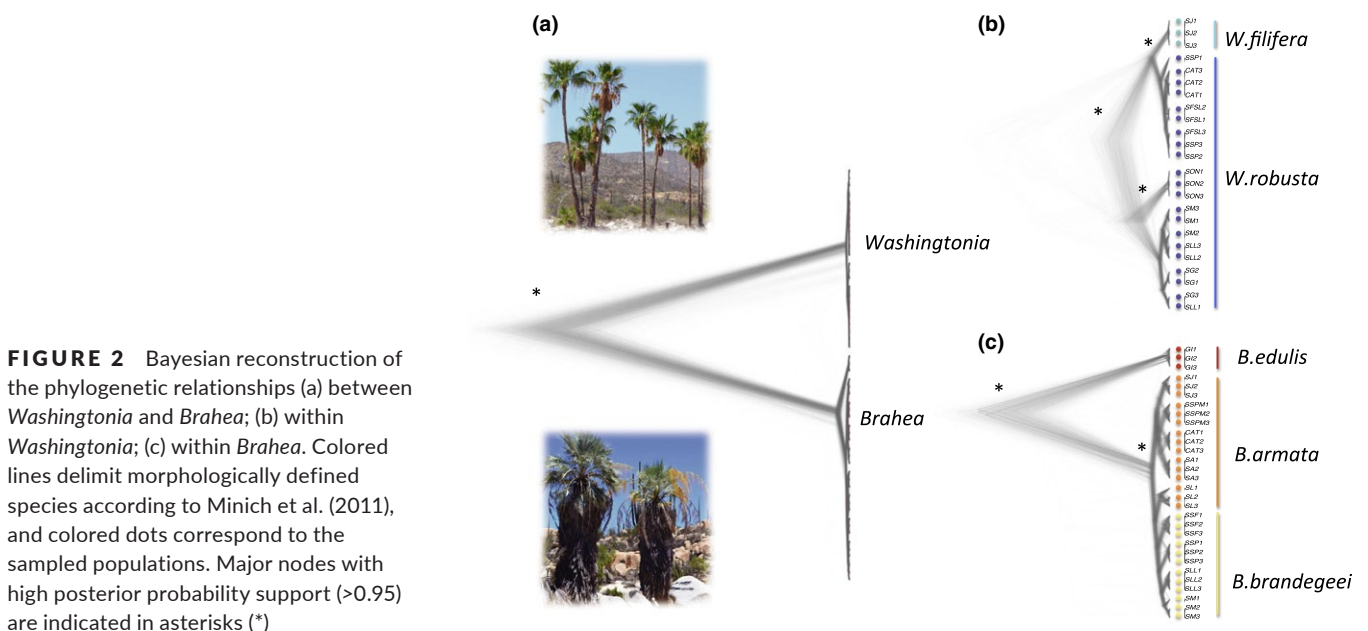
Figure 2b. Furthermore, greater divergence was observed between *W. robusta* from the Mexican mainland (*SON*) and its peninsular conspecifics than between *W. filifera* and *W. robusta* from the northern Baja California peninsula. This suggests that any genetic differences between *W. filifera* and *W robusta* are smaller than the magnitude of intraspecific variation within *W. robusta*.

By contrast, two distinct monophyletic clades were resolved in *Brahea*, the first corresponding to *B. edulis* from Guadalupe Island and the second comprising *B. armata* and *B. brandegeei* (Figure 2c). Within the peninsular clade, individuals diverged from each other not based on morphological species designations but according to the sierras they were collected from. On the bases of the above analyses, we therefore defined three major palm clades on the Baja California peninsula and Guadalupe Island comprising (1) *W. robusta* and *W. filifera*; (2) *B. brandegeei* and *B. armata*; and (3) *B. edulis*.

3.3 | Population structure

Next, we carried out population genetic analyses to investigate the comparative population structure of *Washingtonia* and *Brahea* palms on the Baja California peninsula. These analyses were conducted separately for the two clades identified above corresponding to *W. robusta* and *W. filifera* (forthwith referred to as *Washingtonia*) and *B. brandegeei* and *B. armata* (forthwith referred to as *Brahea*). Strong population structure was found in both genera, with majority of pairwise comparisons among sierras yielding moderately large and highly significant F_{st} values (Tables S2 and S3). For *Washingtonia*, the greatest genetic differences were observed between the peninsular and mainland localities (F_{st} = 0.55–0.86, $p < .001$) as well as between the northernmost population of *SJ* and the other sierras (F_{st} = 0.46–0.86, $p < .001$). Additionally, sierras of the northern (*CAT*, *SFSL*, and *SSP*) and southern (*SG*, *SM*, and *SLL*) regions of the Baja peninsula

were significantly differentiated from one another (F_{st} = 0.27–0.56, $p < .01$), whereas negligible structure was found within each of these regions (Table S2). By contrast, in *Brahea*, all pairwise F_{st} comparisons among sierras within the peninsula were statistically significant (Table S3).

To uncover the main genetic clusters present within *Washingtonia* and *Brahea*, we used admixture estimation and individual clustering within sNMF as well as phylogenetic network inference within SPLITSTREE (see Section 2 for details). Both of these approaches resolved clear groupings and there was general agreement between them on the strength and pattern of population structure. In the case of *Washingtonia*, four main distinct genetic clusters were recovered (Figure 3; Figure S3) corresponding to: (1) the Mexican mainland (*SON*), (2) the southern Baja peninsula (*SLL*, *SM*, and *SG*), (3) the northern Baja peninsula (*SSP*, *SFSL*, and *CAT*), and (4) *Washingtonia* from Sierra Juarez (*SJ*). For *Brahea*, there were some differences in the results depending on the analytical approach used. Specifically, sNMF uncovered three distinct genetic clusters (Figure 4a) representing (1) the southern Baja peninsula (*SLL*, *SM*, and *SSP*), (2) the central Baja peninsula (*SSF*, *SLI*, and *SA*), and (3) the northern Baja peninsula (*CAT*, *SSPM*, and *SJ*). By contrast, the SPLITSTREE analysis (Figure 4b) was more sensitive to the phylogenetic structuring of the peninsular *Brahea* palms and clearly partitioned the samples into nine genetic clusters, each corresponding to a different sierra. The only exceptions were two samples from *SSF* that clustered together with palms from *SL*, two individuals that were misplaced between *SSF* and *SM*, and one individual between *SSP* and *SL*. Intriguingly, individual cluster membership plots (Figure S3) also highlighted the presence of individual palms sampled from the northern sierras that exhibited clear genetic ancestry in the southern sierras. Specifically, four *Washingtonia* palms sampled from two of the northern sierras (*CAT* and *SFSL*) had genotypes indicative of ancestry in the southern
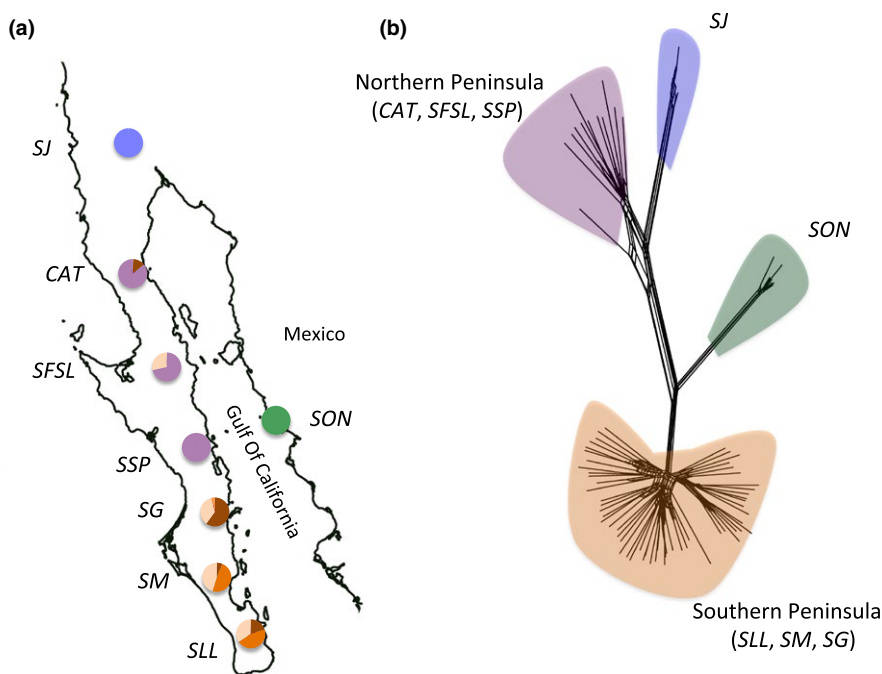
FIGURE 3 Population structure of *Washingtonia* palms. Panel (a) shows the results of cluster analysis within sNMF, with pie charts indicating the geographic distribution of six inferred genetic clusters (each coded by a different color). Panel (b) shows a phylogenetic network generated by SPLITSTREE
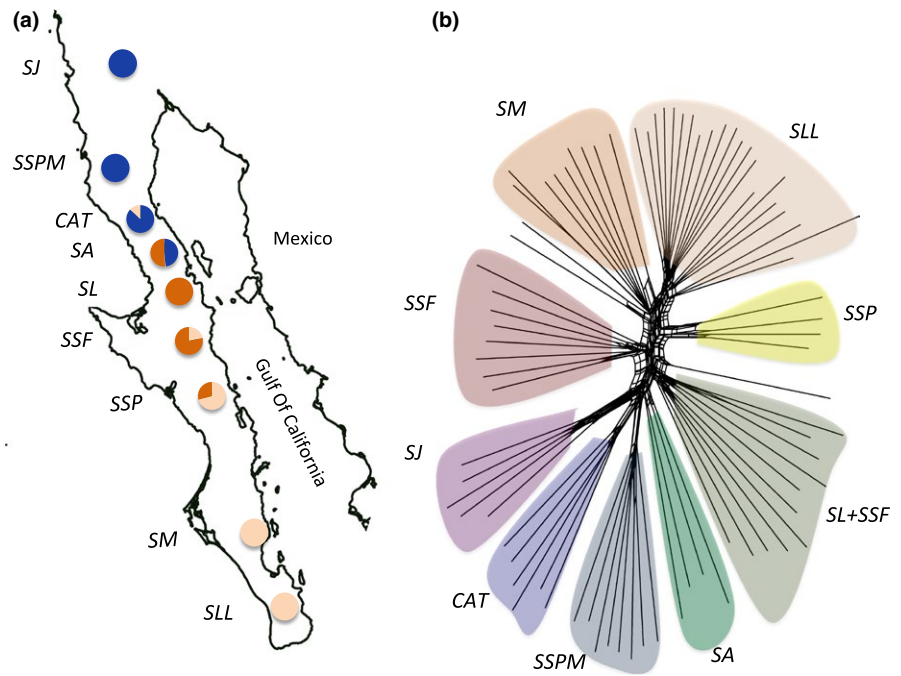
FIGURE 4 Population structure of *Brahea* palms. Panel (a) shows the results of cluster analysis within sNMF, with pie charts indicating the geographic distribution of three inferred genetic clusters (each coded by a different color). Panel (b) shows a phylogenetic network generated by SPLITSTREE

