Geographical isolation, habitat shifts and hybridisation in the diversification of the Macaronesian endemic genus *Argyranthemum* (Asteraceae)

Oliver W. White[1, 2, 3], J. Alfredo Reyes-Betancort[4], Mark A. Chapman[2] and Mark A. Carine[1]

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

[2] Biological Sciences, University of Southampton, Southampton, SO17 1BJ, United Kingdom.

[3] Royal Botanic Gardens, Kew, Richmond, Surrey TW9 3AE.

[4] Jardín de Aclimatación de La Oratava (Instituto Canario de Investigaciones Agrarias - ICIA), C/ Retama 2, 38400 Puerto de la Cruz, Tenerife, Spain

ORCID iDs

Oliver White 0000-0001-6444-0310
Mark A. Chapman 0000-0002-7151-723X

Social media (Twitter)

Olive W. White @Ollie_W_White
J. Alfredo Reyes-Betancort @AlfredoJAOICIA
Mark A. Chapman @CUCsoton
Mark A. Carine @mark_carine

Corresponding author: Mark A. Carine (m.carine@nhm.ac.uk, 02079425541)

Summary	200
Introduction	1,112
Material and Methods	2,373
Results	1,481
Discussion	1,523
Main body	6,489
Number of Figures	5
Number of Tables	4
Supporting Information	Supporting-Information-Table-S8 Supporting-Information-Table-S9 Supporting-Information-Methods-S2

Summary

- Inferring the processes responsible for the rich endemic diversity of oceanic island floras is important for our understanding of plant evolution and setting practical conservation priorities. This requires an accurate knowledge of phylogenetic relationships, which have often been difficult to resolve due to a lack of genetic variation.
- We employed genotyping-by-sequencing (GBS) to investigate how geographical isolation, habitat shifts, and hybridisation have contributed to the evolution of diversity observed in *Argyranthemum* Webb (Asteraceae), the largest genus of flowering plants endemic to the Macaronesian archipelagos.
- Species relationships were resolved, and biogeographical stochastic mapping identified intra-island speciation as the most frequent biogeographic process underlying diversification, contrary to prevailing view in *Argyranthemum* and the Canary Islands. D-statistics revealed significant evidence of hybridisation between lineages co-occurring on the same island, however there was little support for the hypothesis that hybridisation may be responsible for the occurrence of non-monophyletic multi-island endemic (MIE) species.
- Geographic isolation, habitat shifts, and hybridisation have all contributed to the diversification of *Argyranthemum*, with intra-island speciation found to be more frequent than previously thought. Morphological convergence is also proposed to explain the occurrence of non-monophyletic MIE species. This study reveals greater complexity in the evolutionary processes generating Macaronesian endemic diversity.

Keywords

Argyranthemum, genotyping-by-sequencing, biogeography, biogeographic stochastic mapping, hybridisation, Macaronesia, morphological convergence, speciation

Introduction

Investigations of the processes responsible for the unique assemblage of flowering plants on oceanic archipelagos are informative for our understanding of plant evolution and may also inform practical conservation policies (Caujapé-Castells *et al.*, 2010; Bramwell, 2011). The Macaronesian archipelagos of the Azores, Madeira, Selvagens, the Canary Islands and the Cape Verdes in the North Atlantic Ocean have been considered a region ideally suited to investigations of the processes responsible for flowering plant evolution (Kim *et al.*, 2008). Macaronesia is home to approximately 30 endemic genera and 900 endemic species of flowering plants (Caujapé-Castells *et al.*, 2010) and diversification of lineages within the region has played a prominent role in generating the striking levels of endemic diversity.

Hypotheses to explain diversification of plant lineages within Macaronesia and other such archipelagos have largely focussed on geographical isolation (i.e. speciation through isolation following inter-island dispersal between similar ecological niches) and habitat shifts (i.e. speciation associated with the shifts of a lineage to different ecological niches; Francisco-Ortega *et al.*, 2002; Lee *et al.*, 2005; Goodson *et al.*, 2006). Hybridisation has also been recognised as an important process for diversification in oceanic archipelagos, although much less is known about its frequency and evolutionary significance (Jones *et al.*, 2014; Curto *et al.*, 2017). Instances of hybridisation have been inferred from incongruence between chloroplast (cp) and nuclear DNA datasets (Mort *et al.*, 2002; Barber *et al.*, 2007; Jones *et al.*, 2014) or quantified through the use of D-statistics (ABBA-BABA tests) that can discern between incomplete lineage sorting and hybridisation (Eaton & Ree, 2013; Curto *et al.*, 2017).

A significant impediment to our understanding of the relative contribution of these processes in generating the observed diversity on oceanic islands has been the lack of phylogenetic resolution in studies based typically on sequencing the nuclear internal transcribed spacer region (ITS) or analysing one or a few cp loci (Francisco-Ortega *et al.*, 1997b; Mort *et al.*, 2002; Allan *et al.*, 2004; Trusty *et al.*, 2005; Goodson *et al.*, 2006; Barber *et al.*, 2007). To overcome this, recent studies have employed high-throughput sequencing (HTS) approaches, particularly reduced representation style data such as restriction site associated markers (RAD-seq) and genotyping-by-sequencing (GBS; Baird *et al.*, 2008; Elshire *et al.*, 2011). These methods are advantageous as they identify thousands of polymorphic markers across multiple samples in non-model species without the need for a reference

genome. A caveat of these methods is that RAD-seq or GBS loci are typically too short and lack sufficient variation to be analysed individually. However, the loci generated by these methods may be concatenated to create a supermatrix containing thousands of polymorphic sites and have been employed successfully in the resolution of phylogenetic relationships between closely related taxa in Macaronesia, notably *Tolpis* Adans. (Asteraceae; Mort *et al.*, 2015) and *Micromeria* Benth. (Lamiaceae; Puppo *et al.*, 2015), as well as taxa in other island systems including *Dyospiros* (Ebenaceae; Paun *et al.*, 2016). In each of these examples, the reduced representation datasets considerably improved the resolution of species relationships and allowed greater inference of the processes responsible for diversification.

Argyranthemum Webb (Asteraceae) is the largest endemic genus of flowering plants found in the Macaronesian archipelagos, with a total of 24 species and 39 terminal taxa (i.e. species and subspecies;

Fig. 1). Three species are endemic to Madeira, one to Selvagem Pequena and twenty to the Canary Islands. Twenty-one species are single island endemics (SIEs) and three are multiple island endemics (MIEs; *A. frutescens*, *A. broussonetii* and *A. adauctum*). Each MIE is comprised of SIE subspecies. *Argyranthemum* is present in all the major habitat zones in Macaronesia, ranging from coastal to subalpine habitats.

Previous phylogenetic analyses using small numbers of molecular markers have found that *Argyranthemum* is monophyletic and closely related to the continental genera *Glebionis* Cass., *Heteranthemis* Schott and *Ismelia* Cass. that are distributed in the Mediterranean, Southern Iberia and Morocco, respectively (Francisco-Ortega *et al.*, 1995a,b; Oberprieler *et al.*, 2007). However, attempts to resolve species relationships within *Argyranthemum* have been hampered by a lack of genetic variation (Francisco-Ortega *et al.*, 1997a). A cp restriction site analysis identified two main clades, one restricted to Madeira and the Selvagens and one comprising taxa from the Canary Islands (Francisco-Ortega *et al.*, 1996b). Within the Canary Islands clade, two major groups were resolved, one largely corresponding to taxa occupying habitats under the influence of the Northern trade winds and the other not, suggesting that inter-island colonisation between similar habitats was the prominent driver of diversification in the Canary Islands (Francisco-Ortega *et al.*, 1996b). In contrast, habitat shifts were more frequent on Madeira where there was a single colonisation event followed by divergence into different habitat types.

