

The role of geography, ecology, and hybridization in the evolutionary history of Canary Island *Descurainia*

Amy C. Jackson^{1,2} | Oliver W. White^{1,2} | Mark Carine² | Mark A. Chapman¹ 

¹Biological Sciences, University of Southampton, Southampton, SO17 1BJ, United Kingdom

²Algae, Fungi and Plants Division, Department of Life Sciences, The Natural History Museum, Cromwell Road, London, SW7 5BD, United Kingdom

Correspondence

Mark A. Chapman, Biological Sciences, University of Southampton, Southampton, SO17 1BJ, United Kingdom.

Email: m.chapman@soton.ac.uk

Abstract

Premise: Oceanic islands offer the opportunity to understand evolutionary processes underlying rapid diversification. Along with geographic isolation and ecological shifts, a growing body of genomic evidence has suggested that hybridization can play an important role in island evolution. Here we use genotyping-by-sequencing (GBS) to understand the roles of hybridization, ecology, and geographic isolation in the radiation of Canary Island *Descurainia* (Brassicaceae).

Methods: We carried out GBS for multiple individuals of all Canary Island species and two outgroups. Phylogenetic analyses of the GBS data were performed using both supermatrix and gene tree approaches and hybridization events were examined using D-statistics and Approximate Bayesian Computation. Climatic data were analyzed to examine the relationship between ecology and diversification.

Results: Analysis of the supermatrix data set resulted in a fully resolved phylogeny. Species networks suggest a hybridization event has occurred for *D. gilva*, with these results being supported by Approximate Bayesian Computation analysis. Strong phylogenetic signals for temperature and precipitation indicate one major ecological shift within Canary Island *Descurainia*.

Conclusions: Inter-island dispersal played a significant role in the diversification of *Descurainia*, with evidence of only one major shift in climate preferences. Despite weak reproductive barriers and the occurrence of hybrids, hybridization appears to have played only a limited role in the diversification of the group with a single instance detected. The results highlight the need to use phylogenetic network approaches that can simultaneously accommodate incomplete lineage sorting and gene flow when studying groups prone to hybridization; patterns that might otherwise be obscured in species trees.

KEYWORDS

ABBA-BABA, biogeography, *Descurainia*, diversification, genotyping-by-sequencing (GBS), hybridization, phylogenetics, speciation, species networks

Island systems have played a fundamental role in providing insights into the processes underlying the evolution of plant life (Warren et al., 2015). The Canary Islands, an oceanic archipelago composed of eight islands, is a hot spot for biodiversity and exhibits high levels of endemism (Reyes-Betancort et al., 2008). Approximately 570 plant species are endemic to the Canary Islands (25% of the entire flora; Guerra et al., 2008), many of which are found in recent and

rapidly diversifying lineages; the archipelago's flora has been used to investigate evolution, speciation, and adaptation (Juan et al., 2000). Three main processes have been identified to explain the high levels of endemism found across these islands: (1) allopatric speciation arising through geographical isolation, typically after inter island dispersal between similar habitats (Francisco-Ortega et al., 2000; Juan et al., 2000); (2) ecological speciation due to speciation into

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2023 The Authors. *American Journal of Botany* published by Wiley Periodicals LLC on behalf of Botanical Society of America.

new ecological niches (Silvertown et al., 2005); and (3) hybrid speciation owing to the weak reproductive barriers, dynamic landscapes, and the close proximity of related species with different ecological niches (Silvertown, 2004; Fjellheim et al., 2009; White et al., 2018). A significant challenge when constructing the evolutionary history of Canary Island lineages is untangling the relative influence of the different evolutionary processes involved.

Interspecific hybridization has long been viewed as an important driver of speciation in plants (Rieseberg and Willis, 2007). Nevertheless, despite numerous studies demonstrating hybridization and gene flow between Canary Island lineages (Brochmann, 1984; Francisco-Ortega et al., 1996; Kim et al., 1996; Saunders and Gibson, 2005; Carine et al., 2007; Fjellheim et al., 2009), most molecular studies have focused only on the relative roles of ecological speciation and inter-island dispersal in generating the high levels of species diversity observed (Barber et al., 2000; Francisco-Ortega et al., 2002; Allan et al., 2004; Jaén-Molina et al., 2021). Hybridization, as a potential evolutionary force, is often difficult to detect, particularly when using small numbers of genetic markers. The low variation uncovered when using a limited number of markers in recent species radiations generally results in poorly resolved phylogenies. Even when phylogenetic incongruence is identified, for example, when chloroplast [cpDNA] and nuclear [nDNA] DNA give different phylogenetic hypotheses (Francisco-Ortega et al., 1997; Barber et al., 2000; Mort et al., 2002; Jones et al., 2014), the use of only a small number of markers makes it difficult to determine the cause of the incongruence.

Reduced-representation sequencing methods, such as Restriction site Association DNA Sequencing (RAD-seq;

Eaton and Ree, 2013; Eaton et al., 2015) and Genotyping-by-Sequencing (GBS; Elshire et al., 2011), use high-throughput sequencing technologies to simultaneously sample hundreds or thousands of loci throughout the genome (Baird et al., 2008; Rowe et al., 2011; Eaton and Ree, 2013; Fernández-Mazuecos et al., 2018). These methods generate large numbers of phylogenetically informative markers, which can overcome limitations presented when using traditional marker methods. As a result, RAD-seq and GBS data sets are providing new insights into the evolution of the Canary Island flora, as demonstrated in *Micromeria* (Lamiaceae; Puppo et al., 2015) and *Argyranthemum* (Asteraceae; White et al., 2018, 2020). Combined with the development of new statistical methods (Green et al., 2010; Patterson et al., 2012; Sankararaman et al., 2014; Vernot and Akey, 2014), studies are also starting to emerge that offer more comprehensive insight into hybridization in oceanic lineages (Puppo et al., 2015; White et al., 2018, 2020). At present, these developments are reinforcing the notion that gene discordance as a result of hybridization is not only common but is shaping the plant diversity found across the Canary Islands.

Descurainia (Brassicaceae) Webb & Berthel is a genus of 45 species distributed throughout America, Eurasia, and the Canary Islands (Goodson et al., 2011). The seven species in the Canary Islands form a monophyletic group (Goodson et al., 2006), which is distributed across four islands (Figure 1). Five species (*D. gilva* Svent., *D. gonzalezii* Svent., *D. preauxiana* (Webb) Webb ex O.E.Schulz, *D. artemisioides* Svent., and *D. lemsii* Bramwell.) are single island endemics (SIEs) and two species (*D. millefolia* Webb & Berthel. and *D. bourgaeana* (E.Fourn.) Webb ex O.E.Schulz.) are multi-island endemics (MIEs) found on three and two islands,

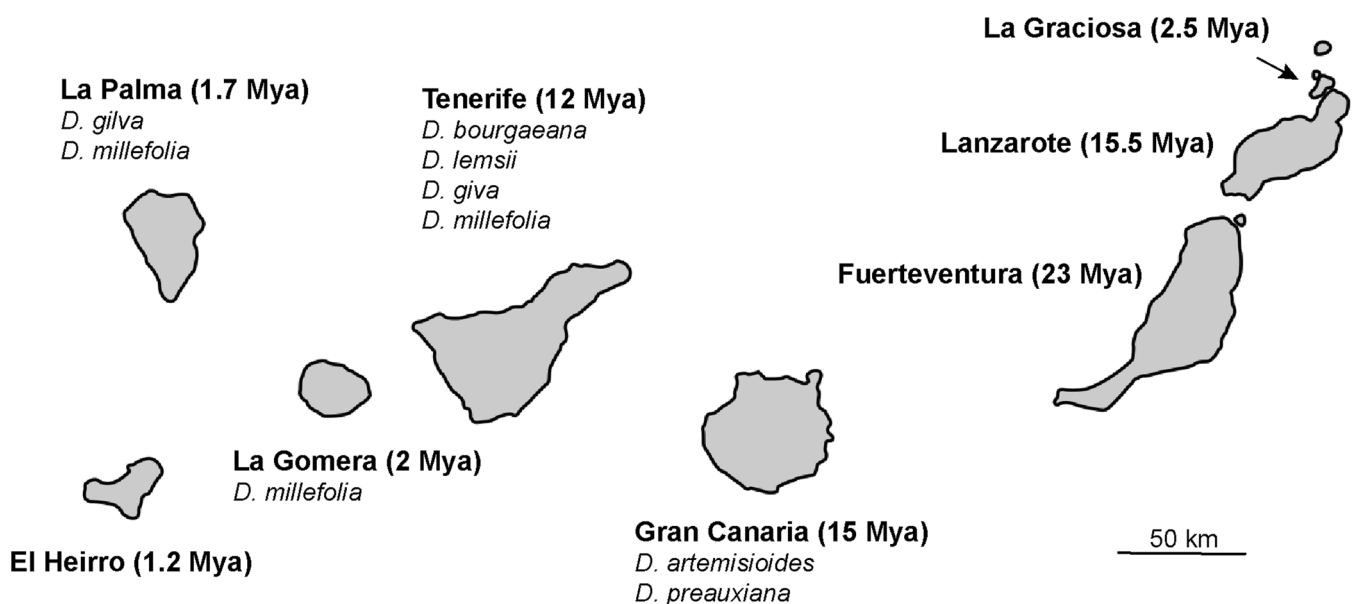


FIGURE 1 Map of the Canary Islands (age of islands in brackets; Schmincke [1976]) and the distribution of each island's seven endemic Canary Island *Descurainia*.