sierras *SLL*, *SM*, and *SG*, whereas two *Brahea* individuals sampled from *CAT* had genetic ancestry consistent with *SM* and *SLL*.

3.4 | Population splits and migration modeling

To investigate the potential cause of some individuals being mis-assigned to their populations of origin, we modeled population divergence with migration within TREEMIX. The resulting maximum-likelihood (ML) tree for *Washingtonia* was concordant with the previous results, revealing deep divergence between the mainland and peninsula populations and partitioning of the latter into southern (*SLL*, *SM*, and *SG*) and northern (*SSP*, *SFSL*, *CAT*) groups (Figure 5a). After sequentially testing for between one and ten discrete migration events, we found that the increase in likelihood beyond two migration events was close to zero (Figure S4) and stepwise comparisons of log-likelihood values lost significance between two and three events (likelihood ratio test, *p* > .05). This indicates that the most likely number of migration events among the sierras was two. Exploring this scenario further, we found evidence for unidirectional long-distance migration spanning around 450 km, from *SG* into *CAT* and from *SM* into *SFSL* (Figure 5a). These migration events were strongly supported by three-population tests (Table S4).

For the *Brahea* palms, TREEMIX analysis supported the partitioning of the peninsular samples into two major groups comprising the northernmost sierras (*SJ* and *SSPM*) and the rest of the peninsula (Figure 5b). The latter in turn was partitioned into the northern sierras (*SL*, *SA*, and *CAT*) and the southern and mid-peninsula sierras (*SLL*, *SM*, *SSP*, and *SSF*). Inferred migration events for *Brahea* pointed toward the possible migration of palms between the southern sierra of *SM* and the northern sierra of *CAT* (Figure 5b). However, the overall pattern of increasing log likelihood with the number of migration events was less pronounced than in *Washingtonia* (Figure S4) and

none of the migration events were supported by three-population tests (*Z* scores < −1.96).

3.5 | Isolation by distance and ecological divergence

Mantel tests revealed strong positive correlations between genetic and geographic distance for both *Washingtonia* and *Brahea* (Mantel's *r* = .685, *p* < .0001 and *r* = .609, *p* < .0001 respectively, Table 1). Furthermore, Partial Mantel tests revealed significant associations between genetic distance and environmental variables after controlling for geographic distance. Specifically, the mean temperature of the coldest quarter correlated significantly with genetic distance in *Washingtonia* (Partial Mantel test, *r* = .267, *p* < .0001), while the amount of precipitation in the driest quarter correlated significantly with genetic distance in *Brahea* (Partial Mantel test, *r* = .09, *p* = .019).
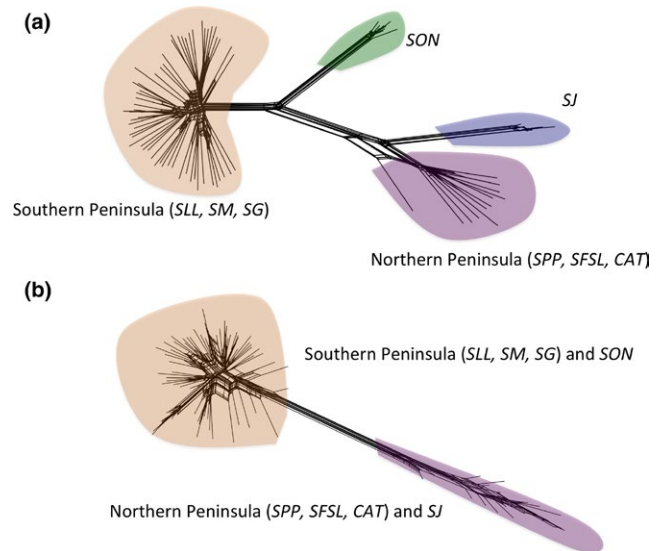
3.6 | Detection of outlier loci associated with environmental variables

In order to investigate the genomic basis of the associations described above, we used two complimentary approaches to test for signatures of local adaptation in *Washingtonia* and *Brahea*. First, latent factor mixed models LFMMs were used to detect loci exhibiting unusual associations with mean temperature of the coldest quarter in *Washingtonia* and precipitation of the driest quarter in *Brahea*. This resulted in the identification of 80 SNPs in *Washingtonia* and 51 SNPs in *Brahea* after table-wide correction of the corresponding *p*-values for the false discovery rate. Second, we used a multivariate approach implemented in SAMβADA to test for genotype by environment associations. This approach identified 18 significant associations in *Washingtonia* and none in *Brahea*.
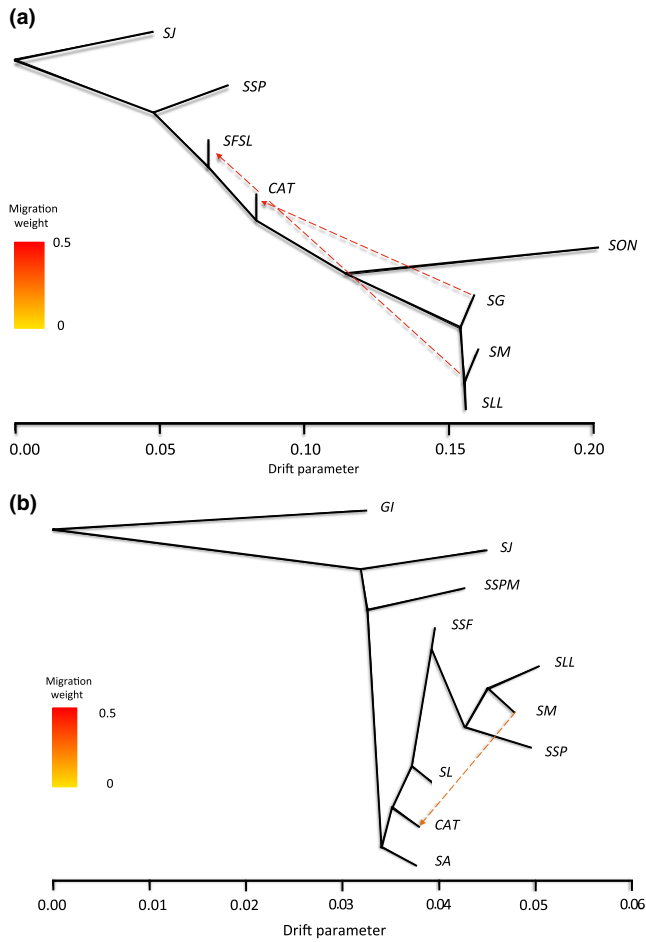
FIGURE 5 Maximum-likelihood trees depicting patterns of genetic divergence among (a) *Washingtonia* palms grouped into eight populations and (b) *Brahea* palms grouped into ten populations (see Section 2 for details). Inferred migration events are indicated by dashed lines, with the direction of gene flow indicated by arrows and color intensity reflecting the intensity of gene flow

FIGURE 6 Phylogenetic networks constructed separately for *Washingtonia* using (a) neutral loci; (b) outlier loci

whereas the outlier loci only resolved two groups, corresponding to the southern Baja peninsula combined with *SON* and the northern Baja peninsula combined with *SJ* (Figure 6b). Phylogenetic trees constructed from the same number of randomly selected neutral loci as there are outliers also recovered four populations (Figure S5), suggesting that the contrasting topologies recovered by the outlier and neutral loci are unlikely to be caused by differences in resolving power. By comparison, the main difference between the neutral and outlier trees for *Brahea* was that the former resolved individual sierras (Figure S6a), whereas the latter did not (Figure S6b). This appears to be related to genetic resolution as phylogenetic trees based on the same number of randomly selected neutral loci as outliers also failed to clearly resolve the sierras (Figure S7).