Hybridisation is also likely to have played a significant role in the diversification of *Argyranthemum*. Intrinsic reproductive barriers between taxa are weak, with reproductive isolation largely due to geographical and ecological differentiation (Francisco-Ortega *et al.*, 1997a). Artificial hybrids can be created with ease in cultivation (Humphries, 1973) and hybrid swarms have been documented between *A. broussonetii* and *A. frutescens* (Tenerife; Brochmann, 1987), *A. coronopifolium* and *A. frutescens* (Tenerife; Brochmann, 1984), *A. adauctum* and *A. filifolium* (Gran Canaria; Borgen, 1976) and *A. tenerifae* and *A. adauctum* (Tenerife; White, pers. obs.). Hybridisation has also been inferred from the presence of polyphyletic taxa in previous phylogenetic analyses, notably *A. adauctum* (Francisco-Ortega *et al.*, 1996b). It has also been hypothesised that *A. escarrei* is of hybrid origin based on morphological similarity to individuals collected from hybrid swarms between *A. adauctum* subsp. *canariense* and *A. filifolium* (Borgen, 1976), although this is not supported by preliminary data (O. White, M.A. Chapman and M.A. Carine, unpublished data). To date, the most robust support for the role of hybridisation in the diversification of *Argyranthemum* relates to *A. sundingii* and *A. lemsii*, for which molecular data clearly support independent homoploid hybrid origins from crosses between *A. broussonetii* and *A. frutescens* (Brochmann *et al.*, 2000; Fjellheim *et al.*, 2009; White *et al.*, 2018).

The aim of this study was to investigate the role of geographical isolation, habitat shifts and hybridisation in the evolution of *Argyranthemum*. To this end, we present a phylogenetic analysis of *Argyranthemum* using GBS and sampling across the entire genus. Ancestral ranges for geography and habitats are inferred and biogeographic stochastic mapping is performed to estimate the frequency of biogeographic events. To detect evidence of hybridisation we employed D-statistics (Eaton & Ree, 2013). We specifically tested two hypotheses to assess the role of hybridisation in the diversification of Canary Island lineages: first, that hybridisation between species on the same island has been frequent, as proposed for *Micromeria* (Curto *et al.*, 2017), and second, that hybridisation explains the non-monophyly of multi-island endemic (MIE) species as proposed for *Pericallis* (Jones *et al.*, 2014).

Materials and Methods

Sampling

Field collections of *Argyranthemum* were made in the Canary Islands and Madeira from July to August 2015 and March 2016 respectively. Outgroup samples of *Glebionis* were collected in

Andalucía, Spain during April 2015. Leaf material was dried and preserved using silica gel for molecular work. Herbarium vouchers were also prepared and deposited at the Natural History Museum, London (BM). Material from the Selvagens was sampled from the University of Madeira Herbarium.

Samples of all *Argyranthemum* taxa recognised by Humphries (1976) and Francisco-Ortega *et al.* (1996b) were included in our study (Table 1) with the exception of *A. sundingii* and *A. lemsii* which are of hybrid origin (Brochmann *et al.*, 2000; Fjellheim *et al.*, 2009; White *et al.*, 2018) and likely to confound phylogenetic analyses (Gruenstaeudl *et al.*, 2017; McVay *et al.*, 2017). "*Argyranthemum vincentii*" has been identified in previous studies as an endemic to Tenerife and restricted to pine forest habitats (Francisco-Ortega *et al.*, 1996c, 2000) and can be distinguished from closely-related taxa by filiform leaf lobes and green (not glaucous) leaves (pers. com., A. Santos-Guerra). This taxon has yet to be formally described but was nevertheless included Where possible we included two samples for each terminal taxon, ideally from different populations. However, only one sample could be included for *A. haematomma*, *A. thalassophilum* and for each subspecies of *A. pinnatifidum*. Additionally, for each of the following taxa, two individuals from the same population were sampled: *A. adauctum* subsp. *dugourii*, subsp. *erythrocarpon*, subsp. *palmensis*, *A. frutescens* subsp. *canariae*, *A. lidii*, *A. sventenii*, *"A. vincentii"* and *A. winteri*. The outgroup taxa, *Glebionis segetum* and *G. coronaria* from the Mediterranean region, were selected on the basis of earlier phylogenetic analyses (Francisco-Ortega *et al.*, 1995b,a; Oberprieler *et al.*, 2007).

DNA isolation and GBS

DNA was extracted from silica-dried leaf material using a modified CTAB method (Doyle & Doyle 1987) identical to the method used by White *et al.* (2016, 2018). DNA samples were sent to the Genomic Diversity Facility at Cornell University for GBS where samples were digested using *Eco*T22I and single-end 100 bp reads were generated using an Illumina Hiseq. Raw sequence data for *A. broussonetii* subsp. *broussonetii*, *A. frutescens* subsp. *frutescens* and subsp. *succulentum* used by White *et al.* (2018) were reanalysed in the present study.

Processing of GBS data

Reads generated by GBS were processed and assembled using ipyrad (version 0.7.15; Eaton & Ree, 2013) employing the same assembly method as White *et al.* (2018). Briefly, the sequence quality of raw GBS reads was assessed and low-quality reads were removed. Samples with less than 500,000

filtered reads were also removed. Reads were clustered at three different similarity thresholds (80%, 85% and 90%) and reads that mapped to the chloroplast genomes of *Arabidopsis thaliana* (L.) Heynh. (Genbank accession number: NC_000932.1), *Helianthus annuus* L. (NC_007977.1) or *Chrysanthemum indicum* L. (NC_020320.1) and the mitochondrial genomes of *A. thaliana* (NC_001282) or *H. annuus* (KF815390.1) were removed. Assembled loci were filtered such that a locus must include 30 or 38 samples, equivalent to 60% and 50% missing data respectively after the removal of low-quality samples (see Results). Additionally, loci with shared heterozygous sites across more than 20% of the samples were removed as likely paralogous loci. Therefore, a total of six assemblies were produced (Supporting Information Table S 1), which were compared by genetic clustering as detailed in the Supporting Information (Methods S 1, Fig. S 1; Fig. S 2; Table S 2).

Phylogenetic reconstruction

Loci were concatenated and missing data were added as Ns to create a supermatrix. The optimal model of sequence evolution was identified using ModelTest-NG v.0.1.3 (https://github.com/ddarriba/modeltest). Maximum likelihood (ML) and Bayesian Inference (BI) trees were generated for each dataset using RAxML Next Generation v.0.5.1 (RAxML-NG; Kozlov *et al.*, 2019) and MrBayes v.3.2.6 (Ronquist *et al.*, 2012), respectively. The ML analyses were performed using the transversion model (TVM) with proportion of invariable sites (I) and a gamma distribution (G) for all six datasets (Supporting Information Table S 3). The best ML tree was selected after 1,000 independent searches and bootstrap values were calculated from 1,000 replicates. The TVM +I+G model was not available in MrBayes so we used the second best supported model which was the general time reversible model (GTR) +I+G (Supporting Information Table S 3) for the Bayesian analyses. The BI analysis had two runs with four chains, 2×10^6 generations and a sampling frequency of 5,000 before computing a 50% majority rule consensus tree. Each run in the MrBayes analyses was found to converge with average standard deviation of split frequencies <0.01 after 2×10^6 generations and effective sample sizes (ESS) ranging from 483 to 635 as reported in Tracer v.1.7.1 (Supporting Information Fig. S 3; Table S 4; Rambaut *et al.*, 2018). For all analyses trees were visualised using ggtree in R (Yu *et al.*, 2017; R Core team, 2020).

Time calibrated species tree estimation

For optimal ancestral range estimation, a time calibrated tree with a single representative per taxon is required, therefore the final step of the ipyrad pipeline was repeated with one representative of

each taxon (excluding outgroup taxa), selected as the one with the least missing data, using a clustering threshold of 90% and a minimum sample number of 16 (equivalent to 60% missing data; see Results). Where taxa were identified as non-monophyletic all samples were included, except for *A. frutescens* subsp. *gracilescens* 179 which appeared to be influenced by hybridisation (see Results). Phylogenetic analysis using RAxML-NG was repeated as above.