respectively (Bramwell, 1977; Goodson et al., 2006). The species are found in a range of different habitats, and Goodson et al. (2006) broadly categorized the ecological zones for Canary Island *Descurainia* as lowland scrub (250–700 m) for *D. millefolia*, *D. preauxiana*, and *D. artemisioides*; pine forest (600–1200 m) for *D. gilva*, *D. gonzalezii* and *D. lemsii*; and subalpine scrub (1900–2100 m) for *D. bourgaeana* (Bramwell, 1977; Goodson et al., 2006). Putative hybrids have been documented on Tenerife (*D. bourgaeana* × *D. lemsii*, and *D. gonzalezii* × *D. bourgaeana*; Goodson et al., 2006). Given the distribution of Canary Island *Descurainia* across multiple islands and ecological zones following a single colonization of the archipelago (Goodson et al., 2006; Goodson et al., 2011), the group offers an opportunity to examine the roles of different evolutionary processes underlying diversification within this archipelago.

Previous phylogenetic analyses of Canary Island *Descurainia* applied nDNA (the internal transcribed spacer [ITS] region) and cpDNA sequences and resolved the Euro-Siberian *Descurainia tanacetifolia* as the continental sister species to the island clade (Goodson et al., 2006, 2011). The ITS phylogeny failed to resolve relationships within the Canary Island clade, however, the cpDNA phylogeny revealed two clades. One clade was restricted to Tenerife (*D. gonzalezii*, *D. lemsii*, and some *D. millefolia* accessions) and the other was composed of taxa from across the four islands (*D. bourgaeana* on Tenerife, *D. artemisioides* and *D. preauxiana* on Gran Canaria, *D. gilva* on La Palma, and *D. millefolia* from La Palma and La Gomera). *Descurainia millefolia* was resolved in both clades, and *D. gonzalezii* was found to be polyphyletic in the Tenerife clade.

Goodson et al. (2006) suggested that the drivers of diversification in Canary Island *Descurainia* were intra-island ecological speciation and inter-island colonization. Studies that have subsequently used the Goodson et al. (2006) phylogenetic hypothesis to examine the adaptive radiation of *Descurainia* provide conflicting results on the role of climate and habitats (Herben et al., 2014; Steinbauer et al., 2016). The polyphyly of *D. millefolia* and *D. gonzalezii* revealed by their study could suggest taxonomic revision is needed, but it also raises questions over the influence of hybridization on the relationships inferred using chloroplast markers (Goodson et al., 2006). Chloroplast capture can give rise to topologies that do not reflect species relationships, leading to incorrect species trees (Rieseberg and Soltis, 1991). This is backed up by field observations of putative hybrids (Goodson et al., 2006). There is also the potential for gene flow because many species occur in close geographical proximity with a lack of karyotypic barriers, very similar floral morphologies, and long overlapping flowering periods (Bramwell, 1977).

Here, we use Genotyping-by-Sequencing (GBS) data in conjunction with ecological and distribution data across all seven Canary Island *Descurainia* species and two continental outgroups to resolve the evolutionary history of the group focusing on determining whether hybridization had played a significant role in the emergence of any taxa. Our

objectives are to (1) investigate phylogenetic relationships using both concatenated data and coalescence-based species networks, and using an Approximate Bayesian Computation (ABC) modelling analysis (Cornuet et al., 2008) to examine alternate phylogenies, (2) quantify the extent of hybridization between taxa using D-statistics (Eaton and Ree, 2013), (3) infer the role of ecology in the diversification of Canary Island *Descurainia* using spatial climatic variables, and (4) in so doing, establish the relative role of geographical isolation, hybridization, and ecological divergence in the diversification of the group.

MATERIALS AND METHODS

Sampling, DNA extraction, and GBS

A total of 18 individuals from eight populations were used to represent the seven Canary Island species of *Descurainia* (Appendix S1). Sampling was composed of one individual of *D. lemsii*, five of *D. bourgaeana*; two each of *D. gilva*, *D. gonzalezii*, *D. artemisioides* and *D. preauxiana*; and four of *D. millefolia* (two from La Palma and two from Tenerife). *Descurainia millefolia* populations from La Gomera and *D. bourgaeana* populations from La Palma were not sampled but all other species/island combinations were represented. *Descurainia tanacetifolia* and *D. depressa* were sampled as outgroups to represent the European sister species and a more distantly related South American taxon, respectively.

DNA was isolated from silica-dried leaves collected from wild plant material (collected in 2016; eight samples, permit numbers in the Acknowledgments section) and from fresh leaves of plants grown from seed from the Banco de Germoplasma Vegetal (BGV) seedbank (four samples; Appendix S1). DNA isolations were performed using a modified cetyltrimethyl ammonium bromide (CTAB) protocol (Doyle, 1991). An additional eight DNA samples used in the Goodson et al. (2006) study were supplied by R. Jansen (University of Texas).

DNA samples were sequenced using genotyping-by-sequencing (GBS) at Novogene (Yuen Long, Hong Kong). DNA was digested with restriction enzymes *MseI* and *HaeIII* and fragments were sequenced for 144 cycles on an Illumina HiSeq (Illumina, San Diego, CA, USA). The sampling of one to five individuals per species for GBS is expected to prove sufficient to provide powerful resolution for plant phylogenetic studies, while proving the depth needed to quantify hybridization, as demonstrated in Escudero et al. (2014), Anderson et al. (2017), Fernández-Mazuecos et al. (2018), and Pérez-Escobar et al. (2020).

GBS Assembly

The GBS data were quality-filtered and assembled using ipyrad v.0.7.23 (Eaton and Ree, 2013; Eaton, 2014). Raw

sequence data were demultiplexed for each sample using barcode sequences with no mismatches allowed. Low-quality base scores (<33) were converted to *N* and reads with more than five *N*s were discarded. Adapter sequences were filtered using the strict setting “2” as recommended for GBS data. The de novo-reference assembly method was used to remove reads that mapped to the chloroplast or mitochondrial genomes of *Arabidopsis thaliana* (L.) Heynh. (National Center for Biotechnology Information [NCBI] Reference Sequence NC_000932.1 and NC_001284.2, respectively) while assembling unmapped nuclear reads de novo. Reads were assembled to identify consensus allele sequences within individuals using three similarity clustering thresholds (80%, 85%, and 90%) before clustering consensus allele sequences across samples to identify loci. For the three levels of clustering employed, loci were filtered for a minimum number of samples per locus of 8 or 10, equivalent to 60% and 50% missing data, respectively. A minimum depth of six reads was required for base calling and any locus with a shared heterozygous site in 20% or more of samples was removed as potentially composed of paralogs. A total of six data sets were therefore generated with three clustering thresholds (hereafter referred to as c80, c85, c90) and two minimum samples per locus thresholds (hereafter m8, m10). For each assembly, ipyrad produces a custom “loci” format, which was used for all subsequent analyses.

Phylogenetic Reconstructions 1—Concatenated Approach

A supermatrix approach was first used for phylogenetic reconstructions to infer evolutionary relationships and to allow comparison between GBS assemblies. This involves the concatenation of loci from the GBS assembly data set into a single alignment. For each of the six assemblies, the optimal model of sequence evolution was identified using ModelTest-NG v.0.1.5 (Darriba et al., 2020) and a Maximum Likelihood (ML) tree was generated using RAxML Next-Generation v.0.6.0 (Stamatakis, 2014), with bootstrap support estimated from 1000 replicate searches from random starting trees.