4 | DISCUSSION

Comparative studies can provide valuable insights into processes that shape population genetic structure and thereby help to improve our understanding of how organisms may respond to ongoing environmental change. We therefore used GBS both to resolve taxonomic uncertainties and to characterize patterns of population

Finally, we asked whether the outlier loci identified by LFMM and SAMβADA (total *n* = 97 for *Washingtonia* and 51 for *Brahea*) resolve contrasting phylogenies to the neutral loci (defined as those loci that were not identified by either program, *n* = 1,365 for *Washingtonia* and 1,999 for *Brahea*). Analysis within SPLITSTREE revealed a striking pattern for *Washingtonia* in which the neutral loci resolved four groups corresponding to the southern Baja peninsula (*SLL*, *SM*, and *SG*), the northern Baja peninsula (*SSP*, *SFSL*, and *CAT*), *SJ* and *SON* (Figure 6a),

	Washingtonia		*Brahea*	
	Mantel	**Partial Mantel**	**Mantel**	**Partial Mantel**
Geographic distance	0.685***	NA	0.609***	NA
Aridity	−0.059	−0.180	0.068*	−0.230
Mean temperature of the coldest quarter	0.620***	0.267***	0.248***	−0.160
Precipitation of the driest quarter	0.177***	0.010	0.127**	0.09*

TABLE 1 Mantel and partial Mantel tests summarizing relationships (*r* and associated *p* values) between genetic distance, geographic distance, and climate variables in *Washingtonia* and *Brahea* P-values * P < 0.05; ** P < 0.01; *** P < 0.001

structure in two closely related North American palm genera, *Washingtonia* and *Brahea*, on the Baja California peninsula, adjacent Mexican mainland and Guadalupe Island. Bayesian phylogenetic analysis supported the classification of *B. edulis* as a distinct species, but this was not the case for *W. filifera–W. robusta* and *B. armata–B. brandegeei*. Furthermore, population genetic analyses clustered the peninsular *Washingtonia* palms into two populations corresponding to the northern and southern peninsula, whereas in *Brahea* every sierra could be genetically distinguished. We also detected a relatively strong influence of ecologically mediated divergence in *Washingtonia* palms, with outlier loci correlated to temperature resolving a markedly different phylogenetic tree to neutral loci. Finally, we found evidence for two unidirectional long-distance migration events in *Washingtonia*, in line with the previous suggestion that human-mediated dispersal could have been disproportionately important in this genus (McClenaghan & Beauchamp, 1986; Minnich et al., 2011). None of these patterns could previously be detected using either allozymes (McClenaghan & Beauchamp, 1986) or classical plastid and nuclear markers (Klimova et al., 2017), suggesting that GBS and related approaches represent powerful tools for uncovering ecologically relevant population subdivision.

4.1 | Phylogenetic relationships

The first aim of our study was to resolve phylogenetic relationships between and within *Washingtonia* and *Brahea* palms sampled from the Baja California peninsula and adjacent areas. As expected, the two genera were found to be deeply divergent, consistent with their having separated from one another at least 25–35 million years ago (Bacon et al., 2012; Baker & Couvreur, 2013. However, mixed support was found for currently recognized taxonomic relationships within each genus (Felger & Joyal, 1999; Henderson et al., 1995; Minnich et al., 2011). Starting with *Washingtonia*, Bayesian phylogenetic reconstruction uncovered two main lineages, the first corresponding to *W. robusta* from the north of the Baja peninsula together with *W. filifera*, and the second corresponding to *W. robusta* from the southern Baja peninsula and the Mexican mainland. Furthermore, the magnitude of divergence between *W. filifera* and *W. robusta* from the northern Baja peninsula was lower than that found between *W. robusta* from the northern Baja peninsula and the Mexican mainland (i.e., it was within the range found within a single species). Our data therefore lead us to the conclusion that *W. filifera* is more likely to represent the northernmost population of *W. robusta* than a separate species. This is in line with a previous study by Bacon et al. (2012), who found no differences between *W. filifera* and *W. robusta* at three plastid and three nuclear genes, and is also consistent with an apparent lack of reproductive isolation between these palms, as hybridization is common in cultivation (Hodel, 2014). Furthermore, a highly detailed morphological study of *Washingtonia* palms from 17 sites on the peninsula recently found no clear support for two distinct species based on 11 morphological characteristics but rather suggested the presence of a latitudinal morphological cline (Villanueva-Almanza & Ezcurra, 2017).

Bayesian phylogenetic reconstruction of *Brahea* provided evidence in support of the species status of Guadalupe Island palms (*B. edulis*) although this was to some extent expected given the geological origin and geographic isolation of Guadalupe Island (Aleixandre, Hernandez-Montoya, & Mila, 2013; Karhu, Vogl, Moran, Bell, & Savolainen, 2006; Klimova et al., 2017). Nonetheless, our data did not support the recognition of two separate species, *B. armata* in the northern peninsula and *B. brandegeei* in the southern peninsula. Again, this is consistent with a previous genetic study based on chloroplast and nuclear sequences, which also failed to separate the peninsular *Brahea* into two monophyletic groups (Klimova et al., 2017). One reason for this could be that the original taxonomy was based on relatively subtle morphological differences such as leaf coloration and inflorescence architecture (Felger & Joyal, 1999; Henderson et al., 1995) and these traits could potentially show plastic variation among populations in response to the prevailing environmental conditions (Roncal, Henderson, Borchsenius, Cardoso, & Balslev, 2012). One way to test this hypothesis would be to use reciprocal transplant or common garden experiments.

4.2 | Patterns of population genetic structure

Washingtonia and *Brahea* palms from the Baja California peninsula provide a unique opportunity to explore the contributions of multiple potential drivers of population structure within an unusually heterogenous natural setting. However, a recent study based on nuclear and chloroplast genes was unable to recover sufficient levels of polymorphism to provide insights at the population level (Klimova et al., 2017). Fortunately, GBS allowed us to genotype over 25,000 SNPs, which after highly stringent filtering to retain only polymorphic, unlinked loci with a small proportion of missing data, left us with a total of 1,462 and 2,050 genome-wide distributed SNPs in *Washingtonia* and *Brahea* respectively. These data allowed us to uncover contrasting and in some cases unexpected patterns of genomewide differentiation, with *Washingtonia* populations showing a clear north–south split, whereas in *Brahea* each individual sierra could be resolved.

We found evidence for four main groups of *Washingtonia* palms. As might be expected given the degree of geographic isolation, palms from the Mexican mainland and *SJ* formed separate clusters, but we did not anticipate finding two distinct palm lineages among sierras of the Baja California peninsula that were roughly evenly spaced along a latitudinal cline. Many plant and animal species on the peninsula show a similar north–south divide that has been linked to the temporary formation of a mid-peninsula seaway around a million years ago (Lindell et al., 2006; Riddle et al., 2000). However, this is not strictly consistent with our results, as palms from *SSP* show membership to the northern cluster but are located to the south of where the seaway is believed to have been located. The reasons for this are not entirely clear. It is possible but fairly unlikely that the true location of the seaway was actually further south than is currently believed. Alternatively, *Washingtonia* could have been locally extirpated at *SSP* and subsequently recolonized from *SFSL* after the closure of the seaway. This explanation is plausible both because

ecological niche modeling has shown major shifts in the distribution of suitable habitat over the past 100,000 years and *SSP* is currently represented by a single isolated oasis situated on the margins of a large area of unsuitable habitat (see Figure 6 in Klimova et al., 2017).

Our results for *Brahea* are in many respects more readily explained (Klimova et al., 2017). SPLITSTREE partitioned the palms into nine genetic clusters, each corresponding to a different sierra, while sNMF detected three main groups whose frequencies followed a clear cline along the peninsula, consistent with a significant pattern of isolation by distance (Mantel's r = .609, p < .001). This probably reflects the fact that the seeds of *Brahea* palms are primarily dispersed by water pulses that wash them short distances along canyons (Wehncke & López-Medellín, 2014; Wehncke et al., 2009). By contrast, *Washingtonia* palms have edible fruit that are eaten by birds and small mammals and which were apparently also used by indigenous people (Cornett, 2008; Luna, 2012), thereby facilitating the dispersal of intact seeds between adjacent oases.