Using the fixed topology species tree, taxon divergence times were estimated using MCMCTree in PAML (Yang, 2007) , to provide an ultrametric tree for optimal ancestral range estimation. This method has been shown to be an effective means of estimating divergence times using genomic sequence alignments (dos Reis *et al.*, 2016; Thode *et al.*, 2019). Two calibration points were used. Firstly, the root age of *Argyranthemum* was previously estimated to be ca. 1.5-3.0 MYA (Francisco-Ortega *et al.*, 1997a). Secondly, the age of El Hierro (0.8 MYA) was used as the maximum age for clades endemic to this island (Francisco-Ortega *et al.*, 1997a). Calibration of divergence times for island lineages is challenging and the use of such recent calibration points will result in recent divergence times but the aim, as noted above, is simply to generate an ultrametric tree and not to estimate divergence times or infer evolutionary rates.

The optimal model of sequence evolution identified by ModelTest-NG was the TVM +I+G model. This was not available in MCMCTree so we used the second best supported model which was GTR +I+G. used. Two independent runs in MCMCTree were performed to check for convergence using independent rates model with a log-normal distribution, sampling frequency of 100 for 200,000 samples after a burn-in of 20,000.

Ancestral area and habitat estimation

The distributions of taxa and the habitats they occupy were determined using 960 occurrence records from fieldwork collections (622 records) and high precision data downloaded from online databases (338 records; https://www.biodiversidadcanarias.es/ accessed June 2019). For each taxon, distribution was scored as the island occupied and 'habitat' as vegetation type. A recent vegetation classification was used to assign taxa to potential vegetation types of taxa in the Canary Islands (del Arco Aguilar & Rodríguez-Delgado, 2018). Shape files for the vegetation of the Canary Islands were kindly provided by Marcelino del Arco Aguilar (pers. comm., June 2019). Vegetation types recognised by del Arco Aguilar & Rodríguez-Delgado (2018) were grouped into: (E) *Euphorbia* scrubland, (T) thermo-sclerophyllous woodland, (L) laurel forest, (P) pine forest and (S) subalpine zone (Supporting Information Table S 5) and occurrence records were scored for these zones using ArcGIS (ESRI, 2019).

Some occurrence records occurred in vegetation classes such as rocky habitats or canary palm groves that span several of the zones defined here. In these instances, the zone in which those habitats occurred was identified and scored to minimise the number of associated parameters. To minimise the impact of erroneous occurrence data, a taxon was only scored for a habitat zone if a certain proportion of occurrences for the taxon were present. To ensure our analyses were not sensitive to this threshold, we varied the threshold used to filter erroneous data from 0% to 20% in steps of 5%. We selected a threshold of 10% because this identified a best supported model which was largely consistent with the other thresholds tested (Supporting Information Table S 6). In addition, there was no obvious threshold at which the number states per taxon plateaued (Supporting Information Fig. S 4-Fig. S 5).

The classification by del Arco Aguilar & Rodríguez-Delgado (2018) does not include Madeira or Selvagem Pequena and the six taxa from these islands were scored using the same categories based on current taxonomic knowledge (Supporting Information Table S 7).

To investigate the biogeographic history of *Argyranthemum*, we estimated ancestral ranges for islands and habitats using the R package BioGeoBEARS (Matzke, 2014). This package allows the comparison of three widely used biogeographic models, namely DEC (dispersal-extinction-cladogenesis; Ree & Smith, 2008), DIVA (dispersal-vicariance analysis; Ronquist, 1997) and BayArea (Bayesian Inference of Historical Biogeography for Discrete Areas; Landis *et al.*, 2013). These models were originally developed in different frameworks: likelihood for DEC, parsimony for DIVA, and Bayesian for BayAREA. However, BioGeoBEARS implements each in a likelihood framework allowing direct comparison and model selection (Dupin *et al.*, 2017). All models in BioGeoBEARS can also be implemented with an additional parameter (J) which accounts for founder event speciation, where jump dispersal events result in new genetically isolated lineages. However, the DEC+J model has been shown to be inappropriate for founder event speciation or for statistical comparison (Ree & Sanmartín, 2018), hence we refrained from using it. The model with the best fit was selected based on log-likelihood values and the corrected Akaike Information Criterion model weights (Matzke, 2014). Using the models with the strongest support for geographical and habitat data, we also performed biogeographical stochastic mapping (BSM) in BioGeoBEARS to investigate the number and type of biogeographical events taking place in the evolution of *Argyranthemum*. A total of 50 biogeographic maps were used. The method identifies anagenetic dispersal events (range contraction or expansion) and cladogenetic events including intra-range speciation (narrow or subset) and vicariance (Fig. 2a; Dupin *et al.*, 2017; Pérez-Escobar *et al.*, 2017). In the context of the

oceanic island setting investigated here, we interpret vicariance as inter-island allopatric speciation for geography and as habitat shifts for habitat data. For ancestral range estimation in BioGeoBEARS, the maximum range sizes for geographic and habitat distributions were three (out of nine) and five (out of five) respectively.

D-statistics

We employed D-statistics as outlined by Eaton and Ree (2013) to test for evidence of hybridisation. Each test takes a four taxon pectinate tree denoted by (((P1,P2),P3),O) and identifies incongruent ancestral (A) and derived (B) alleles denoted as ABBA or BABA (Fig. 2b,c). If ILS is responsible for the incongruence, the proportion of ABBA and BABA alleles will be equal; however, if P3 has hybridised with either P2 or P1, we would expect an asymmetry in the number of ABBA or BABA alleles (Fig. 2b,c). The D statistic quantifies the asymmetry of ABBA and BABA allele frequencies. Following Eaton and Ree (2013), we performed 1,000 bootstrap iterations to measure the standard deviation of the D-statistic, in which loci were re-sampled with replacement to the same number as in the original dataset. The results are reported as Z scores, where Z is the number of standard deviations from 0 (the expected value) for D. Significance is assessed by converting the Z-score into a two-tailed p-value and using 0.01 as a conservative cut-off after correcting for multiple comparisons using Holm-Bonferroni correction.

We first used D-statistics to test for evidence of hybridisation between clades found on the same island (Fig. 2d). Seventeen clades were identified as co-occurring on one of the Canary Islands, with two clades on El Hierro, three on La Palma, two on La Gomera, seven on Tenerife and three on Gran Canaria (Supporting Information Fig. S 6). A total of 29 tests were implemented (Table 4): one on El Hierro (test 1), three on La Palma (tests 2-4), one on La Gomera (test 5), 21 on Tenerife (tests 6-26) and three on Gran Canaria (tests 27-29).

Second, we used D-statistics to test the hypothesis that hybridisation explains the polyphyly of the multi-island endemic species *A. adauctum* and *A. broussonetii* (See Results; Fig. 2e). Four tests were implemented (Table 4): one for *A. broussonetii* that was resolved in two distinct clades (test 30) and three for *A. adauctum* that was resolved in three distinct clades (tests 31-33). We did not test for hybridisation in *A. frutescens* as this was resolved as paraphyletic rather than polyphyletic, with other taxa nested within this species (see Results).

If hybridisation occurred between close relatives not included in the test, it is possible for D-statistics to produce type 1 errors (Eaton & Ree, 2013). Therefore, for any D-statistic which gave a significant result, we additionally performed a five-taxon partitioned D-statistic, to identify more precisely which of the P3 taxa contributed to hybridisation. The partitioned D-statistic is an extension of the D-statistic, where the taxa at the P3 position are separated into $P3_1$ and $P3_2$ giving the overall tree topology of (((P1,P2),($P3_1$,$P3_2$)),O). Similar to the four taxon D-statistic, the asymmetry in allele patterns is measured, with the exception that the test identifies whether the derived allele is present in $P3_1$ and not $P3_2$, present in $P3_2$ and not $P3_1$ or present in both. These allele patterns correspond to three D-statistics for each scenario: D_1, D_2 and D_{12} respectively (Eaton & Ree, 2013). For D-statistics which had a significant result, partitioned D-statistics were implemented by separating the individuals at P3 into all pairwise combinations. For the nineteen D-statistics that were significant (see Results), 183 partitioned D-statistics were performed (Supporting Information Table S 8). Significance is assessed using the same method as above.