Five of the six GBS data sets gave identical phylogenetic topologies (see Results). Of these five, the data set with a clustering threshold of 85% and a minimum sample number of eight (c85m8) was selected for all subsequent analyses because it had the greatest number of single nucleotide polymorphisms (SNPs).

Phylogenetic Reconstructions 2 – Coalescent Approach

The second phylogenetic approach uses a coalescent-based method, which can accommodate reticulated evolutionary histories and incomplete lineage sorting (ILS). Here, we

generated species networks modelling both hybridization and ILS using Phylonet v.3.5 (Than et al., 2008). To prepare the data, the unlinked SNP data set generated by Pyrad was transformed into bi-allelic markers. *Descurainia depressa* was removed from the analysis because of its more distant relationship to the Canary Island ingroup (see Results). The function “MLE_BiMarkers” in Phylonet was used to generate a pseudolikelihood ML estimation of phylogenetic networks with the “-diploid” and “-pseudo” parameters. The maximum number of reticulation nodes in the phylogenetic networks explored during the search was limited to eight. This number was set to a small number to reduce the overall complexity of a network. The number of runs was set to 50 to reduce memory load. Where multiple individuals of a species were present, these were mapped to a single species to reduce computational time to establish a species tree. The remaining parameters were set to their default settings.

Approximate Bayesian Computation (ABC)

We used an ABC approach (Beaumont et al., 2002) to assess the posterior probabilities of the evolutionary scenarios suggested by our analyses (concatenated vs. coalescent approaches; see Results). DIYABC v.2.1 (Cornuet et al., 2008) was used to compare two scenarios: (1) a fully bifurcating phylogeny representing the topology generated using ML analysis of the concatenated data set; and (2) a scenario including a hybrid origin for *D. gilva* between *D. millefolia* and *D. gonzalezii* as inferred in the coalescent analysis.

The unlinked SNP data set was transformed into a DIYABC-friendly format using the python script (vcf2DIYABC.py, available from <https://github.com/loire/> [accessed July 2019]).

Population sizes and divergence rates were set between 10^2 and 10^7 , and admixture rates were 0.001–0.999. A total of 10^7 simulations were performed for each of the two scenarios. Summary statistics were incorporated to compare the observed and simulated data comparisons, which were composed of the mean of genetic diversity, pairwise sample F_{st} and Nei's distance, and admixture summary statistics (Hudson et al., 1992; Choisy et al., 2004). The most probable scenario was identified using the posterior probability for each scenario, which was computed using a direct approach and logistic regression. The scenario probabilities were calculated using 500 simulations for the direct estimate and 10000 for the logistic regression.

Once the most probable scenario was identified, posterior analyses were carried out to evaluate the robustness of the simulation of the selected scenario with our GBS data set. Because DIYABC is a computationally and memory-intensive method, it was necessary to perform the remaining postsimulation analyses on a subset of 1000 simulations. First, the prior distribution of the scenario was performed using a logistic regression from which false positives and false negatives in the choice of scenario were

estimated. False positives and negatives were calculated by measuring the proportion of the simulated data set for the best scenario when assigned to the other scenario, and the proportion of data sets simulated under the other scenario that was assigned to the best scenario, respectively. Second, goodness of fit for the selected scenario was performed using DIYABC model-checking analysis. This is to assess how close the simulated data set fits our GBS data set by producing summary statistics and ranking the observed value among the values obtained with simulated data sets.

Patterson's D-statistics

Patterson's D-statistics (also known as ABBA-BABA tests) are used to infer introgression between two lineages in a given phylogeny by testing for shared derived alleles between taxa. ABBA-BABA tests are strictly a tree-based test, from which species are assigned to the topology (((P1, P2), P3), O). P1 and P2 are species that belong to a monophyletic group, P3 corresponds to a taxon from a different ingroup species, and O is the outgroup. The D-statistic measures the asymmetry in the number of alleles supporting "ABBA" and "BABA" patterns of allele distributions (where A is ancestral, and B is derived). If the proportion of alleles is equal (ABBA = BABA), then we cannot rule out ILS as the likely cause of gene incongruence. In contrast, a significant asymmetry between the number of alleles for each topology suggests introgression has occurred between P3 and either P2 (ABBA > BABA) or P1 (BABA > ABBA).

To test the hybridization event indicated in our species networks (see results), we used Patterson's D-statistic to test whether there is an excess of shared alleles between *D. millefolia* and *D. gilva*. Because of the tree's topology, it was impossible to test whether there is gene flow between *D. gonzalezii* and *D. gilva*, and *D. millefolia*.

SNPs from the c85m8 assembly were used to generate the D-statistic using the *baba* tool in ipyrad v.0.9.16 (Eaton, 2014). A *P*-value was calculated from the *Z*-score and adjusted for multiple comparisons using Bonferroni correction.

Ecological niche occupation and conservatism

Global positioning system (GPS) coordinates were taken from 216 Canary Island *Descurainia* herbarium specimens held at the Natural History Museum, London, United Kingdom. Climatic variables for these localities were obtained from WorldClim v.2.1 (years 1970–2000; Fick and Hijmans, 2017), using 19 bioclimatic (bioclim) variables derived from temperature and rainfall values at a spatial resolution of 30 s (~1 km²). The GIS data sets were used to interpolate climate values for all individuals using ArcMap within ArcGIS (ESRI, 2011) and the means for each bioclimatic variable per species were calculated. The *pairs*

function in R was used to identify correlated variables. Where correlated variables were identified, the one that explained the largest differences between species was retained for further analysis. A phylogenetic principal component analysis (pPCA) was generated with the R package Phytools (Revell, 2012), using the phylogeny generated from the analysis of the concatenated data set and the five bioclim variables retained. The averages for each climatic variable were represented in a dotTree using Phytools, along with the results of the phylogeny PCA (PC1 and PC2).

We tested for phylogenetic signal in each trait using Blomberg's *K* statistic, as implemented with the function "phylosig" in the R package Phytools. The *K* values closer to zero correspond to a random or convergent pattern of evolution, while values greater than 1 indicate a strong phylogenetic signal and conservatism of traits.

RESULTS

GBS assembly

Between 0.46 and 2.14 million reads were generated for each sample. After quality control and removal of poor-quality reads, six data sets were assembled from the GBS raw reads (Table 1). Between 113,107 and 122,724 loci were generated in each assembly (Table 1). An increase in the clustering threshold (*c*; 80%, 85%, and 90%) resulted in a greater number of loci. Increasing the minimum sample coverage (*m*) from 8 to 10 resulted in a considerable decrease in the number of total SNPs and parsimony informative (PI) SNPs available (Table 1).

Phylogenetic Reconstructions 1—Concatenated Approach

The GTR+I+G model of nucleotide evolution was found to be the best model for all six assemblies (Appendix S2) and was employed for ML phylogeny reconstruction.

All phylogenies resolved two main clades (bootstrap support, BS = 100%), "clade A" and "clade B" (Figure 2A; Appendix S3). Clade A is composed of *D. preauxiana*, *D. millefolia*, and *D. artemisioides*, and in five of the six phylogenies (i.e., excluding c90m8), *D. millefolia* is resolved as monophyletic (BS = 54–82%). *D. millefolia* is sister to *D. artemisioides*, with *D. preauxiana* resolved as sister to this pair. Within *D. millefolia*, Tenerife populations are resolved as paraphyletic with respect to the La Palma accessions that are resolved as monophyletic (BS = 100%). For the data set with the alternate topology (c90m8), *D. millefolia* is resolved as polyphyletic with one individual nested within *D. artemisioides* (BS = 48%).

The clade B topology was consistent for all six assemblies and is composed of subclades B1 (*D. gilva* and *D. gonzalezii*) and B2 (*D. bourgaeana* and *D. lemsii*)—both

TABLE 1 Summary of the results of filtering and clustering of genotyping-by-sequencing (GBS) raw sequences using ipyrad (Eaton, 2014). The GBS reads were obtained from 18 individuals from seven Canary Island *Descurainia* species and two continental *Descurainia* individuals. Loci = unique GBS DNA cluster; PI = parsimony informative; SNPs (Single nucleotide polymorphisms) = PI SNPs and autapomorphies; PI SNPs = minor allele in >1 sample; PI uSNPs = unlinked PI SNP.