Another factor that could have contributed toward differences between *Washingtonia* and *Brahea* is demographic history. In particular, historical bottlenecks can lead to strong genetic drift and thereby contribute toward both the pattern and strength of population genetic structure (Futuyma & Kirkpatrick, 2017). Unfortunately, however, GBS and related approaches are not well suited to demographic reconstruction as tests for bottlenecks and population expansion are extremely sensitive to MAF thresholds and other aspects of the bioinformatic pipeline(s) used (Shafer et al., 2017). Nevertheless, we have little reason to believe that the two genera experienced markedly different recent demographic histories, as ecological niche modeling has shown that both *Washingtonia* and *Brahea* would have been locally restricted to similar areas during the LGM (about 22,000 years ago) and subsequently re-established themselves across most of the peninsula (Klimova et al., 2017).

4.3 | Ecologically mediated divergence

There is growing interest and empirical support for the notion that strong divergent natural selection can drive genomic divergence, ultimately leading in some cases to reproductive isolation and speciation (Beheregaray, Cooke, Chao, & Landguth, 2015; Lexer et al., 2014; Sork et al., 2016). The palms of Baja California are interesting in this regard because they reside at the extreme northern distributional limits of the mostly tropical Arecaceae and are therefore subjected to unusually dry, cold and generally suboptimal conditions (Hampe & Jump, 2011; Woolbright, Whitham, Gehring, Allan, & Beiley, 2014). Under such conditions, local adaptation can be a particularly important force in shaping patterns of divergence across the genome (Pannell & Fields, 2014; Savolainen, Lascoux, & Merilä, 2013). Our results are consistent with this notion and suggest that local adaptation may have contributed toward the population structure of *Washingtonia* in particular.

We originally hypothesized that the influence of ecologically mediated selection should be strongest on *Washingtonia* palms due to their stricter ecological requirements and relatively restricted

geographic distribution (Minnich et al., 2011). In line with this, we detected significant associations between environmental variables and genetic distance in both genera, but these were stronger in *Washingtonia*. We also found that genetic distance was correlated to the mean temperature of the coldest quarter in *Washingtonia*, whereas in *Brahea* genetic distance was associated with precipitation of the driest quarter. Taken at face value, this difference would imply that, even though many of the *Washingtonia* and *Brahea* palms were sampled from the same oases, natural selection has influenced the two genera in different ways.

To explore this further, we used two different approaches to test for loci showing unusual associations with environmental variables. In both cases, we attempted to minimize the occurrence of false positives, either by controlling for the underlying population structure by introducing latent factors (in LEA) or using multivariate logistic regression (in SAMβADA), which reduces the occurrence of spurious genotype by environment associations (Stucki et al., 2016). These approaches identified different sized and largely non-overlapping subsets of loci, but this is consistent with previous studies and reflects differences in the underlying methodologies and assumptions (Benestan et al., 2016; Feng, Jiang, & Fan, 2016; Nadeau, Meirmans, Aitken, Ritland, & Isabel, 2016). To capture as many outliers as possible, we therefore pooled all of the loci flagged by at least one approach and classified the remaining loci as neutral. Constructing phylogenetic trees separately for these two classes of locus revealed a clear difference in *Washingtonia*, with the neutral loci resolving four groups but the outlier loci only two groups. This finding is reminiscent of similar studies that likewise resolved different trees based on neutral and outlier loci (Funk et al., 2016; Keller et al., 2013; Matala, Ackerman, Campbel, & Narum, 2014). Such a pattern could be considered a footprint of selection as *SJ* and the northern peninsular sierras have diverged at the genomic background, while stabilizing selection appears to have resulted in very similar genotypes at the outlier loci.

The equivalent results for *Brahea* were less clear cut. Within the Baja peninsula, the outlier loci failed to resolve the individual sierras. However, phylogenetic trees based on the same number of randomly selected neutral loci also grouped the sierras together, suggesting that genetic differences between the sierras may be too weak to be resolved by a relatively small subset of SNPs. Either this could reflect weaker selective pressures on *Brahea* palms or alternatively the association between genetic distance and precipitation of the driest quarter in *Brahea* could be a type I error. Further insights into this and related questions could be gained from detailed physiological studies aiming to establish more clearly how the two palm genera respond to climatic extremes.

4.4 | Patterns of long-distance dispersal

A further unexpected pattern was revealed by cluster analyses of the GBS data. Overall, the majority of individuals were confidently clustered to their respective geographic groups and levels of admixture between the identified populations were low. However,

several palms sampled from the northern sierras had genotypes that were clearly consistent with ancestry in the southern sierras, a pattern that is strongly suggestive of recent long-distance dispersal. To explore this further, we used TREEMIX to infer the most likely number of migration events in both palm species. The results for *Washingtonia* were clear cut, with strong support being found for two distinct south to north migration events spanning around 450 km. By contrast, although one migration event was inferred for *Brahea*, this was not supported by three-population tests and should therefore be treated as putative at best.

It is unlikely that these patterns could have resulted from differences in the natural dispersal abilities of the two palm genera because, even though *Brahea* is more dispersal limited, *Washingtonia* cannot be naturally dispersed over more than a few tens of kilometers. Furthermore, if natural agents such as birds or mammals were involved, one would not necessarily expect to find a bias in the direction of migration from north to south, which appears to be the case for the long-distance migration events inferred in *Washingtonia*. However, as palms have a long history of relationship with humans and have been extensively used as a source of food, construction materials and more recently as ornaments in cities and gardens, several authors have speculated that humans may have been involved in spreading palms on the Baja peninsula (Cornett, 2008; Felger & Joyal, 1999; Levis et al., 2017; McClenaghan & Beauchamp, 1986; Minnich et al., 2011). Furthermore, the fruits of *Washingtonia* were extensively used as a food source by native people (Cornett 1987; Felger & Moser, 1985; Felger & Joyal, 1999), whereas *Brahea* fruit were less appreciated and have even been referred to as "useless" (Minnich et al., 2011), leading some authors to suggest that the indigenous people of the Baja peninsula may have dispersed *Washingtonia* but not *Brahea* prior to European contact (Cornett, 2008; Minnich et al., 2011). Our results are consistent with this hypothesis as long-distance migration events were only inferred unequivocally in *Washingtonia*. However, the two northern oases containing *Washingtonia* palms of southern ancestry are also both sites of colonial Spanish missions where agriculture and trade would have been especially well developed (Minnich et al., 2011). Consequently, it is not inconceivable that these genetic introductions could have occurred more recently, which is supported by the observation that the palms in question exhibited negligible admixture. Although there are clear precedents for human-mediated dispersal of palms (Aschmann, 1957; Kondo et al., 2012; Rivera et al., 2013), we prefer not to speculate further at this point without additional archeological evidence or more detailed genetic data that would allow us to reliably date these migration events.

4.5 | Limitations of the study

The recent development of cost-effective methods for obtaining high-quality genome-scale data has stimulated growing interest in the genomic basis of ecological divergence. By greatly increasing genotyping coverage, approaches like GBS have made it possible to identify genomic regions and in some cases specific loci responsible for adaptive differences among populations (Savolainen et al., 2013). Nonetheless, a number of caveats need to be taken into account. For example, population structure, demographic history and the quality of the environmental data and biases caused by the genetic markers themselves can all lead to false-positive results in outlier scans (Hoban et al., 2016).

First of all, disentangling IBE from neutral patterns of genetic variation can be challenging (Wang & Bradburd, 2014) because IBD can produce patterns similar to IBE when geography is correlated with environmental variation (Meirmans, 2012; Nadeau et al., 2016). We attempted to reduce this potential source of bias by analyzing only ecological variables that were significantly associated with genetic distance after having controlled for geographic distance. We further minimized the occurrence of false positives by controlling for population structure and demographic history by introducing latent factors into the LEA analysis and by implementing a multivariate logistic regression approach in SAMβADA.

Second, the accurate detection of locally adapted loci through genotype–environment associations also depends on accurate measures of different aspects of the ecological landscape. Low-resolution environmental data may reduce the accuracy of results even if the selective environment is known (Hoban et al., 2016), so the resolution of the environmental data must be fine enough to adequately characterize each sampling locality. For this reason, we cross-referenced fine resolution (~1 km scale) environmental data with GPS coordinates collected by ourselves at each oasis.

Third, in common with virtually all genotyping approaches including restriction enzyme-based methods, GBS suffers from a number of potential sources of genotyping error (Andrews et al., 2016; Hoban et al., 2016). One of the most important of these is the presence of null alleles. These occur when a polymorphism within the restriction enzyme recognition site results in failure to cut the genomic DNA at that location. Alleles lacking the complete recognition site are not sequenced, which results in individuals who are heterozygous for the null allele appearing as homozygotes. Null alleles can downwardly bias estimates of genetic diversity, overestimate F_{st}, and result in an increase in false positives in F_{st} outlier tests (Andrews et al., 2016). To overcome this limitation, we used two approaches that detect associations between genetic and environmental distances without estimating F_{st} (Frichot & François, 2015; Stucki et al., 2016).