For all tests we used outgroup (O) accessions from the Madeira and Selvagem Pequena clade which were resolved as sister to the Canary Islands clade (see Results). We used the dataset based on clustering threshold of 90% and a minimum sample number of 30 (see justification in Results; Phylogenetic reconstruction). The methodology for running D-statistics in ipyrad can be found in a Jupyter notebook (https://jupyter.org/) within Supporting Information Methods S 2.

Results

Processing of GBS data

Approximately 219 M raw reads were generated using GBS with an average of 2.65 M per sample (range 0.02-9.27 M). Filtering of low-quality reads removed 5.71-11.25% of the total reads per sample resulting in an average of 2.45 M reads retained across all samples (0.02-8.44 M). Seven samples with less than 0.5 M reads were removed (Table 1; Supporting Information Fig. S 7), leaving a total of 76 samples. An average of 2237 (0.08%) reads per sample mapped to either the mitochondrial or chloroplast reference genomes and were excluded (Supporting Information Table S 9).

Increasing the similarity threshold for *de novo* clustering resulted in a higher number of clusters and consensus sequences. The average number of clusters per individual were 23,745, 26,497 and 32,946 for thresholds of 80%, 85% and 90% respectively (Supporting Information Table S 9). The

equivalent numbers for the average number of consensus sequences were 20,583, 23,706 and 30,865 respectively (Supporting Information Table S 9). Increasing the minimum number of samples required to include a locus from 30 (60% missing data) to 38 (50% missing data) reduced the average number of loci across samples included in the assembly by 32.0-34.2% (Supporting Information Table S 9).

Phylogenetic reconstruction

Seven main clades were consistently recovered across all datasets in both ML and BI analyses with very high or maximal ML bootstrap (BS) and Bayesian posterior probability (PP) (ML: Supporting Information Fig. S 8-Fig. S 13 and BI: Fig. S 14-Fig. S 19). Clade A contains all taxa from Madeira and Selvagem Pequena. Clade B comprises accessions of the Tenerife endemic *A. broussonetii* subsp. *broussonetii*. Clade C includes the multi-island endemic *A. frutescens* which is resolved as paraphyletic with respect to *A. gracile* and *"A. vincentii"* from Tenerife. Clade D includes *A. tenerifae*, *A. adauctum* subsp. *adauctum* and subsp. *dugourii* from Tenerife. Clade E comprises all accessions from Gran Canaria, excluding *A. frutescens* subsp. *canariae*. The Tenerife endemics *A. coronopifolium* and *A. foeniculaceum* are grouped together in clade F. Finally, clade G is composed of taxa endemic to the Eastern and Western Canary Islands.

Whilst clades A-G were consistently recovered in all analyses, the relationships between those clades differed in some analyses. For example, clades B and C formed a sister relationship in one ML analysis (Supporting Information Fig. S 12), however, clade B was sister to all other clades for the remaining ML datasets. Similarly, a sister group relationship was recovered between clades D and clade E in ML analyses of three datasets (Supporting Information Fig. S 8, Fig. S 10, Fig. S 12), but not for the remaining three ML datasets where clade D was resolved as sister to clades F and G. In addition, for the ML analyses based on an 80% clustering threshold, one outgroup (*G. coronaria* 797) was resolved as sister to the Madeiran clade (Supporting Information Fig. Fig. S 8, Fig. S 9), although these relationships were poorly supported (BS <70). We carried out all further analyses on the dataset based on a clustering threshold of 90% and minimum sample number of 30 (Fig. 3) because it had the highest number of SNPs and the modal number of PCA and STRUCTURE clusters (Supporting Information Table S 10; Methods S 1).

Our analysis revealed three non-monophyletic species (Fig. 3). Firstly, accessions of *A. broussonetii* were resolved as polyphyletic, with subsp. *broussonetii* from Tenerife resolved as sister to clade C containing the multiple island endemic *A. frutescens*, and subsp. *gomerensis* from La Gomera sister

to *A. callichrysum*, also endemic to La Gomera (Fig. 3). Secondly, accessions of *A. adauctum* were polyphyletic and were resolved in three clades (Fig. 3). The first included *A. adauctum* subsp. *adauctum* and subsp. *dugourii*, resolved in a Tenerife clade (clade D) with *A. tenerifae*. The second was composed of *A. adauctum* subsp. *jacobaeifolium*, subsp. *canariense* and subsp. *gracile* which was resolved in the Gran Canarian clade (clade E). The third clade included *A. adauctum* subsp. *erythrocarpon* and subsp. *palmensis* from El Hierro and La Palma respectively. This was resolved in the Eastern-Western island clade G. Finally, and as noted above, samples of *A. frutescens* were identified as paraphyletic, with *A. gracile* and *"A. vincentii"* nested within the *A. frutescens* clade (clade C). In addition, within *A. frutescens,* subsp. *gracilescens* was resolved as polyphyletic, with one of the two samples (177) sister to the La Gomera endemics *A. frutescens* subsp. *parviflorum* and subsp. *foeniculaceum*, and the other (179) nested within subsp. *frutescens* (Fig. 3).

Time calibrated species tree estimation

After selecting a single representative sample per taxon (except for those identified as being non-monophyletic or, in the case of *A. frutescens* subsp. *gracilescens* 179 potentially influenced by hybridisation for which both samples were retained; see above), a total of 40 samples remained (Supporting information Table S 11). This resulted in a dataset with 7,424 loci and 7,277 SNPs, and phylogenetic analysis using RAxML-NG identified identical taxon relationships as the full dataset (Fig. 3; Supporting information Fig. S 20). Independent runs of MCMCTree found similar node dates suggesting good convergence (Supporting information Fig. S 21). Crown group ages for the main clades identified in *Argyranthemum* were as follows; A=1.72 MYA (highest probability density [HPD]: 1.20-2.22), C=1.75 MYA (HPD: 1.24-2.22), D=1.69 MYA (HPD: 1.19-2.14); E=1.67 MYA (HPD: 1.19-2.1); F=1.59 MYA (HPD: 1.09-2.07) and G=1.84 MYA (HPD: 1.34-2.29).

Ancestral area and habitat estimation

Model comparisons in BiogeoBEARS identified the best supported model for geographic distribution as DIVALIKE (Lnl -64.20; AICc_wt 2.4×10^{-7}; Table 2). For habitat distribution, the best supported model was BAYAREALIKE (Lnl -112.10; AICc_wt 0.76; Table 2), and this was robust to varying the threshold for filtering erroneous data (Supporting Information Table S 6). These models were used for all subsequent analyses (Fig. 4; Supporting Information Fig. S 22).

The most likely ancestral area for Clade A was Madeira (Fig. 4) and an inter-island allopatric speciation event was inferred for the Selvagem Pequena endemic *A. thalassophylum*. Clades B, C, D

and F originated on Tenerife and within clade C there were inter-island allopatric events to Gran Canaria and La Gomera. Inter-island allopatry between Tenerife and Gran Canaria was inferred for clade E (Fig. 4). The ancestral area inferred for clade G is La Palma (Fig. 4) and several inter-island allopatry events within this clade are inferred; El Hierro was colonised twice independently and there was a single colonisation of La Gomera and the Eastern islands (Lanzarote and Fuerteventura). The diversification of clades A-G are predominately associated with habitat range contraction, becoming increasing specialised toward the tips of the tree (Fig. 4).