Assembly	Clustering Threshold (%)	Minimum Taxon Coverage	N loci	SNPs	PI SNPs	PI uSNPs
c80m8	80	8	113,107	69,339	29,037	12,779
c80m10	80	10	113,107	41,235	18,002	7201
c85m8	85	8	118,553	75,603	31,714	13,797
c85m10	85	10	118,553	45,347	19,794	7828
c90m8	90	8	122,724	76,875	32,488	14,652
c90m10	90	10	122,724	45,526	19,981	8174

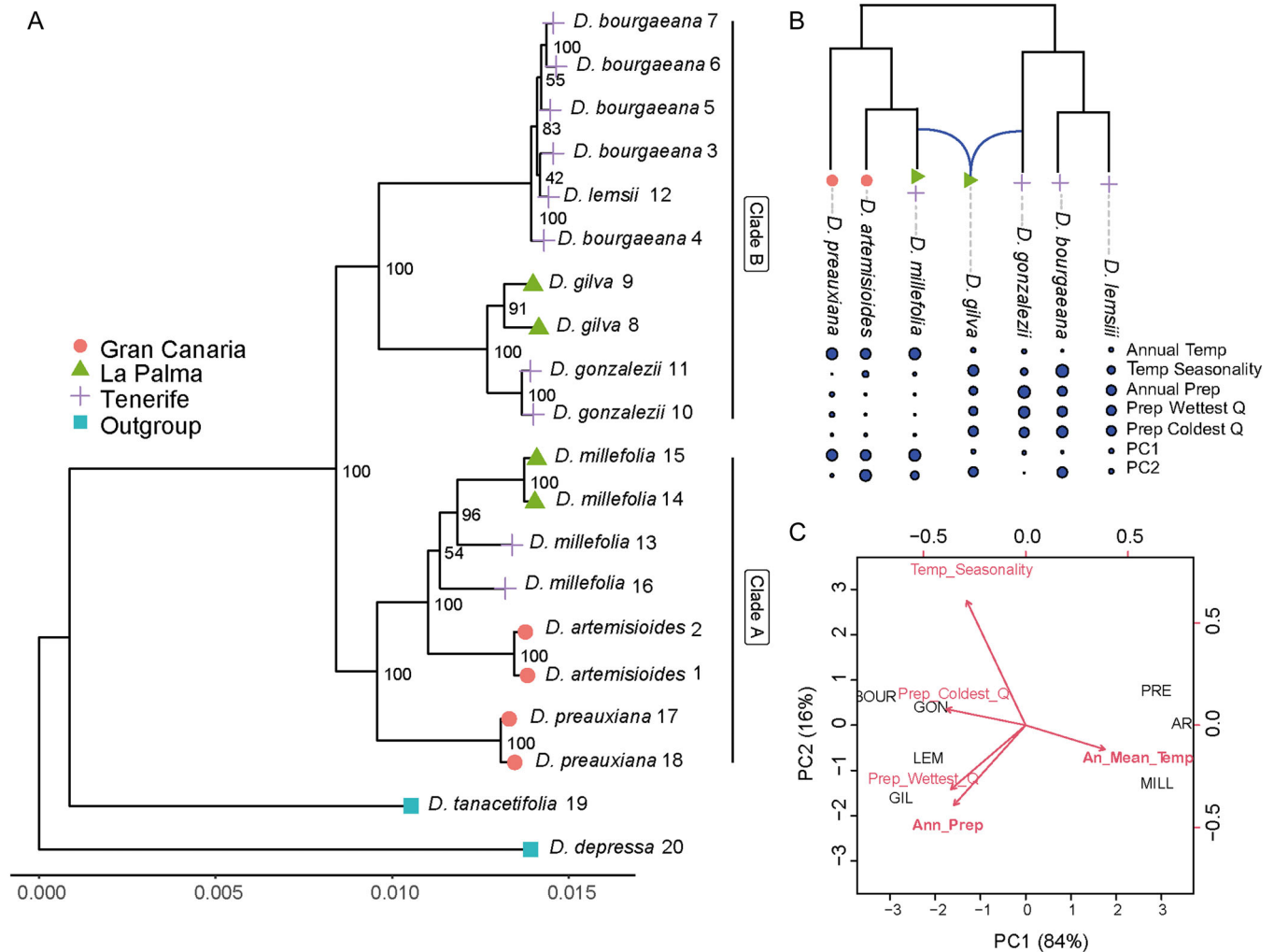


FIGURE 2 (A) Maximum-likelihood phylogenetic tree inferred from genotyping-by-sequencing (GBS) data set (c85m8) for 18 Canary Island *Descurainia* individuals and two continental relatives. Numbers represent bootstrap values as inferred from 1000 bootstrap (BS) repetitions. Branch lengths are represented by the bottom bar. The continental relatives, *D. depressa* and *D. tanacetifolia*, are the outgroups. ID numbers next to taxa names represent samples in Appendix S1, and shapes represent the island of origin. (B) Schematic representing the phylogenetic network generated by Phylonet under pseudo-likelihood when the maximum reticulation is set to eight. Dots represent the average value of bioclimatic variables and phylogenetic principal component analysis (pPCA) loads for each species. (C) The pPCA plots of PC1 and PC2 from five bioclimatic variables. Arrows in bold indicate a significant phylogenetic signal ($K > 1$ at $P < 0.05$).

of which have maximum support (BS = 100%). In B1, both *D. gonzalezii* and *D. gilva* are resolved as monophyletic (BS > 96%). Within Clade B, *D. lemsii* is found nested within *D. bourgaeana*, although the position is variable and weakly supported (BS = 42–79%).

Phylogenetic Reconstructions 2 – Coalescent Approach

A total of 13,797 unlinked SNPs were used as input data for Phylonet. The phylogenetic network with the lowest log-likelihood score from all 50 runs resulted in a network with one reticulation (Figure 2B). The resulting topology was similar to that of the concatenated approach but *D. gilva* is suggested to be of hybrid origin between *D. gonzalezii* (clade B) and *D. millefolia* (clade A).

ABC

The c85m8 SNP data set was filtered, so that there was at least one individual per species at each SNP site (as required by DIYABC), after which 1722 SNPs remained. Scenario 2, in which *D. gilva* is of hybrid origin between *D. millefolia* and *D. gonzalezii*, had the highest posterior probability (0.786, 95% confidence interval 0.43–1.00; Table 2). Scenario 1, which represented the bifurcating phylogeny without hybridization, had a much lower posterior probability (0.214; confidence interval 0.00–0.57).

The summary statistics calculated from the scenario 2 simulated data set were found to be close to the summary statistics calculated with our GBS data set (Appendix S4). The estimation of error rates provided confidence in scenario choice; the false positive and false negative error rates were estimated to be 10.2% and 0.1%, respectively.

Our analysis also indicated that the parental contributions to the genetic composition of the hybrid *D. gilva* were not equal (Appendix S4). There was a large bias in contribution with *D. gonzalezii* contributing 98% (95% CI: 98.4–99.6) and *D. millefolia* 2%. The parameter estimation for the time of this hybridization event was 3.7 ± 2.1 Mya.

Patterson's D-statistics

Patterson's D-statistic examining gene flow between *D. gilva* (P3) and *D. millefolia* (P2) resulted in a significant D-statistic (BABA > ABBA) when both *D. artemisioides*

and *D. preauxiana* were used as P1 (Table 3; $P < 0.005$, after multiple correction).

Ecological niche occupation and conservatism

After filtering the most highly correlated variables, five BioClim variables were included in the ecological analyses: annual temperature, temperature seasonality, annual precipitation, precipitation in the wettest quarter, and precipitation in the coldest quarter (Appendix S5). The first two axes, principal components 1 and 2, of the phylogenetically corrected PCA explained 84% and 16% of the variance in climatic variables across the seven species (Figure 2C, Table 4). Precipitation in the coldest quarter (−0.99) and annual mean temperature (0.98) have the highest loads for PC1. *Descurainia bourgaeana* (−3.45) was on one extreme of PC1, followed by *D. gilva*, *D. lemsii*, and *D. gonzalezii*, all demonstrating high average precipitation, whereas *D. artemisioides* (3.44), *D. millefolia* (2.77), and *D. preauxiana* (2.77) are on the opposite end of the PC1 axis and represent higher annual mean temperature.

Temperature seasonality (0.67) and annual precipitation (−0.43) represented the highest loads for PC2 with *D. gilva* (−1.59) and *D. preauxiana* (0.72) on the extremes of this axis.

The *K* statistic indicated a strong phylogenetic signal for all bioclimatic variables ($K > 1$; Table 4). Annual mean temperature, annual precipitation, and precipitation in the wettest quarter were significant ($p < 0.05$), indicating these traits have a nonrandom phylogenetic signal.