A related issue is that stochastic processes during PCR can cause one allele to amplify more readily than the other at a given locus. This can lead to downstream genotyping errors, as heterozygotes can appear as homozygotes, or alleles containing PCR errors can be interpreted as true alleles. However, at least in theory, PCR should not systematically favor one allele over another at a given locus, and therefore, parameters estimated from a large number of loci are unlikely to be substantially biased (Andrews et al., 2016). Furthermore, although reduced genome representation approaches such as GBS provide far greater genomic coverage than their predecessors such as microsatellites or amplified fragment length polymorphisms,

it is nevertheless important to bear in mind that only a fraction of the genome can be screened, which may result in some important targets of selection being missed (Narum, Buerkle, Davey, Miller, & Hohenlohe, 2013). However, we were more interested in broad patterns rather than in the nature of the specific loci under selection, and the contrasting topographies of neutral versus selected loci in _Washingtonia_ suggest that at least for this genus our study was successful at detecting genomic regions influenced by selection. Classical approaches like common garden or reciprocal transplant experiments would provide a suitable basis for future confirmatory studies (Anderson et al., 2010; Savolainen et al., 2013).

5 | CONCLUSIONS

Comparative studies can shed light on species-specific properties that may influence dispersal and ecological divergence, while GBS offers an unprecedentedly detailed window on genome-wide patterns of differentiation. By combining these approaches in iconic North American palms, we were able to uncover markedly different patterns of population structure in _Washingtonia_ and _Brahea_, reveal associations between genetic distance and climatic variables, identify subsets of loci that appear to be under divergent ecologically mediated selection, and identify long-distance migration events consistent with human-mediated dispersal. None of these patterns could be detected with classical molecular markers, indicating the promise of approaches like GBS to dissect apart the contributions of different processes toward genome-wide patterns of divergence.

CONFLICT OF INTEREST

None declared.

AUTHOR CONTRIBUTIONS

AK, AOR, and JIH designed the research; AK collected samples; AK performed molecular laboratory work; AK, DLJV, and JIH analyzed the molecular data; AK and JIH wrote the manuscript; AOR provided funding; all authors approved of the final version of the manuscript.

DATA ACCESSIBILITY

All of the genotypes together with environmental data for the sampling locations are available from Dryad https://doi.org/10.5061/dryad.5vk6219.

ORCID

Anastasia Klimova (iD) http://orcid.org/0000-0002-1502-2910

Alfredo Ortega-Rubio (iD) http://orcid.org/0000-0002-6365-287X

Joseph I. Hoffman (iD) http://orcid.org/0000-0001-5895-8949

REFERENCES

Aleixandre, P., Hernandez-Montoya, J., & Mila, B. (2013). Speciation on Oceanic Islands: Rapid adaptive divergence vs. cryptic speciation in a Guadalupe Island Songbird (Aves: Junco). _PLoS ONE_, _8_, e63242. https://doi.org/10.1371/journal.pone.0063242

Anderson, J. T., Willis, J. H., & Mitchell-Olds, T. (2011). Evolutionary genetics of plant adaptation. _Trends in Genetics_, _27_, 258–266. https://doi.org/10.1016/j.tig.2011.04.001

Andrews, K. R., Good, J. M., Miller, M. R., Luikart, G., & Hohenlohe, P. A. (2016). Harnessing the power of RADseq for ecological and evolutionary genomics. _Nature Reviews Genetics_, _17_, 81–92. https://doi.org/10.1038/nrg.2015.28

Aschmann, H. (1957). The introduction of the date palm to Baja California. _Economic Botany_, _11_, 174–177. https://doi.org/10.1007/BF02985306

Avise, J. C. (2000). _Phylogeography: The history and formation of species_. Cambridge, UK: Harvard University Press.

Avise, J. C., Bowen, B. W., & Ayala, F. J. (2016). In the light of evolution X: Comparative phylogeography. _Proceedings of the National Academy of Sciences of the United States of America_, _113_, 7957–7961. https://doi.org/10.1073/pnas.1604338113

Axelrod, D. I. (1958). Evolution of the Madro-Tertiary geoflora. _Botanical Review_, _24_, 433–509. https://doi.org/10.1007/BF02872570

Bacon, C. D., Baker, W. J., & Simmons, M. P. (2012). Miocene dispersal drives island radiations in the palm tribe Trachycarpeae (Arecaceae). _Systematic Biology_, _61_, 426–442. https://doi.org/10.1093/sysbio/syr123

Bacon, C. D., Look, S. L., Gutierrez-Pinto, N., Antonelli, A., Tan, H. T. W., Kumar, P. P., … Baker, W. J. (2016). Species limits, geographical distribution and genetic diversity in Johannesteijsmannia (Arecaceae). _Botanical Journal of the Linnean Society_, _182_, 318–347. https://doi.org/10.1111/boj.12470

Baker, W. J., & Couvreur, T. L. P. (2013). Global biogeography and diversification of palms sheds light on the evolution of tropical lineages. Historical biogeography. _Journal of Biogeography_, _40_, 274–285. https://doi.org/10.1111/j.1365-2699.2012.02794.x

Batiza, R. (1977). Petrology and chemistry of Guadalupe Island: An alkalic seamount on a fossil ridge crest. _Geology_, _5_, 760–764. https://doi.org/10.1130/0091-7613(1977)5<760:PACOGI>2.0.CO;2

Beheregaray, L. B., Cooke, G., Chao, N. L., & Landguth, E. L. (2015). Ecological speciation in the tropics: Insights from comparative genetic studies in Amazonia. _Frontiers in Genetics_, _5_, 1–19.

Benestan, L., Quinn, B. K., Maaroufi, H., Laporte, M., Clark, F. K., Greenwood, S. J., … Bernatchez, L. (2016). Seascape genomics provides evidence for thermal adaptation and current-mediated population structure in

American lobster (*Homarus americanus*). *Molecular Ecology*, 25, 5073–5092. https://doi.org/10.1111/mec.13811

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B*, 57, 289–300.

Bernardi, G., Ruiz-Campos, G., & Camarena-Rosales, F. (2007). Genetic isolation and evolutionary history of oases populations of the Baja California killifish, *Fundulus lima*. *Conservation Genetics*, 8, 547–554. https://doi.org/10.1007/s10592-006-9190-1

Bouckaert, R. (2010). DensiTree: Making sense of sets of phylogenetic trees. *Bioinformatics*, 26, 1372–1373. https://doi.org/10.1093/bioinformatics/btq110

Bouckaert, R., Heled, J., Kuhnert, D., Vaughan, T. G., Wu, C. H., & Xie, D. (2014). BEAST2: A software platform for Bayesian evolutionary analysis. *PLoS Computational Biology*, 10, e1003537. https://doi.org/10.1371/journal.pcbi.1003537

Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., & Buckler, E. S. (2007). TASSEL: Software for association mapping of complex traits in diverse samples. *Bioinformatics*, 23, 2633–2635. https://doi.org/10.1093/bioinformatics/btm308

Bryant, D., Bouckaert, R., Felsenstein, J., Rosenberg, N. A., & Choudhury, A. (2012). Inferring species trees directly from biallelic genetic markers: Bypassing gene trees in a full coalescent analysis. *Molecular Biology and Evolution*, 29, 1917–1932. https://doi.org/10.1093/molbev/mss086

Bryant, D., & Moulton, V. (2004). Neighbor-Net: An agglomerative method for the construction of phylogenetic networks. *Molecular Biology and Evolution*, 21, 255–265.

Chang, C. C., Chow, C. C., Tellier, L. C. A. M., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: Rising to the challenge of larger and richer datasets. *GigaScience*, 4, 7. https://doi.org/10.1186/s13742-015-0047-8

Cornett, J. W. (1985a). Reading the fan palms. *Natural History*, 94, 64–73.

Cornett, J. W. (1985b). Nutritional value of desert fan palm fruits. *Principes (USA)*, 31, 159–161.

Cornett, J. W. (1987). Three palm species at Cataviña. *Principes*, 31, 12–13.

Cornett, J. W. (2008). The desert fan palm oasis. In L. E. Stevens, & V. J. Meretsky (Eds.), *Aridland springs in North America: Ecology and conservation* (pp. 158–184). Tucson, AZ: University of Arizona Press.

Cornett, J. W., Glen, T., & Stewart, J. M. (1986). The largest desert fan palm oases. *Principes (USA)*, 30, 82–84.

Couvreur, T. L. P., & Baker, W. J. (2013). Tropical rain forest evolution: Palms as a model group. *BMC Biology*, 11, 48. https://doi.org/10.1186/1741-7007-11-48

Couvreur, T. L. P., Forest, F., & Baker, W. J. (2011). Origin and global diversification patterns of tropical rain forests: Inferences from a complete genus-level phylogeny of palms. *BMC Biology*, 9, 44. https://doi.org/10.1186/1741-7007-9-44

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., & 1000 Genomes Project Analysis Group (2011). The variant call format and VCFtools. *Bioinformatics*, 27, 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

De Donato, M., Peters, S. O., Mitchell, S. E., Hussain, T., & Imumorin, I. G. (2013). Genotyping-by-sequencing (GBS): A novel, efficient and cost-effective genotyping method for cattle using next-generation sequencing. *PLoS ONE*, 8, e62137. https://doi.org/10.1371/journal.pone.0062137

Dolby, G. A., Bennett, S. E. K., Lira-Noriega, A., Wilder, B. T., & Munguia-Vega, A. (2015). Assessing the geological and climatic forcing of biodiversity and evolution surrounding the Gulf of California. *Journal of the Southwest*, 57, 391–455. https://doi.org/10.1353/jsw.2015.0005

Dransfield, J., Uhl, N. W., Asmussen, C. B., Baker, W. J., Harley, M., & Lewis, C. (2008). *Genera palmarum: The evolution and classification of palms*. Kew, UK: Royal Botanic Gardens.