Based on the biogeographic stochastic mapping analyses of geography, intra-island speciation (narrow; 56.62%) was the most frequent, with inter-island allopatry (vicariance; 22.87%) and range expansion (20.51%) less frequent (Table 3). Note that DIVALIKE models, the most strongly supported model for geography, do not include estimates of intra-range (subset) speciation (Matzke, 2020). For habitats, range contraction (55.99%) was the most frequent event, followed by intra-habitat (narrow; 36.70%) and range contraction (7.30%; Table 3). For BAYAREALIKE models, the most strongly supported for habitats, it is not possible to infer estimates of intra-range (subset) speciation and vicariance.

In terms of dispersal between islands, Tenerife and La Palma act as the main sources with 39.76% and 23.66% of all dispersals occurring from these island respectively (Fig. 5a). There are also some notable differences in directionality. For example, Tenerife never acts as a sink (Fig. 5a) and Gran Canaria is solely a sink (Fig. 5a). For dispersal between habitats, *Euphorbia* scrub is the greatest source of dispersal events (30.15%) and the subalpine zone is the habitat most frequently identified as a sink (42.01%; Fig. 5b).

D-statistics

Nineteen of the 29 D-statistics (ca. 65%) between lineages found on the same island provided evidence of hybridisation (Table 4). Of these, 14 were also significant using the partitioned D-statistic. On El Hierro, there was support for hybridisation between the *A. sventenii*-*A. hierrense* clade and *A. adauctum* subsp. *erythrocarpon* (test 1; Table 4), which was confirmed by a significant result using partitioned D-statistics. For La Palma, there were significant D-statistics supporting hybridisation between: *A. adauctum* subsp. *palmensis* and *A. webbii* (test 3) and *A. haouarytheum* and *A. webbii* (test 4). However, only the former was significant when tested with the partitioned D-statistic. Of the 21 D-statistics performed between clades on Tenerife (tests 6-26), 13 provided significant support for hybridisation and all clades examined were admixed with at least two others. Of these,

11 were also significant using the partitioned D-statistic. On Gran Canaria, there was significant evidence of hybridisation between all three clades based on D-statistics (tests 27-29), but only one of these tests was also significant with partitioned D-statistics.

Between non-monophyletic lineages of MIE species, only one of the four D-statistics performed provided support for evidence of hybridisation, between *A. adauctum* on Tenerife and Gran Canaria (test 31). This test was also significant for the partitioned D-statistic.

Discussion

Previous phylogenetic analyses of *Argyranthemum* based on a few molecular markers suggested that diversification was largely explained by geographical isolation by means of inter-island dispersal between similar habitats (Francisco-Ortega *et al.*, 1996b). Several taxa in *Argyranthemum* were also resolved as non-monophyletic, including the multi-island endemics (MIEs) *A. broussonetii*, *A. adauctum* and *A. frutescens* (Francisco-Ortega *et al.*, 1996a,b). However, phylogenetic inferences in these analyses were limited by a lack of genetic variation and consequentially poor resolution.

In keeping with the analysis of Francisco-Ortega *et al.* (1996b), two major clades corresponding to taxa from (1) Madeira and the Selvagens and (2) the Canary Islands were recovered in the present study. However, the two Canary Island subclades previously identified, and predominately restricted to arid and humid habitats respectively, were not recovered. Rather, the ML analysis resolved two clades composed of (1) *A. broussonetii* subsp. *broussonetii* and the multi island endemic *A. frutescens*, and (2) all reaming taxa from the Canary Islands (Fig. 3).

The ancestral area for the most recent common ancestor of *Argyranthemum* was inconclusive, resolved as either Tenerife and Madeira, and the basal split separates a Canary Island clade from that of Madeira and the Selvagem Pequena. Our results suggest that the Selvagem Pequena was most likely colonised from Madeira. Within the Canary Islands, Tenerife acted as the centre of diversity; a pattern suggested for other Canary island radiations such as *Lotus* (Allan *et al.*, 2004), *Cheirolophus* (Vitales *et al.*, 2014) and *Descurainia* (Goodson *et al.*, 2006). Our biogeographic stochastic mapping analyses identifies Tenerife as the main source of dispersal events (39.76%; Fig. 5), with La Palma also a common source (23.66%; Fig. 5). The colonisation of the Eastern islands of Lanzarote and Fuerteventura from the Western island of El Hierro revealed by our results is counter-intuitive, but was also recovered by Francisco-Ortega *et al.* (1996b). As an alternative to long distance dispersal, Francisco-Ortega *et al.* (1996b) hypothesised that this lineage may have followed

418 a stepping stone colonisation pattern from West to East, with the lineages on the central islands
419 since becoming extinct. As far as we are aware, there are no other Canary Island endemic lineages
420 showing this West to East relationship. However, there are comparable examples of floristic links
421 between the Western and Eastern islands of the Azorean archipelago (e.g. Schaefer *et al.*, 2011).

422 Intra-island speciation was found to be most frequent in our biogeographic analysis of geography,
423 accounting for 56.62% of all events (Table 3). For habitats, biogeographic events were
424 predominantly associated with range contraction (55.99%) and intra-habitat speciation (narrow) is
425 also common (36.70%). The pattern of basal nodes having widespread habitat ranges and becoming
426 increasingly specialised (Fig. 4) was a pattern consistently associated with the BAYAREALIKE model
427 for our analysis. BAYAREALIKE models are typically favoured where there are shared widespread
428 ranges in sister species (for example see Litsios *et al.*, 2014). It should be noted, however, that
429 habitat shifts cannot be inferred with BAYAREALIKE models since they do not estimate vicariance.

430 Taken together, intra-island speciation was found to be predominant in our analysis of geography,
431 contrary to the view that speciation in Canary Island groups is largely driven by inter-island dispersal
432 (Francisco-Ortega *et al.*, 1997a; Baldwin *et al.*, 1999). Within islands, our results do not suggest that
433 habitat shifts are responsible for speciation as suggested for some Macaronesian groups (Barber *et*
434 *al.*, 2007; Gruenstaeudl *et al.*, 2013). Instead, our habitat analysis suggests that the ancestors of
435 *Argyranthemum* were potentially widespread generalists that have become increasingly specialised
436 over time.

437 The topographic complexity of the islands is likely to be an important factor in explaining the
438 prevalence of intra-island speciation processes revealed by our analysis. Such complex landscapes
439 could promote the isolation of lineages at a fine geographic scale leading to narrow endemism.
440 However, we propose that hybridisation could also be an important factor in explaining these
441 patterns. D-statistics (ABBA-BABA tests) suggest that hybridisation between lineages co-occurring on
442 the same islands has been common. D-statistics supported hybridisation in 52% (13/21) of tests
443 performed on Tenerife, two of three tests on La Palma, the single test on El Hierro and all three tests
444 on Gran Canaria. Using partitioned D-statistics which are less susceptible to false positives,
445 hybridisation was supported in in 52% (11/21) of tests performed on Tenerife, one of three tests on
446 La Palma, the single test on El Hierro (1/1) and one of three tests in Gran Canaria. Taken together
447 with evidence that hybridisation has generated two species by homoploid hybrid speciation
448 (Brochmann *et al.*, 2000; Fjellheim *et al.*, 2009; White *et al.*, 2018), it is clear that hybridisation has

449 played a significant role in the evolutionary history and diversification of *Argyranthemum*, a pattern
450 consistent with the findings in *Micromeria* (Curto *et al.*, 2017).

451 A recent review highlighted the role of hybridisation in generating new genetic combinations by
452 reassembly of old variation, thereby facilitating rapid speciation and adaptive radiations (Marques *et*
453 *al.*, 2019). This "combinatorial view" of speciation has been associated with several adaptive
454 radiations including Darwin's finches (Lamichhaney *et al.*, 2016, 2017; Han *et al.*, 2017), Lake Victoria
455 Cichlids (Meier *et al.*, 2017) and the Hawaiian silversword alliance (Baldwin & Sanderson, 1998;
456 Barrier *et al.*, 1999). The results presented here suggest that combinatorial processes may also be
457 significant in the diversification of Macaronesian lineages.