DISCUSSION

The GBS analysis presented provides new insights into phylogenetic relationships among Canary Island *Descurainia* (Brassicaceae) and the evolutionary processes that underpin the group's diversification. Phylogenetic analyses of the concatenated data set revealed two major and well-resolved clades, with ABBA-BABA tests and species networks both supporting a role for hybridization in the diversification of the group. Our results suggest that a hybridization event has occurred for *D. gilva*. Furthermore, the results suggest that diversification of the group has involved a single shift in climatic preferences associated with a shift in elevation, with the low-elevation Clade A species occupying high temperature and low precipitation environments and the higher-elevation Clade B species

TABLE 2 Posterior probability of parameters, and their confidence intervals, for each four scenarios tested with approximate Bayesian computation (ABC) in DIYABC (Cornuet et al., 2008). The corresponding phylogeny or network for each scenario is found in Figure 2.

Scenario	Test Description	Posterior Probability	95% Confidence Interval
1	Bifurcating phylogeny	0.214	[0.000, 0.574]
2	<i>D. gilva</i> is a hybrid origin between parents <i>D. millefolia</i> and <i>D. gonzalezii</i> .	0.786	[0.427, 1.000]

TABLE 3 Patterson's D-statistic (ABBA-BABA), which compares *D. gilva* (gil; P3) with *D. millefolia* (mil; P2) and its sister taxa (P1), either *D. artemisioides* (art) or *D. preauxiana* (pre). *Descurainia tanacetifolia* and *D. depressa* were selected as the outgroups (O). D-statistics in bold represent significant *P* values at < 0.05.

Test	D stat	Z	ABBA	BABA	# loci	P value
art (P1) mil (P2) gil (P3)	-0.121	3.929	1031.256	1313.998	5817	0.000
pre (P1) mil (P2) gil (P3)	-0.132	3.221	610.708	796	3530	0.001

TABLE 4 The loadings of each trait for the first two axes from the phylogenetic principal component analysis and Blomberg's *K* statistic under a Brownian motion model of evolution for our five bioclimatic variables. Traits that are significant where *P* < 0.05 have a nonrandom phylogenetic signal. The *K* values greater than 1 indicate a strong phylogenetic signal and phylogenetically conservation.

Bio Variable	PC1 (84%)	PC2 (16%)	K	K P value
Annual Mean Temp (BIO1)	0.989	-0.131	1.365	0.045
Temp Seasonality (BIO4)	-0.738	0.672	1.447	0.037
Annual Precipitation (BIO12)	-0.898	-0.432	1.150	0.116
Precipitation in Wettest Quarter (BIO16)	-0.935	-0.346	1.395	0.041
Precipitation in Coldest Quarter (BIO19)	-0.992	0.087	1.530	0.053

occupying lower temperature and higher precipitation environments. The study, therefore, indicates the importance of geographic isolation, hybridization and ecological diversification in driving the diversification of Canary Island *Descurainia*.

Inferring species relationships using coalescent vs. concatenated methods

Previous studies have suggested that phylogenies based on concatenated GBS data sets suffer strong systematic bias because high statistical support for multiple alternative topologies—depending on clustering threshold values used in the sequence analysis—is often found (Fernández-Mazuecos et al., 2018). Here, our GBS assemblies show congruent and well-supported topologies across five of six data sets. The incongruent data set differed only in the resolution of *D. millefolia*. Phylogenetic analysis of all concatenated data sets revealed two major clades (A and B) and two subclades within B (B1 and B2) with maximum bootstrap support. This suggests the different clustering and minimum sample coverage thresholds used have little effect on phylogenetic inferences, proving confidence in the results.

In the topology resolved in five of the six analyses, *D. bourgaeana* was the only species revealed as nonmonophyletic, with *D. lemsii* nested within this species, although branch lengths in this clade were short and relationships were not well supported. The occurrence of paraphyly such as this is often explained by taxonomic error, ILS, or hybridization (Holder et al., 2001). Morphologically, *D. bourgaeana* and *D. lemsii* may be distinguished because the former has decurrent leaves and ascending siliquae, whereas the latter lacks decurrent leaf segments and has erect

siliquae (Bramwell, 1977). However, hybridization has played a role in Canary Island *Descurainia* more generally (see below) and putative hybrids of *D. bourgaeana* and *D. lemsii* have been observed in the wild (Goodson et al., 2006; A.C.J., personal field observations). Because of the topology of our phylogeny, it was not possible to test for hybridization (versus ILS) involving these two taxa with Patterson's D-statistics. Given that neither ILS nor hybridization can be ruled out, further morphological and genetic studies with greater sampling of *D. lemsii* are necessary to resolve the status of these two taxa.

The topology of our concatenated phylogeny was incongruent with the previous phylogenetic reconstructions of Goodson et al. (2006) based on 44 cpDNA characters from seven loci. Specifically, our analysis did not support the polyphyly of *D. gonzalezii* and *D. millefolia* found by Goodson et al. (2006); instead, both were resolved as monophyletic. The Goodson et al. phylogeny may be affected by chloroplast (cp) capture (the transfer of cpDNA between species due to hybridization). Given that our results indicate hybridization in the evolutionary history of the Canary Islands *Descurainia*, cpDNA capture likely explains the significant differences between the results of Goodson et al. (2006) and those presented here. Furthermore, our phylogenetic reconstructions were generated using a significantly larger number of genetic markers used in our GBS data sets generated through next-generation sequencing technology. For instance, our ML phylogeny was generated using 76,875 SNPs (c85m8 GBS assembly; Figure 2.2), which significantly exceeds the 44 cpDNA characters employed in the Goodson et al. (2006) study. Our data sets provide a far more powerful phylogenetic signal to resolve the relationships within Canary Island *Descurainia* and likely explains the conflicting topologies between studies.

In recent years, evolutionary networks have been explored as an alternative method for phylogenetic reconstructions (McCluskey and Postlethwait, 2015). Because reduced-representation data sets (i.e., RADseq and GBS) sample data from hundreds to tens of thousands of loci across the entire genome, these sequencing methods can be applied to generate multilocus species trees. While Phylonet has been shown to be a promising tool in estimating hybridization between taxa (Hibbins and Hahn, 2021), few studies have attempted to use reduced-representation data sets to estimate evolutionary networks (Eaton and Ree, 2013; Blanco-Pastor et al., 2019). Several issues have been highlighted with this approach, notably the short length of each locus means that there is variable phylogenetic information between loci, and the data sets may be subjected to phylogenetically structured patterns of missing data (Salichos and Rokas, 2013). However, a coalescent analysis of reduced-representation data allows for variation across loci in a genealogy, whereas a concatenated approach assumes a shared phylogenetic history for all genes (Rokas et al., 2003). As a result, a coalescent approach can be useful for understanding hybridization.

Here, the nonreticulated species network demonstrated a similar topology to that of the concatenated analysis. However, we found evidence that the most likely evolutionary scenario for Canary Island *Descurainia* is reticulate and includes interspecies gene flow. The reticulated phylogenies generated by Phylonet were also supported by D-statistics and DIYABC models. This implies that the concatenated approach, despite the larger number of loci included and the high bootstrap support, has masked incongruence between loci within the data set.

Our analyses support a hybridization event for *D. gilva*, between the *D. gonzalezii* and *D. millefolia* lineages. This is at odds with the hypothesis of Bramwell (1977), who, based on morphological similarities, considered *D. gilva* to be a vicariant of *D. bourgaeana*, a species that also occupies high altitudes.

Our ABC analysis indicated there is a large bias in putative parental contributions, with 98% contribution from *D. gonzalezii* and 2% from *D. millefolia*. Inter-island hybridization has been documented in other Canary Island endemic plant lineages (van Hengstum et al., 2012; Puppo et al., 2015; White et al., 2020). However, it seems most likely that *D. gilva* is rather a vicariant form of *D. gonzalezii*, which has been exposed to gene flow from *D. millefolia* upon establishing itself on La Palma. It is notable that *D. gilva* and *D. gonzalezii* are somewhat ecologically distinct; *D. gilva* habitats exhibit annual precipitation and precipitation in the wettest quarter >80% higher than *D. gonzalezii* habitats (Figure 2B; Table S4). Our work therefore suggests that introgressive hybridization, despite only representing a minor fraction of the hybrid genome, was enough for a lineage to establish in a new niche, as has been reported in *Senecio* (Kim et al., 2008) and *Helianthus* (Whitney et al., 2006).