Eiserhardt, W. L., Svenning, J. C., Kissling, W. D., & Balslev, H. (2011). Geographical ecology of the palms (Arecaceae): Determinants of diversity and distributions across spatial scales. *Annals of Botany*, 108, 1391–1416. https://doi.org/10.1093/aob/mcr146

Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., & Mitchell, S. E. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE*, 6, e19379. https://doi.org/10.1371/journal.pone.0019379

Ersts, P. J. (2016). *Geographic Distance Matrix Generator (version 1.2.3)*. American Museum of Natural History, Center for Biodiversity and Conservation. Retrieved from http://biodiversityinformatics.amnh.org/open_source/gdmg (accessed July 2016).

Felger, R. S., & Broyles, B. (2007). *Dry borders, great natural reserves of the Sonoran Desert*. Salt Lake City, UT: The University of Utah Press.

Felger, R. S., Johnson, M. S., & Wilson, M. F. (2001). *Trees of Sonora, Mexico*. New York, NY: Oxford University Press.

Felger, R. S., & Joyal, E. (1999). The palms (Arecaceae) of Sonora, Mexico. *Aliso*, 18, 1–18. https://doi.org/10.5642/aliso

Felger, R.S., & Moser, M.B. (1985). *People of the desert and sea: Ethnobotany of the Seri Indians*. Tucson, AZ: University of Arizona

Feng, X. J., Jiang, G. F., & Fan, Z. (2016). Identification of outlier sinagenomic scan for selection along environmental gradients in the bamboo locust, Ceracris kiangsu. *Scientific Reports*, 5, 13758.

Fischer, M. C., Rellstab, C., Leuzinger, M., Roumet, M., Gugerli, F., Shimizu, K. K., ... Widmer, A. (2017). Estimating genomic diversity and population differentiation—An empirical comparison of microsatellite and SNP variation in Arabidopsis halleri. *BMC Genomics*, 18, 69. https://doi.org/10.1186/s12864-016-3459-7

Franco-Vizcaíno, E., López-Beltrán, A. C., & Salazar-Ceseña, M. (2007). Water relations and community composition in three blue fan palm oases across the Californian-Sonoran biome transition. *The Southwestern Naturalist*, 52, 191–200. https://doi.org/10.1894/0038-4909(2007)52[191:WRACCI]2.0.CO;2

Frankham, R., Ballou, J. D., & Briscoe, D. A. (2002). *Introduction to conservation genetics*. Cambridge, UK: Cambridge University Press. https://doi.org/10.1017/CBO9780511808999

Frichot, E., & François, O. (2015). LEA: An R package for landscape and ecological association studies. *Methods in Ecology and Evolution*, 6, 925–929. https://doi.org/10.1111/2041-210X.12382

Frichot, E., Mathieu, F., Trouillon, T., Bouchard, G., & François, O. (2014). Fast and efficient estimation of individual ancestry coefficients. *Genetics*, 196, 973–983. https://doi.org/10.1534/genetics.113.160572

Frichot, E., Schoville, S. D., Bouchard, G., & François, O. (2013). Testing for associations between loci and environmental gradients using latent factor mixed models. *Molecular Biology and Evolution*, 30, 1687–1699. https://doi.org/10.1093/molbev/mst063

Funk, W. C., Lovich, R. E., Hohenlohe, P. A., Hofman, C. A., Morrison, S. A., Sillett, T. S., ... Andelt, W. F. (2016). Adaptive divergence despite strong genetic drift: Genomic analysis of the evolutionary mechanisms causing genetic differentiation in the island fox (*Urocyon littoralis*). *Molecular Ecology*, 25, 2176–2194. https://doi.org/10.1111/mec.13605

Futuyma, D. J., & Kirkpatrick, M. (2017). *Evolution*, Fourth edn. Oxford, UK: Sinauer Associates is an imprint of Oxford University Press.

Garcillán, P. P., Vega, E., & Martorell, C. (2012). The Brahea edulis palm forest in Guadalupe Island: A North American fog oasis? *Revista Chilena de Historia Natural*, 85, 1137–1145.

Garrick, R. C., Rowell, D. M., Simmons, C. S., Hillis, D. M., & Sunnucks, P. (2008). Fine-scale phylogeographic congruence despite demographic incongruence in two low-mobility saproxylic springtails. *Evolution*, 62, 1103–1118. https://doi.org/10.1111/j.1558-5646.2008.00349.x

Glaubitz, J. C., Casstevens, T. M., Lu, F., Harriman, J., Elshire, R. J., Sun, Q., & Buckler, E. S. (2014). TASSEL-GBS: A high capacity genotyping

by sequencing analysis pipeline. *PLoS ONE*, *9*, e90346. https://doi.org/10.1371/journal.pone.0090346

Grismer, L. L. (2000). Evolutionary biogeography on Mexico's Baja California peninsula: A synthesis of molecules and historical geology. *Proceedings of the National Academy of Sciences*, *97*, 14017–14018. https://doi.org/10.1073/pnas.260509697

Grismer, L. L., & McGuire, J. A. (1993). The oases of Central Baja California, Mexico. Part I. A preliminary account of the relict mesherpetofauna and the status of the oases. *Bulletin of the Southern California Academy of Sciences*, *92*, 2–24.

Grismer, L.L. (2002). A re-evaluation of the evidence for a Mid-Pleistocene seaway in Baja California: A reply to Riddle et al. *Herpetological Review*, *33*, 15–16.

Hafner, D. J., & Riddle, B. R. (2011). Boundaries and barriers of North American warm deserts: An evolutionary perspective. In P. Upchurch, A. McGowan, & C. Slater (Eds.), *Paleogeography and paleobiogeography: Biodiversity in space and time* (pp. 75–113). Boca Raton, FL: CRC Press. https://doi.org/10.1201/CRCSYSASSSPE

Hampe, A., & Jump, A. S. (2011). Climate relicts: Past, present, future. *Annual Review of Ecology, Evolution, and Systematics*, *42*, 313–333. https://doi.org/10.1146/annurev-ecolsys-102710-145015

Harley, M. M. (2006). A summary of fossil records for Arecaceae. *Botanical Journal of the Linnean Society*, *151*, 39–67. https://doi.org/10.1111/j.1095-8339.2006.00522.x

Heath, T. A., Hedtke, S. M., & Hillis, D. M. (2008). Taxon sampling and the accuracy of phylogenetic analyses. *Journal of Systematics and Evolution*, *46*, 239–257.

Henderson, A., Galeano, G., & Bernal, R. (1995). *Field guide to the palms of the Americas*. Princeton, NJ: Princeton University Press.

Hickerson, M. J., Carstens, B. C., Cavender-Bares, J., Crandall, K. A., Graham, C. H., Johnson, J. B., … Yoder, A. D. (2010). Phylogeography's past, present, and future: 10 years after Avise, 2000. *Molecular Phylogenetics & Evolution*, *54*, 291–301. https://doi.org/10.1016/j.ympev.2009.09.016

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, *25*, 1965–1978. https://doi.org/10.1002/(ISSN)1097-0088

Hoban, S., Kelley, J. L., Lotterhos, K. E., Antolin, M. F., Bradburd, G., Lowry, D. B., … Whitlock, M. C. (2016). Finding the genomic basis of local adaptation: Pitfalls, practical solutions, and future directions. *American Naturalist*, *188*, 379–397. https://doi.org/10.1086/688018

Hodel, D. R. (2014). Washingtonia filibusta (Arecaceae: Coryphoideae), a new hybrid from cultivation. *Phytoneuron*, *68*, 1–7.