458 Our phylogenetic analysis confirmed that all three MIE species, *A. broussonetii*, *A. adauctum* and *A.*
459 *frutescens*, are non-monophyletic (Fig. 3). Although these findings are in agreement with the analysis
460 of Francisco-Ortega *et al.* (1996a,b), our analysis provided greater resolution of species and
461 subspecies relationships. *Argyranthemum frutescens* was resolved as paraphyletic with *A. gracile*
462 and "*A. vincentii"* nested within (Fig. 3; clade C). Indeed, based on our phylogenetic analysis, a
463 reconsideration of current taxonomic circumscriptions in the *A. frutescens* clade would be
464 appropriate.

465 *Argyranthemum broussonetii* is resolved as polyphyletic with subsp. *broussonetii* on Tenerife sister
466 to the multi-island endemic *A. frutescens*, whereas subsp. *gomerensis* on La Gomera is distantly
467 related and is sister to *A. callichrysum*, also from La Gomera. Based on D-statistics, there was no
468 evidence of hybridisation between *A. broussonetii* subsp. *broussonetii* and subsp. *gomerensis* or *A.*
469 *callichrysum* (Table 4; test 30).

470 *Argyranthemum adauctum* was also polyphyletic with three independent clades that corresponded
471 to (1) Gran Canaria (subsp. *gracile*, *canariense* and *jacobaeifolium*), (2) Tenerife (subsp. *adauctum*
472 and *dugourii*) and (3) La Palma and El Hierro (subsp. *palmensis* and *erythrocarpon,* respectively). Of
473 the three D-statistics performed between clades of *A. adauctum*, only one supported hybridisation
474 (between clades of *A. adauctum* on Tenerife and Gran Canaria; Table 4, test 31).

475 Curto *et al.* (2017) was able to infer evidence of hybridisation between lineages of *Micromeria*
476 distributed across different islands using D-statistics, suggesting that inter-island hybridisation might
477 be a significant process in the diversification of Macaronesian lineages, and Jones *et al.* (2014)
478 proposed that hybridisation may explains the polyphyly of MIEs in *Pericallis.* In *Argyranthemum*, this
479 was only supported in one of the four D-statistics implemented and in the absence of evidence for

hybridisation in the other MIE tests performed, morphological convergence of distinct evolutionary lineages may better explain the patterns observed.

In the case of *A. broussonetii*, the two subspecies are similar in leaf characteristics, but subsp. *gomerensis* shows greater affinity to *A. callichrysum* in capitula width and cypselae (dry single seeded fruits) traits (Supporting Information Fig. S 23; White *et al.*, submitted). This suggests that the two have converged on similar leaf traits, potentially in response to the similar habitats in which they occur.

The morphological characters that differentiate the three lineages of *A. adauctum* from other taxa (Humphries, 1976) appear to be more consistent, including hispid or tomentose indumentum, sessile leaves with primary lobes or teeth at the leaf base and wingless and fused ray cypselae (pers. obs.). Although the morphological convergence of *A. adauctum* is less easily explained than in *A. broussonetii*, it remains the most parsimonious explanation for the relationships observed.

Humphries (1976; page 162) noted the potential for convergent evolution in *Argyranthemum* given the "limited repertoire of leaf shape in the genus". In addition, Lee *et al.* (2005) proposed convergent evolution as a potential explanation for the non-monophyletic relationships of the genus *Taeckholmia* in the *Sonchus* alliance (Asteraceae: Sonchinae), although they were unable to rule out hybridisation as an alternative explanation. Convergent morphological evolution in response to similar habitats on islands has also been reported in *Nesotes* beetles in the Canary Islands (Rees *et al.*, 2001) and *Anolis* Lizards of the Greater Antilles (Losos *et al.*, 1998). Nevertheless, convergent morphological evolution has received much less attention in the diversification of flowering plants across Macaronesia compared with geographic isolation, habitat shifts and hybridisation.

Conclusions

The use of GBS has significantly improved the resolution of phylogenetic relationships in *Argyranthemum* and revealed greater complexity in the processes responsible for its diversification with geographical isolation, habitat shifts, hybridisation and convergent morphological evolution all inferred. High throughput sequencing (HTS) is increasingly employed for investigations of oceanic island endemic lineages (Mort *et al.*, 2015; Paun *et al.*, 2016; Curto *et al.*, 2017) and further studies of Macaronesian lineages using HTS are likely to provide the power to discern the complexity of the processes acting to generate flowering plant diversity in the region.

Acknowledgements

The authors would like to acknowledge Cabildo de Tenerife (permit number 18297), Cabildo de La Gomera (3102), Cabildo La Palma (2015005976) and Gobierno de Canarias (260209, 2015/939) for providing permits for fieldwork collections in Canary Islands. In addition, we would like to thank Juan Antonio Devesa Alcaraz (University of Córdoba) and Manuel de la Estrella (Royal Botanic Gardens Kew) for assisting with permits for collections in Andalusia. Fieldwork collections were performed with the help of Rachael Graham, Arnoldo Santos-Guerra, Giancarlo Torre, Magui Olangua-Corral and Miguel Menezes de Sequeira. Members of the Chapman lab at the University of Southampton and the biogeography discussion group at the Royal Botanic Gardens Kew provided instructive comments on the earlier drafts of the manuscript. We would like to thank Marcelino del Arco for sharing shape files for the vegetation of the Canary Islands. We would also like to 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. Finally, we would like to thank the anonymous reviewers for their feedback on our initial submission. This work was funded by an NHM-University of Southampton PhD studentship to OW.

Author contributions

Oliver White (OW), Mark Chapman (MCh) and Mark Carine (MCa) designed the study; OW, Alfredo Reyes-Betancort (AR-B) and MCa planned and undertook fieldwork; OW and MCh led the data analysis; all authors contributed to the interpretation of results; OW wrote the first draft of the manuscript; all co-authors contributed to the preparation of the final manuscript.

Data availability statement

Raw demultiplexed GBS sequence data used in this study are available from NCBI GenBank (BioProject ID: PRJNA664547) and assembled datasets are available upon request.

References

Allan GJ, Francisco-Ortega J, Santos-Guerra A, Boerner E, Zimmer EA. **2004**. Molecular phylogenetic evidence for the geographic origin and classification of Canary Island *Lotus* (Fabaceae: Loteae). *Molecular Phylogenetics and Evolution* **32**: 123–138.

del Arco Aguilar MJ, Rodríguez-Delgado O. **2018**. *Vegetation of the Canary Islands* (MJA Werger, Ed.). Cham: Spinger.

Baird NA, Etter PD, Atwood TS, Currey MC, Shiver AL, Zachary A, Selker EU, Cresko WA, Johnson EA. **2008**. Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS Genetics* **3**: 1–7.

Baldwin BG, Crawford DJ, Francisco-Ortega J, Kim S-C, Stuessy TF. **1999**. Molecular phylogenetic insights on the origin and evolution of oceanic island plants. In: Soltis DE, Soltis PS, Doyle JJ, eds. Molecular systematics of plants II. DNA sequencing. Boston, MA: Kluwer Academic Press, 410–441.

Baldwin BG, Sanderson MJ. **1998**. Age and rate of diversification of the Hawaiian silversword alliance (Compositae). *Proceedings of the National Academy of Sciences* **95**: 9402–9406.

Barber JC, Finch CC, Francisco-Ortega J, Santos-Guerra A, Jansen RK. **2007**. Hybridization in Macaronesian *Sideritis* (Lamiaceae): Evidence from incongruence of multiple independent nuclear and chloroplast sequence datasets. *Taxon* **56**: 74–88.

Barrier M, Baldwin BG, Robichaux RH, Purugganan MD. **1999**. Interspecific hybrid ancestry of a plant adaptive radiation: allopolyploidy of the Hawaiian silversword alliance (Asteraceae) inferred from floral homeotic gene duplications. *Molecular biology and evolution* **16**: 1105–1113.