Diversification and the role of ecology and geography

Inter-island dispersal and ecological shifts are the two processes most commonly invoked to explain the high levels of diversity seen within Canary Island lineages (Francisco-Ortega et al., 2002; Mort et al., 2002; Allan et al., 2004; Jones et al., 2014). Previous phylogenetic analysis of Canary Island *Descurainia* suggested that inter-island dispersal was more influential than ecological speciation (Goodson et al., 2006), with three ecological shifts inferred alongside multiple inter-island colonization events. Our phylogeny also supports the importance of geographic speciation in the evolution of Canary Island *Descurainia*. Within clade B1, there is a dispersal event from Tenerife to La Palma for *D. gilva*, and in clade A, there are likely two dispersal events: one from Gran Canaria to Tenerife, followed by another from Tenerife to La Palma.

Climatic variation within the Canary Islands has often been seen as a major driver of diversification within endemic plant lineages (Harter et al., 2015; Irl et al., 2015). Climatic variation, especially along elevation gradients, promotes ecological shifts and speciation (Hua and Wiens, 2013); for example, precipitation gradients have played a significant role in the radiation of *Aeonium* (Crassulaceae; Harter et al., 2015). Our results indicate that for *Descurainia*, precipitation and temperature are strongly linked phylogenetically. We observe only one shift in climatic preferences, supporting the idea of niche conservatism between closely related *Descurainia* species. Taxa in each of our two clades occupy similar ecological zones. *Descurainia gonzalezii*, *D. lemsii*, and *D. bourgaeana* in clade B are species that inhabit Tenerife with similar environmental pPCA axes. They are largely allopatric, suggesting intra-island geographic diversification without large ecological divergence. Similarly, *D. artemisioides* and *D. preauxiana*, both endemic to Gran Canaria, occupy similar habitats but are largely allopatric: *D. artemisioides* is restricted to the northwest of the island and *D. preauxiana* is more widespread but they do not overlap geographically. Inter- and intra-island isolation within similar habitats has played a more significant role in the diversification of the group than ecological shifts based on the data to hand. However, we acknowledge that other factors—for example, wind exposure and soil types—should be examined.

The phylogenetic framework of Goodson et al. (2006) has been used to investigate the role of adaptive divergence in the evolution of Canary Island *Descurainia*. Steinbauer et al. (2016) examined climatic niche differentiation between pairs of species existing in sympatry with that for pairs of species in allopatry for a range of Canary Island radiations, including *Descurainia*. They concluded that *Descurainia* showed niche conservatism for temperature. In contrast, Herben et al. (2014) proposed that morphological traits related to water availability were not phylogenetically linked and therefore suggested adaptive divergence to differing

habitats based on water availability. However, both of these studies used the Goodson et al. (2006) phylogenetic framework, which conflicts with our own phylogeny.

While several Macaronesian plant radiations have been subject to molecular phylogenetic analysis, resolution of relationships and support for those relationships are often limited and many groups would benefit from further investigation with more extensive molecular sampling. Where reduced representation sequencing approaches have been used, e.g., to investigate the Macaronesian endemics *Argyranthemum*, *Micromeria*, and *Lavatera* (Villa-Machío et al., 2020; White et al., 2020), better resolved phylogenies have resulted and more complex patterns of relationships than earlier assessments suggested have been revealed.

Our results suggest that ecological shifts are not entirely absent from the diversification of Canary Island *Descurainia*. A significant ecological distinction is found between the two clades, and our pPCA results show that species from clade A are found in warm and arid habitats, and species from clade B are found in cooler, wetter habitats. Upslope migrations, involving adaptation to higher altitudes, are a common observation in Canary Island lineages, e.g., in *Helianthemum* (Cistaceae) and *Echium* (Albaladejo et al., 2021; Graham et al., 2021). This fits with the finding that the initial colonization of the Canary Islands is often at low altitude where more diverse habitats are available, followed by diversification upslope to more specialized high-altitude habitats (Steinbauer et al., 2012). However, in *Descurainia*, the sister group of the Canary Islands clade, *D. tanacetifolia* (Goodson et al., 2006), is a species restricted to montane regions of northern Iberia and the Alps. Given our data, we cannot rule out “downslope” migration from cooler wetter habitats to warmer, drier habitats. A more comprehensive analysis of the evolution of climatic preferences in the genus would be necessary to provide insights into this.

CONCLUSIONS

Despite only analyzing a few individuals per species, we produced a new multilocus framework for Canary Island *Descurainia*, which indicates that hybridization has occurred within this group, likely explaining species polyphyly identified in an earlier study (Goodson et al., 2006). Inter-island dispersal, within-island isolation, and to a lesser extent, ecological shifts and hybridization, are all implicated in the diversification of Canary Island *Descurainia*. The findings of this study reinforce the notion that hybridization is one of the drivers contributing to the complex history of Canary Island flora. It also highlights how phylogenetic analyses of island lineages should employ multiple analytical approaches to test for alternative scenarios.

AUTHOR CONTRIBUTIONS

A.C.J., M.C., and M.A.C. designed the study; M.A.C. extracted DNA, O.W.W. assembled the raw GBS data with

A.C.J. performing the rest of the analyses; A.C.J. wrote the manuscript with contributions from all authors.

ACKNOWLEDGMENTS

The authors would like to thank Rachael Graham and Alfredo Reyes-Bentacourt for Canary Island field sampling. We also thank Barbara Goodson and Bob Jansen (University of Texas) for providing DNA samples, and the Banco de Germoplasma Vegetal (BGV) Seedbank for providing *Descurainia* seeds. We are grateful to the following national parks for permits to collect. Cabildo de Tenerife (permit no. 22835), Parque Nacional del Teide (permit no. 24339), Cabildo de La Palma (permit no. 2016005709), Parque Nacional de la Caldera de Taburiente (permit no. 671303), Parque Nacional Garajonay (permit no. 25771), Cabildo de Gran Canaria (permit no. 16267), and Gobierno de Canarias (permit no. 671303). We would like to thank the reviewers for their helpful and insightful comments, which helped improve our manuscript. This work was supported by a Ph.D. studentship to ACJ from the Natural Environmental Research Council (grant number NE/L002531/1). We acknowledge the use of the IRIDIS High Performance Computing Facility, and associated support services at the University of Southampton, in the completion of this work.

DATA AVAILABILITY STATEMENT

Climate data for Canary Island *Descurainia* individuals has been uploaded to Figshare (<https://doi.org/10.6084/m9.figshare.22270987.v1>) and raw Genotyping-by-Sequencing to the NCBI SRA (BioProject number PRJNA894797).

ORCID

Mark A. Chapman  <http://orcid.org/0000-0002-7151-723X>

REFERENCES

- Albaladejo, R. G., S. Martín-Hernanz, J. A. Reyes-Betancort, A. Santos-Guerra, M. Olangua-Corral, and A. Aparicio. 2021. Reconstruction of the spatio-temporal diversification and ecological niche evolution of *Helianthemum* (Cistaceae) in the Canary Islands using genotyping-by-sequencing data. *Annals of Botany* 127: 597–611.
- Allan, G. J., J. Francisco-Ortega, A. Santos-Guerra, E. Boerner, and E. A. Zimmer. 2004. Molecular phylogenetic evidence for the geographic origin and classification of Canary Island *Lotus* (Fabaceae: Loteae). *Molecular Phylogenetics and Evolution* 32: 123–138.
- Anderson, B. M., K. R. Thiele, S. L. Krauss, and M. D. Barrett. 2017. Genotyping-by-sequencing in a species complex of Australian hummock grasses (*Triodia*): Methodological insights and phylogenetic resolution. *PLoS One* 12: e0171053.
- Baird, N. A., P. D. Etter, T. S. Atwood, M. C. Currey, A. L. Shiver, Z. A. Lewis, E. U. Selker, et al. 2008. Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS One* 3: e3376.
- Barber, J. C., J. F. Ortega, A. Santos-Guerra, A. Marrero, and R. K. Jansen. 2000. Evolution of endemic *Sideritis* (Lamiaceae) in Macaronesia: Insights from a chloroplast DNA restriction site analysis. *Systematic Botany* 25: 633–647.
- Beaumont, M. A., W. Zhang, and D. J. Balding. 2002. Approximate Bayesian computation in population genetics. *Genetics* 162: 2025–2035.
- Blanco-Pastor, J. L., Y. J. K. Bertrand, I. M. Liberal, Y. Wei, E. C. Brummer, and B. E. Pfeil. 2019. Evolutionary networks from RADseq loci point