Hoffman, J. I., Clarke, A., Linse, K., & Peck, L. S. (2011). Effects of brooding and broadcasting reproductive modes on the population genetic structure of two Antarctic gastropod molluscs. *Marine Biology*, *158*, 287–296. https://doi.org/10.1007/s00227-010-1558-6

Hoffman, J. I., Simpson, F., David, P., Rijks, J. M., Kuiken, T., Thorne, M. A. S., & Dasmahapatra, K. K. (2014). High-throughput sequencing reveals inbreeding depression in a natural population. *Proceedings of the National Academy of Sciences of the United States of America*, *111*, 3775–3780. https://doi.org/10.1073/pnas.1318945111

Huson, D. H., & Bryant, D. (2006). Application of phylogenetic networks in evolutionary studies. *Molecular Biology & Evolution*, *23*, 254–267. https://doi.org/10.1093/molbev/msj030

Jakobsson, M., & Rosenberg, N. A. (2007). CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, *23*, 1801–1806. https://doi.org/10.1093/bioinformatics/btm233

Jensen, J. L., Bohona, A. J., & Kelley, S. T. (2005). Isolation by distance, web service. *BMC Genetics*, *6*, 13. https://doi.org/10.1186/1471-2156-6-13

Joost, S., Bonin, A., Bruford, M. W., Despres, L., Conord, C., Erhardt, G., & Taberlet, P. (2007). A spatial analysis method (SAM) to detect candidate loci for selection: Towards a landscape genomics approach to adaptation. *Molecular Ecology*, *16*, 3955–3969. https://doi.org/10.1111/j.1365-294X.2007.03442.x

Karhu, A., Vogl, C., Moran, G. F., Bell, J. C., & Savolainen, O. (2006). Analysis of microsatellite variation in *Pinus radiata* reveals effects of genetic drift but no recent bottlenecks. *Evolutionary Biology*, *19*, 167–175. https://doi.org/10.1111/j.1420-9101.2005.00982.x

Keller, I., Wagner, C. E., Greuter, L., Mwaiko, S., Selz, O. M., Sivasundar, A., … Seehausen, O. (2013). Population genomic signatures of divergent adaptation, gene flow and hybrid speciation in the rapid radiation of Lake Victoria cichlid fishes. *Molecular Ecology*, *22*, 2848–2863. https://doi.org/10.1111/mec.12083

Kissling, W. D., Eiserhardt, W. L., Baker, W. J., Borchsenius, F., Couvreur, T. L. P., Balslev, H., & Svenning, J. S. (2012). Cenozoic imprints on the phylogenetic structure of palm species assemblages worldwide. *Proceedings of the National Academy of Sciences of the United States of America*, *109*, 7379–7384. https://doi.org/10.1073/pnas.1120467109

Klimova, A., Hoffman, I. J., Gutierrez-Rivera, J., Leon de la Luz, J., & Ortega-Rubio, A. (2017). Molecular genetic analysis of two native desert palm genera, *Washingtonia* and *Brahea*, from the Baja California Peninsula and Guadalupe Island. *Ecology & Evolution*, *7*, 4919–4935. https://doi.org/10.1002/ece3.3036

Kondo, T., Crisp, M. D., Linde, C., Bowman, D. M., Kawamura, K., Kaneko, S., & Isagi, Y. (2012). Not an ancient relic: The endemic Livistona palms of arid central Australia could have been introduced by humans. *Proceedings of the Royal Society B: Biological Sciences*, *279*, 2652–2661. https://doi.org/10.1098/rspb.2012.0103

León de la Luz, J. L. (2014). Floristic analysis in oases at central Baja California Sur, Mexico. In E. V. Wehncke, S. Alvarez Borrego, & E. Ezcurra (Eds.), *Conservation Science in Mexico's Northwest Ecosystem status and trends in the Gulf of California* (pp. 217–234). Mexico, Mexico: Instituto Nacional de Ecología y Cambio Climático.

León de la Luz, J. L., Rebman, J. P., & Oberbauer, I. (2003). On the urgency of conservation of Guadalupe Island, Mexico: Is it a lost paradise? *Biodiversity & Conservation*, *12*, 1073–1082. https://doi.org/10.1023/A:1022854211166

Levis, C., Costa, F. R. C., Bongers, F., Peña-Claros, M., Clement, C. R., Junqueira, A. B., & ter Steege, H. (2017). Persistent Effects of Pre-Columbian Plant Domestication on Amazonian Forest Composition. *Science*, *355*, 925–931. https://doi.org/10.1126/science.aal0157

Lexer, C., Wüest, R. O., Mangili, S., Heuertz, M., Stölting, K. N., &Bossolini, E. (2014). Genomics of the divergence continuum in an African plant biodiversity hotspot, I: Drivers of population divergence in Restio capensis (Restionaceae). *Molecular Ecology*, *23*, 4373–4386. https://doi.org/10.1111/mec.12870

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*, *25*, 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Lindell, J., Ngo, A., & Murphy, R. W. (2006). Deep genealogies and the mid-peninsular seaway of Baja California. *Journal of Biogeography*, *33*, 1327–1331. https://doi.org/10.1111/j.1365-2699.2006.01532.x

Luna, T. (2012). Propagation methods for *Washingtonia filifera* (Linden ex André) H. Wendl. (Arecaceae). *Native Plants Journal*, *13*, 216–221. https://doi.org/10.3368/npj.13.3.216

Martín-Barajas, A. (2014). The geological foundations of the Gulf of California Region. In E. V. Wehncke, S. Alvarez Borrego, & E. Ezcurra (Eds.), *Conservation science in Mexico's Northwest Ecosystem status and trends in the Gulf of California* (pp. 53–80). Mexico, Mexico: Instituto Nacional de Ecología y Cambio Climático.

Matala, A. P., Ackerman, M. W., Campbel, M. R., & Narum, S. R. (2014). Relative contributions of neutral and non-neutral genetic differentiation to inform conservation of steelhead trout across highly variable landscapes. *Evolutionary Applications*, *7*, 682–701. https://doi.org/10.1111/eva.12174

McClenaghan, L. R., & Beauchamp, A. C. (1986). Low genetic differentiation among isolated populations of California fan palm (Washingtonia filifera). *Evolution*, 40, 315–322.

McClintock, E. (1978). The Washington fan palm. *Fremontia*, 6, 3–5.

Meirmans, P. G. (2012). The trouble with isolation by distance. *Molecular Ecology*, 21, 2839–2846. https://doi.org/10.1111/j.1365-294X.2012.05578.x

Meirmans, P. G., & Van Tienderen, P. H. (2004). GENOTYPE and GENODIVE two programs for the analysis of genetic diversity of asexual organisms. *Molecular Ecology Notes*, 4, 792–794. https://doi.org/10.1111/j.1471-8286.2004.00770.x

Migliore, J., Baumel, A., Juin, M., Fady, B., Roig, A., Duong, N., & Medail, F. (2013). Surviving in Mountain Climate Refugia: New Insights from the Genetic Diversity and Structure of the Relict Shrub *Myrtus nivellei* (Myrtaceae) in the Sahara Desert. *PLoS ONE*, 8, e73795. https://doi.org/10.1371/journal.pone.0073795

Minnich, R. A., Franco-Vizcaíno, E., & Salazar-Ceseña, M. (2011). Distribution and regional ecology of Californian palm oases interpreted from Google Earth images. *Aliso*, 29, 1–12. https://doi.org/10.5642/aliso

Moore, W. S. (1995). Inferring phylogenies from mtDNA variation: Mitochondrial-gene trees versus nuclear-gene trees. *Evolution*, 49, 718–726.

Moran, R. (1996). *The flora of Guadalupe Island, Mexico*. Memoirs of the California Academy of Sciences, 19. San Francisco, CA: California Academy of Sciences.

Morin, P. A., Luikart, G., & Wayne, R. K. (2004). Applications of SNPs in ecology, evolution, and conservation. *Trends in Ecology & Evolution*, 19, 208–216. https://doi.org/10.1016/j.tree.2004.01.009

Mueller, K., Kier, G., Rockwell, T., & Jones, C. (2009). Quaternary rift-flank uplift of the peninsular ranges in baja and southern California by removal of mantle lithosphere. *Tectonics*, 28, 17.

Nadeau, S., Meirmans, P. G., Aitken, S. N., Ritland, K., & Isabel, N. (2016). The challenge of separating signatures of local adaptation from those of isolation by distance and colonization history: The case of two white pines. *Ecology & Evolution*, 6, 8649–8664. https://doi.org/10.1002/ece3.2550

Narum, S. R., Buerkle, C. A., Davey, J. W., Miller, M. R., & Hohenlohe, P. (2013). Genotyping-by-sequencing in ecological and conservation genomics. *Molecular Ecology*, 22, 2841–2847. https://doi.org/10.1111/mec.12350

Oberbauer, T. (2005). A comparison of estimated historic and current vegetation community structure on Guadalupe Island, Mexico. In D. K. Garcelon, & C. A. Schwemm (Eds.), *Proceedings of the Sixth California Islands Symposium*. National Park Service Technical Publication CHIS-05-01 (pp. 143–153). Arcata, CA: Institute for Wildlife Studies.

Oksanen, J., Blanchet, F.G., Legendre, P., McGlinn, D., Minchin, P.R., Keerthi, S., & Wagner, H. (2013). vegan: Community ecology package. Retrieved from http://CRAN.R-project.org/package=vegan (accessed May 2016).