Borgen L. **1976**. Analysis of a hybrid swarm between *Argyranthemum adauctum* and *A. filifolium* in the Canary islands. *Norwegian Journal of Botany* **23**: 121–137.

Bramwell D. **2011**. Introduction: islands and plants. In: Bramwell D, Caujapé-Castells J, eds. The Biology of Island Floras. Cambridge: Cambridge University Press, 1–10.

Brochmann C. **1984**. Hybridization and distribution of *Argyranthemum coronopifolium* (Asteraceae, Anthemideae) in the Canary Islands. *Nordic Journal of Botany* **4**: 729–736.

Brochmann C. **1987**. Evaluation of some methods for hybrid analysis, exemplified by hybridisation in *Argyranthemum*. *Nordic Journal of Botany* **7**: 609–630.

Brochmann C, Borgen L, Stabbetorp OE. **2000**. Multiple diploid hybrid speciation of the Canary Island endemic *Argyranthemum sundingii* (Asteraceae). *Plant Systematics and Evolution* **220**: 77–92.

Caujapé-Castells J, Tye A, Crawford DJ, Santos-Guerra A, Sakai A, Beaver K, Lobin W, Vincent Florens FB, Moura M, Jardim R, *et al.* 2010. Conservation of oceanic island floras: Present and future global challenges. *Perspectives in Plant Ecology, Evolution and Systematics* **12**: 107–129.

Curto M, Schachtler C, Puppo P, Meimberg H. **2017**. Using a new RAD-sequencing approach to study the evolution of *Micromeria* in the Canary Islands. *Molecular Phylogenetics and Evolution* **119**: 160–169.

Dupin J, Matzke NJ, Särkinen T, Knapp S, Olmstead RG, Bohs L, Smith SD. **2017**. Bayesian estimation of the global biogeographical history of the Solanaceae. *Journal of Biogeography* **44**: 887–899.

Eaton D, Ree RH. **2013**. Inferring phylogeny and introgression using RADseq data: an example from flowering plants (*Pedicularis*: Orobanchaceae). *Systematic Biology* **62**: 689–706.

Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, Mitchell SE. **2011**. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE* **6**: 1–10.

ESRI. **2019**. ArcGIS Desktop: Release 10.7. *Redlands CA: Environmental Systems Research Institute.*

Fjellheim S, Jørgensen MH, Kjos M, Borgen L. **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, Crawford DJ, Santos-Guerra A, Carvalho JA. **1996a**. Isozyme differenctiation in the endemic genus *Argyranthemum* (Asteraceae: Anthemideae) in the Macaronesian Islands. *Plant Systematics and Evolution* **202**: 137–152.

Francisco-Ortega J, Crawford DJ, Santos-Guerra A, Jansen RK. **1997a**. Origin and Evolution of *Argyranthemum* (Asteraceae: Anthemideae) in Macaronesia. In: Givnish TJ, Sytsma KJ, eds. In Molecular Evolution and Adaptive Radiation. Cambridge University Press, 407–432.

Francisco-Ortega J, Crawford DJ, Santos-Guerra A, Sa-Fontinha S. **1995a**. Genetic divergence among Mediterranean and Macronesian genera of the subtribe Chrysantheminae (Asteraceae). *American Journal of Botany* **82**: 1321–1328.

Francisco-Ortega J, Fuertes-Aguilar J, Kim SC, Santos-Guerra A, Crawford DJ, Jansen RK. **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, Jansen RK, Crawford DJ, Santos-Guerra A. **1995b**. Chloroplast DNA evidence for intergeneric relationships of the macaronesian endemic genus *Argyranthemum* (Asteraceae).

Systematic Botany **20**: 413–422.

Francisco-Ortega J, Jansen RK, Santos-Guerra A. **1996b**. Chloroplast DNA evidence of colonization, adaptive radiation, and hybridization in the evolution of the Macaronesian flora. *Proceedings of the National Academy of Sciences of the United States of America* **93**: 4085–4090.

Francisco-Ortega J, Santos-Guerra A, Hines A, Robert K. **1997b**. Molecular evidence for a Mediterranean origin of the Macaronesian endemic genus *Argyranthemum* (Asteraceae). *American Journal of Botany* **84**: 1595–1613.

Francisco-Ortega J, Santos-Guerra A, Kim SC, Crawford DJ. **2000**. Plant genetic diversity in the Canary Islands: A conservation perspective. *American Journal of Botany* **87**: 909–919.

Francisco-Ortega J, Santos-Guerra A, Mesa-Coello R, Gonzalez-Feria E, Crawford DJ. **1996c**. Genetic resource conservation of the endemic genus *Argyranthemum* Sch. Bip. (Asteraceae: Anthemideae) in the Macaronesian Islands. *Genetic Resources and Crop Evolution* **43**: 33–39.

Goodson BE, Santos-guerra A, Jansen RK, Goodson E, Jansen K. **2006**. Molecular systematics of *Descurainia* (Brassicaceae) in the Canary Islands: Biogeographic and taxonomic implications. *Taxon* **55**: 671–682.

Gruenstaeudl M, Carstens BC, Santos-Guerra A, Jansen RK. **2017**. Statistical hybrid detection and the inference of ancestral distribution areas in *Tolpis* (Asteraceae). *Biological Journal of the Linnean Society* **121**: 133–149.

Gruenstaeudl M, Santos-Guerra A, Jansen RK. **2013**. Phylogenetic analyses of *Tolpis* Adans. (Asteraceae) reveal patterns of adaptive radiation, multiple colonization and interspecific hybridization. *Cladistics* **29**: 416–434.

Han F, Lamichhaney S, Grant BR, Grant RP, Andersson L, Webster MT. **2017**. Gene flow, ancient polymorphism, and ecological adaptation shape the genomic landscape of divergence among Darwin's finches. *Genome Research* **27**: 1004–1015.

Humphries CJ. **1973**. A Taxonomic study of the genus Argyranthemum Webb Ex Sch. BIP. PhD thesis, The University of Reading, Berkshire, UK.

Humphries CJ. **1976**. A Revision of the Macaronesian Genus *Argyranthemum* Webb Ex Schultz BIP. (Compositae-Anthemideae). *Bulletin of the British Museum Natural History* **5**: 145–240.

Jones KE, Reyes-Betancort JA, Hiscock SJ, Carine MA. **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.

Kim SC, McGowen MR, Lubinsky P, Barber JC, Mort ME, Santos-Guerra A. **2008**. Timing and tempo of early and successive adaptive radiations in Macaronesia. *PLoS ONE* **3**: 1–7.

Kozlov AM, Darriba D, Flouri T, Morel B, Stamatakis A, Wren J. **2019**. RAxML-NG: A fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics* **35**: 4453–4455.

Lamichhaney S, Han F, Berglund J, Wang C, Almén MS, Webster MT, Grant BR, Grant PR, Andersson L. **2016**. A beak size locus in Darwin's finches facilitated character displacement during a drought. *Science* **352**: 470–474.

Lamichhaney S, Han F, Webster MT, Andersson L, Grant BR, Grant PR. **2017**. Rapid hybrid speciation in Darwin's finches. *Sience* **4593**: 224–228.

Landis MJ, Matzke NJ, Moore BR, Huelsenbeck JP. **2013**. Bayesian analysis of biogeography when the number of areas is large. *Systematic Biology* **62**: 789–804.

Lee C, Kim SC, Lundy K, Santos-Guerra A. **2005**. Chloroplast DNA phylogeny of the woody *Sonchus* alliance (Asteraceae: Sonchinae) in the Macaronesian Islands. *American Journal of Botany* **92**: 2072–2085.

Litsios G, Pearman PB, Lanterbecq D, Tolou N, Salamin N. **2014**. The radiation of the clownfishes has two geographical replicates. *Journal of Biogeography* **41**: 2140–2149.