- to hybrid origins of *Medicago carstiensis* and *Medicago cretacea*. *American Journal of Botany* 106: 1219–1228.
- Bramwell. 1977. A revision of *Descurainia* Webb and Berth. Section *Sisymbriodendron* (Christ) O.E. Schulz in the Canary Islands [Cruciferae]. *Botanica Macaronésica* 4: 31–53.
- Brochmann, C. 1984. Hybridization and distribution of *Argyranthemum coronopifolium* (Asteraceae–Anthemideae) in the Canary Islands. *Nordic Journal of Botany* 4: 729–736.
- Carine, M. A., L. Robba, R. Little, S. Russell, and A. S. Guerra. 2007. Molecular and morphological evidence for hybridization between endemic Canary Island *Convolvulus*. *Botanical Journal of the Linnean Society* 154: 187–204.
- Choisy, M., P. Franck, and J.-M. Cornuet. 2004. Estimating admixture proportions with microsatellites: comparison of methods based on simulated data. *Molecular Ecology* 13: 955–968.
- Cornuet, J.-M., F. Santos, M. A. Beaumont, C. P. Robert, J.-M. Marin, D. J. Balding, T. Guillemaud, and A. Estoup. 2008. Inferring population history with DIY ABC: a user-friendly approach to approximate Bayesian computation. *Bioinformatics* 24: 2713–2719.
- Darriba, D., D. Posada, A. M. Kozlov, A. Stamatakis, B. Morel, and T. Flouri. 2020. ModelTest-NG: A new and scalable tool for the selection of DNA and protein evolutionary models. *Molecular Biology and Evolution* 37: 291–294.
- Doyle, J. 1991. DNA protocols for plants. In: G. M. Hewitt, A. W. B. Johnston, and J. P. W. Young [eds.], *Molecular techniques in taxonomy*, NATO ASI Series, 283–293. Springer, Berlin, Germany.
- Eaton, D. A., A. L. Hipp, A. González-Rodríguez, and J. Cavender-Bares. 2015. Historical introgression among the American live oaks and the comparative nature of tests for introgression. *Evolution* 69: 2587–2601.
- Eaton, D. A., and R. H. Ree. 2013. Inferring phylogeny and introgression using RADseq data: an example from flowering plants (*Pedicularis*: Orobanchaceae). *Systematic Biology* 62: 689–706.
- Eaton, D. A. R. 2014. PyRAD: assembly of de novo RADseq loci for phylogenetic analyses. *Bioinformatics* 30: 1844–1849.
- Elshire, R. J., J. C. Glaubitz, Q. Sun, J. A. Poland, K. Kawamoto, E. S. Buckler, and S. E. Mitchell. 2011. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One* 6: e19379.
- Escudero, M., D. A. R. Eaton, M. Hahn, and A. L. Hipp. 2014. Genotyping-by-sequencing as a tool to infer phylogeny and ancestral hybridization: A case study in *Carex* (Cyperaceae). *Molecular Phylogenetics and Evolution* 79: 359–367.
- ESRI. 2011. ArcGIS Desktop: Release 10. Redlands, CA: Environmental Systems Research Institute.
- Fernández-Mazuecos, M., G. Mellers, B. Vigalondo, L. Sáez, P. Vargas, and B. J. Glover. 2018. Resolving recent plant radiations: power and robustness of genotyping-by-sequencing. *Systematic Biology* 67: 250–268.
- Fick, S. E., and R. J. Hijmans. 2017. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology* 37: 4302–4315.
- Fjellheim, S., M. H. Jørgensen, M. Kjos, and L. Borgen. 2009. A molecular study of hybridization and homoploid hybrid speciation in *Argyranthemum* (Asteraceae) on Tenerife, the Canary Islands. *Botanical Journal of the Linnean Society* 159: 19–31.
- Francisco-Ortega, J., J. Fuertes-Aguilar, S.-C. Kim, A. Santos-Guerra, D. J. Crawford, and R. K. Jansen. 2002. Phylogeny of the Macaronesian endemic *Crambe* Section *Dendrocrambe* (Brassicaceae) based on internal transcribed spacer sequences of nuclear ribosomal DNA. *American Journal of Botany* 89: 1984–1990.
- Francisco-Ortega, J., R. K. Jansen, and A. Santos-Guerra. 1996. Chloroplast DNA evidence of colonization, adaptive radiation, and hybridization in the evolution of the Macaronesian flora. *Proceedings of the National Academy of Sciences* 93: 4085–4090.
- Francisco-Ortega, J., A. Santos-Guerra, A. Hines, and R. K. Jansen. 1997. Molecular evidence for a Mediterranean origin of the Macaronesian endemic genus *Argyranthemum* (Asteraceae). *American Journal of Botany* 84: 1595–1613.
- Francisco-Ortega, J., A. Santos-Guerra, S.-C. Kim, and D. J. Crawford. 2000. Plant genetic diversity in the Canary Islands: a conservation perspective. *American Journal of Botany* 87: 909–919.
- Goodson, B. E., S. K. Rehman, and R. K. Jansen. 2011. Molecular systematics and biogeography of *Descurainia* (Brassicaceae) based on nuclear ITS and non-coding chloroplast DNA. *Systematic Botany* 36: 957–980.
- Goodson, B. E., A. Santos-Guerra, and R. K. Jansen. 2006. Molecular systematics of *Descurainia* (Brassicaceae) in the Canary Islands: biogeographic and taxonomic implications. *Taxon* 55: 671–682.
- Graham, R. E., J. A. Reyes-Betancort, M. A. Chapman, and M. A. Carine. 2021. Inter-island differentiation and contrasting patterns of diversity in the iconic Canary Island sub-alpine endemic *Echium wildpretii* (Boraginaceae). *Systematics and Biodiversity* 19: 507–525.
- Green, R. E., J. Krause, A. W. Briggs, T. Maricic, U. Stenzel, M. Kircher, N. Patterson, et al. 2010. A draft sequence of the Neandertal genome. *Science* 328: 710–722.
- Guerra, A., J. Reyes-Betancort, M. Carine, C. Humphries, and I. Guma. 2008. Diversity, rarity and the evolution and conservation of the Canary Islands endemic flora. *Anales del Jardín Botánico de Madrid* 65: 25–45.
- Harter, D. E. V., M. Thiv, A. Weig, A. Jentsch, and C. Beierkuhnlein. 2015. Spatial and ecological population genetic structures within two island-endemic *Aeonium* species of different niche width. *Ecology and Evolution* 5: 4327–4344.
- Herben, T., V. Rydlová, T. Fér, J. Suda, Z. Münzbergová, R. Wildová, and J. Wild. 2014. Phylogenetic signal in growth and reproductive traits and in their plasticity: the *Descurainia* radiation in the Canary Islands. *Botanical Journal of the Linnean Society* 174: 384–398.
- Hibbins, M., and M. Hahn. 2021. Phylogenomic approaches to detecting and characterizing introgression. *Genetics* 220: iyab173.
- Holder, M. T., J. A. Anderson, and A. K. Holloway. 2001. Difficulties in detecting hybridization. *Systematic Biology* 50: 978–982.
- Hua, X., and J. J. Wiens. 2013. How does climate influence speciation? *The American Naturalist* 182: 1–12.
- Hudson, R. R., M. Slatkin, and W. P. Maddison. 1992. Estimation of levels of gene flow from DNA sequence data. *Genetics* 132: 583–589.
- Irl, S. D. H., D. E. V. Harter, M. J. Steinbauer, D. G. Puyol, J. M. Fernández-Palacios, A. Jentsch, and C. Beierkuhnlein. 2015. Climate vs. topography—spatial patterns of plant species diversity and endemism on a high-elevation island. *Journal of Ecology* 103: 1621–1633.
- Jaén-Molina, R., Á. Marrero-Rodríguez, J. Caujapé-Castells, and D. I. Ojeda. 2021. Molecular phylogenetics of *Lotus* (Leguminosae) with emphasis in the tempo and patterns of colonization in the Macaronesian region. *Molecular Phylogenetics and Evolution* 154: 106970.
- Jones, K. E., J. A. Reyes-Betancort, S. J. Hiscock, and M. A. Carine. 2014. Allopatric diversification, multiple habitat shifts, and hybridization in the evolution of *Pericallis* (Asteraceae), a Macaronesian endemic genus. *American Journal of Botany* 101: 637–651.
- Jørgensen, T. H., and J. M. Olesen. 2001. Adaptive radiation of island plants: evidence from *Aeonium* (Crassulaceae) of the Canary Islands. *Perspectives in Plant Ecology, Evolution and Systematics* 4: 29–42.
- Juan, C., B. C. Emerson, P. Oromí, and G. M. Hewitt. 2000. Colonization and diversification: towards a phylogeographic synthesis for the Canary Islands. *Trends in Ecology & Evolution* 15: 104–109.
- Kim, M., M. L. Cui, P. Cubas, A. Gillies, K. Lee, M. A. Chapman, R. J. Abbott, and E. Coen. 2008. Regulatory genes control a key morphological and ecological trait transferred between species. *Science* 322(5904): 1116–1119.
- Kim, S. C., D. J. Crawford, J. Francisco-Ortega, and A. Santos-Guerra. 1996. A common origin for woody *Sonchus* and five related genera in the Macaronesian islands: molecular evidence for extensive radiation. *Proceedings of the National Academy of Sciences* 93: 7743–7748.
- McCluskey, B. M., and J. H. Postlethwait. 2015. Phylogeny of Zebrafish, a “Model Species,” within *Danio*, a “Model Genus”. *Molecular Biology and Evolution* 32: 635–652.