Pannell, J. R., & Fields, P. D. (2014). Evolution in subdivided plant populations: Concepts, recent advances and future directions. *New Phytologist*, 201, 417–432. https://doi.org/10.1111/nph.12495

Papadopoulou, A., & Knowles, L. L. (2016). Toward a paradigm shift in comparative phylogeography driven by trait-based hypotheses. *Proceedings of the National Academy of Sciences of the United States of America*, 113, 8018–8024. https://doi.org/10.1073/pnas.1601069113

Patterson, N., Moorjani, P., Luo, Y., Mallick, S., Rohland, N., Zhan, Y., … Reich, D. (2012). Ancient admixture in human history. *Genetics*, 192, 1065–1093. https://doi.org/10.1534/genetics.112.145037

Pembleton, L. W., Cogan, N. O. I., & Forster, J. W. (2013). StAMPP: An R package for calculation of genetic differentiation and structure of mixed-ploidy level populations. *Molecular Ecology Resources*, 13, 946–952. https://doi.org/10.1111/1755-0998.12129

Pickrell, J. K., & Pritchard, J. K. (2012). Inference of population splits and mixtures from genome-wide allele frequency data. *PLoS Genetics*, 8, e1002967. https://doi.org/10.1371/journal.pgen.1002967

Popescu, A. A., Harper, A. L., Trick, M., Bancroft, I., & Huber, K. T. (2014). A novel and fast approach for population structure inference using kernel-PCA and optimization (PSIKO). *Genetics*, 198, 1421–1431. https://doi.org/10.1534/genetics.114.171314

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

Prunier, R., Akman, M., Kremer, C. T., Aitken, N., Chuah, A., Borevitz, J., & Holsinger, K. E. (2017). Isolation by distance and isolation by environment contribute to population differentiation in *Protea repens* (Proteaceae L.), a widespread South African species. *American Journal of Botany*, 104, 1–11.

Quero, H. J. (1992). Current status of Mexican palms. *Principes*, 36, 203–216.

Raeymaekers, J. A. M., Chaturvedi, A., Hablützel, P. I., Verdonck, I., Hellemans, B., Maes, G. E., … Volckaert, F. A. M. (2017). Adaptive and non-adaptive divergence in a common landscape. *Nature Communications*, 8, 267. https://doi.org/10.1038/s41467-017-00256-6

Rambaut, A., Suchard, M.A., Xie, D., & Drummond, A.J. (2014). Tracer 1.6. Retrieved from http://tree.bio.ed.ac.uk/software/tracer/ (Accessed July 2017).

Riddle, B. R., Hafner, D. J., Alexander, L. F., & Jaeger, J. R. (2000). Cryptic vicariance in the historical assembly of a Baja California peninsular desert biota. *Proceedings of the National Academy of Sciences of the United States of America*, 97, 14438–14443. https://doi.org/10.1073/pnas.250413397

Riemann, H., & Ezcurra, E. (2007). Endemic regions of the vascular flora of the peninsula of Baja California, Mexico. *Journal of Vegetation Science*, 18, 327–336. https://doi.org/10.1111/j.1654-1103.2007.tb02544.x

Rivera, D., Johnson, D., Delgadillo, J., Carrillo, M. H., Obón, C., Krueger, R., … Carreño, E. (2013). Historical evidence of the Spanish introduction of date palm (*Phoenix dactylifera* L., Arecaceae) into the Americas. *Genetic Resources and Crop Evolution*, 60, 1433–1452. https://doi.org/10.1007/s10722-012-9932-5

Rokas, A., & Carroll, S. B. (2005). More genes or more taxa? The relative contribution of gene number *and* taxon number *to* phylogenetic accuracy. *Molecular Biology & Evolution*, 22, 1337–1344. https://doi.org/10.1093/molbev/msi121

Roncal, J., Henderson, A., Borchsenius, F., Cardoso, S. R. S., & Balslev, H. (2012). Can phylogenetic signal, character displacement, or random phenotypic drift explain the morphological variation in the genus *Geonoma* (Arecaceae)? *Biological Journal of the Linnean Society*, 106, 528–539. https://doi.org/10.1111/j.1095-8312.2012.01879.x

Rundle, H., & Nosil, P. (2005). Ecological speciation. *Ecology Letters*, 8, 336–352. https://doi.org/10.1111/j.1461-0248.2004.00715.x

Savolainen, V., Anstett, M. C., Lexer, C., Hutton, I., Clarkson, J. J., Norup, M. V., & Baker, W. J. (2006). Sympatric speciation in palms on an oceanic island. *Nature*, 441, 210–213. https://doi.org/10.1038/nature04566

Savolainen, O., Lascoux, M., & Merilä, J. (2013). Ecological genomics of local adaptation. *Nature Reviews Genetics*, 14, 807–820. https://doi.org/10.1038/nrg3522

Schluter, D. (2000). *The ecology of adaptive radiation*. Oxford, UK: Oxford University Press.

Shafer, A. B. A., Peart, C., Tusso, S., Maayan, I., Brelsford, A., Wheat, C., & Wolf, J. B. W. (2017). Bioinformatic processing of RAD-seq data dramatically impacts downstream population genetic inference. *Methods in Ecology & Evolution*, 8, 907–917. https://doi.org/10.1111/2041-210X.12700

Shafer, A. B. A., & Wolf, J. B. W. (2013). Widespread evidence for incipient ecological speciation: A meta-analysis of isolation-by-ecology. *Ecology Letters*, 16, 940–950. https://doi.org/10.1111/ele.12120

Silverstein, R. P. (2005). Germination of native and exotic plant seeds dispersed by coyotes (*Canis latrans*) in Southern California. *The Southwestern Naturalist*, *50*, 472–478. https://doi.org/10.1894/0038-4909(2005)050[0472:GONAEP]2.0.CO;2

Sork, V. L., Gugger, P. F., Chen, J. M., & Werth, S. (2016). Evolutionary lessons from California plant phylogeography. *Proceedings of the National Academy of Sciences of the United States of America*, *113*, 8064–8071. https://doi.org/10.1073/pnas.1602675113

Spinks, P. Q., Thomson, R. C., & Shaffer, H. B. (2014). The advantages of going large: Genome-wide SNPs clarify the complex population history and systematics of the threatened western pond turtle. *Molecular Ecology*, *23*, 2228–2241. https://doi.org/10.1111/mec.12736

Stucki, S., Orozco-Wengel, P., Forester, B. R., Duruz, S., Colli, L., Masembe, C., … Joost, S. (2016). High performance computation of landscape genomic models including local indices of spatial association. *Molecular Ecology Resources*, https://doi.org/10.1111/1755-0998.12629

Tomlinson, P. B. (2006). The uniqueness of palms. *Botanical Journal of the Linnean Society*, *151*, 5–14. https://doi.org/10.1111/j.1095-8339.2006.00520.x

Tregear, J. W., Rival, A., & Pintaud, J. C. (2011). A family portrait: Unravelling the complexities of palms. *Annals of Botany*, *108*, 1387–1389. https://doi.org/10.1093/aob/mcr269

Vendrami, D. L. J., Telesca, L., Schweyen, H., Weigand, H., Weiss, M., Fawcett, K., … Hoffman, J. I. (2017). RAD sequencing resolves fine-scale population structure in a benthic invertebrate: Implications for understanding phenotypic plasticity. *Royal Society Open Science*, *4*, 160548. https://doi.org/10.1098/rsos.160548

Villanueva-Almanza, L., & Ezcurra, E. (2017). *Historia botánica temprana y variación clinal del genero Washingtonia (Arecaceae)*. XIV Simposio de Binacional de Botánica Baja California y Áreas Adyacentes.

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

Wehncke, E. V., & López-Medellín, X. (2014). Living at the edge: Blue fan palm desert oases of northern Baja California. In E. V. Wehncke, S. Alvarez Borrego, & E. Ezcurra (Eds.), *Conservation science in Mexico's Northwest ecosystem status and trends in the Gulf of California* (pp. 217–234). Mexico, Mexico: Instituto Nacional de Ecología y Cambio Climàtico.

Wehncke, E. V., López-Medellín, X., & Ezcurra, E. (2009). Patterns of frugivory, seed dispersal and predation of blue fan palms (*Brahea armata*) in oases of Northern Baja California. *Journal of Arid Environments*, *73*, 773–783. https://doi.org/10.1016/j.jaridenv.2009.03.007

Wehncke, E. V., López-Medellín, X., & Ezcurra, E. (2010). Blue fan palm distribution and seed removal patterns in three desert oases of northern Baja California, Mexico. *Plant Ecology*, *208*, 1–20. https://doi.org/10.1007/s11258-009-9682-4

Wiggins, I. L. (1980). *Flora of Baja California*. Stanford, CA: Stanford University Press.

Wollstein, A., & Lao, O. (2015). Detecting individual ancestry in the human genome. *Investigative Genetics*, *6*, 1–12.

Woolbright, S. A., Whitham, T. G., Gehring, C. A., Allan, G. J., & Beiley, J. K. (2014). Climate relicts and their associated communities as natural ecology and evolution laboratories. *Trends in Ecology & Evolution*, *29*, 406–416. https://doi.org/10.1016/j.tree.2014.05.003

Wright, S. (1943). Isolation by distance. *Genetics*, *28*, 114–138.

Zomer, R. J., Trabucco, A., Bossio, D. A., & Verchot, L. V. (2008). Climate change mitigation: A spatial analysis of global land suitability for clean development mechanism afforestation and reforestation. *Agriculture, Ecosystems & Environment*, *126*, 67–80. https://doi.org/10.1016/j.agee.2008.01.014

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.