- Mort, M. E., D. E. Soltis, P. S. Soltis, J. Francisco-Ortega, and A. Santos-Guerra. 2002. Phylogenetics and evolution of the Macaronesian clade of Crassulaceae inferred from nuclear and chloroplast sequence data. *Systematic Botany* 27: 271–288.
- Patterson, N., P. Moorjani, Y. Luo, S. Mallick, N. Rohland, Y. Zhan, T. Genschoreck, et al. 2012. Ancient admixture in human history. *Genetics* 192: 1065–1093.
- Pérez-Escobar, O. A., D. Bogarín, R. Schley, R. M. Bateman, G. Gerlach, D. Harpke, J. Brassac, et al. 2020. Resolving relationships in an exceedingly young Neotropical orchid lineage using genotyping-by-sequencing data. *Molecular Phylogenetics and Evolution* 144: 106672.
- Puppo, P., M. Curto, J. Gusmão-Guedes, J. Cochofel, P. L. Pérez de Paz, C. Bräuchler, and H. Meimberg. 2015. Molecular phylogenetics of *Micromeria* (Lamiaceae) in the Canary Islands, diversification and inter-island colonization patterns inferred from nuclear genes. *Molecular Phylogenetics and Evolution* 89: 160–170.
- Revell, L. J. 2012. phytools: an R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution* 3: 217–223.
- Reyes-Betancort, J. A., A. S. Guerra, I. R. Guma, C. J. Humphries, and M. A. Carine. 2008. Diversity, rarity and the evolution and conservation of the Canary Islands endemic flora. *Anales del Jardín Botánico de Madrid* 65: 25–45.
- Rieseberg, L., and D. Soltis. 1991. Phylogenetic consequences of cytoplasmic gene flow in plants. *Evolutionary Trends in Plants* 5: 65–84.
- Rieseberg, L. H., and J. H. Willis. 2007. Plant speciation. *Science* 317: 910–914.
- Rokas, A., B. L. Williams, N. King, and S. B. Carroll. 2003. Genome-scale approaches to resolving incongruence in molecular phylogenies. *Nature* 425: 798–804.
- Rowe, H. C., S. Renaut, and A. Guggisberg. 2011. RAD in the realm of next-generation sequencing technologies. *Molecular Ecology* 20: 3499–3502.
- Salichos, L., and A. Rokas. 2013. Inferring ancient divergences requires genes with strong phylogenetic signals. *Nature* 497: 327–331.
- Sankararaman, S., S. Mallick, M. Dannemann, K. Prüfer, J. Kelso, S. Pääbo, N. Patterson, and D. Reich. 2014. The genomic landscape of Neanderthal ancestry in present-day humans. *Nature* 507: 354–357.
- Saunders, N. E., and D. J. Gibson. 2005. Breeding system, branching processes, hybrid swarm theory, and the humped-back diversity relationship as additional explanations for apparent monophyly in the Macaronesian island flora. *Journal of Ecology* 93: 649–652.
- Schmincke, H.-U. 1976. The geology of the Canary Islands. In: G. Kunkel [ed.], *Biogeography and ecology in the Canary Islands*, Monographiae Biologicae, 67–184. Springer, Dordrecht, Netherlands.
- Silvertown, J. 2004. The ghost of competition past in the phylogeny of island endemic plants. *Journal of Ecology* 92: 168–173.
- Silvertown, J., J. Francisco-Ortega, and M. A. Carine. 2005. The monophyly of island radiations: an evaluation of niche pre-emption and some alternative explanations. *Journal of Ecology* 93: 653–657.
- Stamatakis, A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30: 1312–1313.
- Steinbauer, M. J., O. Rüdiger, A. Naranjo-Cigala, C. Beierkuhnlein, and J.-M. Fernández-Palacios. 2012. Increase of island endemism with altitude–speciation processes on oceanic islands. *Ecography* 35: 23–32.
- Steinbauer, M. J., R. Field, J. M. Fernández-Palacios, S. D. H. Irl, R. Otto, H. Schaefer, and C. Beierkuhnlein. 2016. Biogeographic ranges do not support niche theory in radiating Canary Island plant clades. *Global Ecology and Biogeography* 25: 792–804.
- Than, C., D. Ruths, and L. Nakhleh. 2008. PhyloNet: a software package for analyzing and reconstructing reticulate evolutionary relationships. *BMC Bioinformatics* 9: 322.
- van Hengstum, T., S. Lachmuth, J. G. B. Oostermeijer, H. (J.) C. M. den Nijs, P. G. Meirns, and P. H. van Tienderen. 2012. Human-induced hybridization among congeneric endemic plants on Tenerife, Canary Islands. *Plant Systematics and Evolution* 298: 1119–1131.
- Vernot, B., and J. M. Akey. 2014. Resurrecting surviving Neandertal lineages from modern human genomes. *Science* 343: 1017–1021.
- Villa-Machío, I., A. G. Fernández de Castro, J. Fuertes-Aguilar, and G. Nieto Feliner. 2020. Colonization history of the Canary Islands endemic *Lavatera acerifolia* (Malvaceae) unveiled with genotyping-by-sequencing data and niche modelling. *Journal of Biogeography* 47: 993–1005.
- Warren, B. H., D. Simberloff, R. E. Ricklefs, R. Aguilée, F. L. Condamine, D. Gravel, H. Morlon, et al. 2015. Islands as model systems in ecology and evolution: prospects fifty years after MacArthur-Wilson. *Ecology Letters* 18: 200–217.
- White, O. W., A. Reyes-Betancort, M. A. Chapman, and M. A. Carine. 2018. Independent homoploid hybrid speciation events in the Macaronesian endemic genus *Argyranthemum*. *Molecular Ecology* 27: 4856–4874.
- White, O. W., J. A. Reyes-Betancort, M. A. Chapman, and M. A. Carine. 2020. Geographical isolation, habitat shifts and hybridisation in the diversification of the Macaronesian endemic genus *Argyranthemum* (Asteraceae). *New Phytologist* 228: 1953–1971.
- Whitney, K. D., R. A. Randell, and L. H. Rieseberg. 2006. Adaptive introgression of herbivore resistance traits in the weedy sunflower *Helianthus annuus*. *The American Naturalist* 167: 794–807.

SUPPORTING INFORMATION

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

APPENDIX S1. Twenty Canary Island *Descurainia* individuals from eight species, and their origin, which were used within this study. Field-collected leaf material was dried in silica gel, and voucher specimens were deposited at the Natural History Museum (BM), London. Origin abbreviation as follows: GC = Gran Canaria, TEN = Tenerife, LP = La Palma, SP = Spain, SA = South America.

APPENDIX S2. Most likely Maximum-Likelihood model for all six genotyping-by-sequencing (GBS) assemblies using ModelTest-NG.

APPENDIX S3. Maximum-likelihood phylogenies using 18 individuals of Canary Island *Descurainia* and two outgroups from six assembly parameters using genotyping-by-sequencing (GBS) data. Bootstrap values are next to nodes in red. Branch length bar above phylogeny. Individual ID numbers are represented in Appendix S1.

APPENDIX S4. The mean, median, mode and quantiles of posterior distribution samples for effective population sizes (parameter N), time (t) and admixture (r) for the scenario tested in Approximate Bayesian computation (ABC) analysis generated by DIYABC (Cornuet et al., 2008).

APPENDIX S5. Averages for bioclimatic variables for each the seven species across 217 individuals of Canary Island *Descurainia*.

How to cite this article: Jackson, A. C., O. W. White, M. Carine, and M. A. Chapman. 2023. The role of geography, ecology, and hybridization in the evolutionary history of Canary Island *Descurainia*. *American Journal of Botany* 110(5): e16162. <https://doi.org/10.1002/ajb2.16162>