Losos JB, Jackman TR, Larson A, Queiroz K de, Rodríguez-Schettino L. **1998**. Contingency and determinismin replicated adaptative radiations of island lizards. *Science* **279**: 2115–2118.

Marques DA, Meier JI, Seehausen O. **2019**. A combinatorial view on speciation and adaptive radiation. *Trends in Ecology & Evolution* **34**: 531-544.

Matzke NJ. **2014**. Model selection in historical biogeography reveals that founder-event speciation is a crucial process in island clades. *Systematic Biology* **63**: 951–970.

Matzke NJ. **2020**. Stochastic mapping under biogeographical models. *http://phylo.wikidot.com/biogeobears#stochastic_mapping* [accessed 01 April 2020].

McVay JD, Hipp AL, Manos PS. **2017**. A genetic legacy of introgression confounds phylogeny and biogeography in oaks. *Proceedings of the Royal Society of London B* **284**: 20170300.

Meier JI, Marques DA, Mwaiko S, Wagner CE, Excoffier L, Seehausen O. **2017**. Ancient hybridization fuels rapid cichlid fish adaptive radiations. *Nature Communications* **8**: 14363.

Mort ME, Crawford DJ, Kelly JK, Santos-Guerra A, Menezes de Sequeira M, Moura M, Caujape-Castells J. **2015**. Multiplexed-shotgun-genotyping data resolve phylogeny within a very recently derived insular lineage. *American Journal of Botany* **102**: 634–641.

Mort M, Soltis DE, Soltis PS, Francisco-Ortega J. **2002**. Phylogenetics and evolution of the Macaronesian clade of Crassulaceae inferred from nuclear and chloroplast sequence data. *Systematic Botany* **27**: 271–288.

Oberprieler C, Himmelreich S, Vogt R. **2007**. A new subtribal classification of the tribe Anthemideae. *Willdenowia* **37**: 89–114.

Paun O, Turner B, Trucchi E, Munzinger J, Chase MW, Samuel R. **2016**. Processes driving the adaptive radiation of a tropical tree (*Diospyros*, Ebenaceae) in New Caledonia, a biodiversity hotspot. *Systematic Biology* **65**: 212–227.

Pérez-Escobar OA, Chomicki G, Condamine FL, Karremans AP, Bogarín D, Matzke NJ, Silvestro D, Antonelli A. **2017**. Recent origin and rapid speciation of Neotropical orchids in the world's richest plant biodiversity hotspot. *New Phytologist* **215**: 891–905.

Puppo P, Curto M, Gusmão-Guedes J, Cochofel J, Pérez de Paz PL, Bräuchler C, Meimberg H. **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.

R Core team. **2020**. R (v3.6.3): A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA. **2018**. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology* **67**: 901–904.

Ree RH, Sanmartín I. **2018**. Conceptual and statistical problems with the DEC+J model of founder-event speciation and its comparison with DEC via model selection. *Journal of Biogeography* **45**: 741–

749.

Ree RH, Smith SA. **2008**. Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. *Systematic Biology* **57**: 4–14.

Rees DJ, Emerson BC, Oromí P, Hewitt GM. **2001**. The diversification of the genus *Nesotes* (Coleoptera: Tenebrionidae) in the Canary Islands: Evidence from mtDNA. *Molecular Phylogenetics and Evolution* **21**: 321–326.

dos Reis M, Donoghue PCJ, Yang Z. **2016**. Bayesian molecular clock dating of species divergences in the genomics era. *Nature Reviews Genetics* **17**: 71–80.

Ronquist F. **1997**. Dispersal-Vicariance Analysis: A New Approach to the Quantification of Historical Biogeography. *Systematic Biology* **46**: 195–203.

Ronquist F, Teslenko M, Van Der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. **2012**. Mrbayes 3.2: Efficient bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology* **61**: 539–542.

Schaefer H, Moura M, Belo Maciel MG, Silva L, Rumsey FJ, Carine MA. **2011**. The Linnean shortfall in oceanic island biogeography: A case study in the Azores. *Journal of Biogeography* **38**: 1345–1355.

Thode VA, Sanmartín I, Lohmann LG. **2019**. Contrasting patterns of diversification between Amazonian and Atlantic forest clades of Neotropical lianas (*Amphilophium*, Bignonieae) inferred from plastid genomic data. *Molecular Phylogenetics and Evolution* **133**: 92–106.

Trusty JL, Olmstead RG, Santos-Guerra A, Sá-Fontinha S, Francisco-Ortega J. **2005**. Molecular phylogenetics of the Macaronesian-endemic genus *Bystropogon* (Lamiaceae): Palaeo-islands, ecological shifts and interisland colonizations. *Molecular Ecology* **14**: 1177–1189.

Vitales D, Garnatje T, Pellicer J, Vallès J, Santos-Guerra A, Sanmartín I. **2014**. The explosive radiation of *Cheirolophus* (Asteraceae, Cardueae) in Macaronesia. *BMC Evolutionary Biology* **14**: 118.

White OW, Doo B, Carine MA, Chapman MA. **2016**. Transcriptome sequencing and simple sequence repeat marker development for three macaronesian endemic plant species. *Applications in Plant Sciences* **4**: 1600050.

White OW, Reyes-Betancort JA, Chapman MA, Carine MA. **2018**. Independent homoploid hybrid speciation events in the Macaronesian endemic genus *Argyranthemum*. *Molecular Ecology* **27**:

4856–4874.

Yang Z. **2007**. PAML 4: Phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution* **24**: 1586–1591.

Yu G, Smith DK, Zhu H, Guan Y, Lam TTY. **2017**. ggtree: An R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods in Ecology and Evolution* **8**: 28–36.

Supporting Information Figures

Fig. S 1 Principal Component Analysis (PCA) for each dataset.

Fig. S 2 Delta K and STRUCTURE plots for K one to 10 for datasets assembled.

Fig. S 3 Trace files for MrBayes analysis for each dataset.

Fig. S 4 Number of occurrence records for each vegetation type for all taxa in *Argyranthemum*.

Fig. S 5 Number of habitat types scored against the threshold used for filtering potentially erroneous data for all taxa in *Argyranthemum*.

Fig. S 6 Clades within islands used for D-statistics.

Fig. S 7 Number of reads for each sample following the filtering step of the ipyrad pipeline.

Fig. S 8 Maximum likelihood tree generated using RAxML-NG for the dataset based on a clustering threshold of 80 % and minimum sample number of 30.

Fig. S 9 Maximum likelihood tree generated using RAxML-NG for the dataset based on a clustering threshold of 80 % and minimum sample number of 38.

Fig. S 10 Maximum likelihood tree generated using RAxML-NG for the dataset based on a clustering threshold of 85 % and minimum sample number of 30.

Fig. S 11 Maximum likelihood tree generated using RAxML-NG for the dataset based on a clustering threshold of 85 % and minimum sample number of 38.

Fig. S 12 Maximum likelihood tree generated using RAxML-NG for the dataset based on a clustering threshold of 90 % and minimum sample number of 30.

Fig. S 13 Maximum likelihood tree generated using RAxML-NG for the dataset based on a clustering threshold of 90 % and minimum sample number of 38.

Fig. S 14 Bayesian inference tree generated using MrBayes for the dataset based on a clustering threshold of 80 % and minimum sample number of 30.

Fig. S 15 Bayesian inference tree generated using MrBayes for the dataset based on a clustering threshold of 80 % and minimum sample number of 38.

Fig. S 16 Bayesian inference tree generated using MrBayes for the dataset based on a clustering threshold of 85 % and minimum sample number of 30.

Fig. S 17 Bayesian inference tree generated using MrBayes for the dataset based on a clustering threshold of 85 % and minimum sample number of 38.

Fig. S 18 Bayesian inference tree generated using MrBayes for the dataset based on a clustering threshold of 90 % and minimum sample number of 30.

Fig. S 19 Bayesian inference tree generated using MrBayes for the dataset based on a clustering threshold of 90 % and minimum sample number of 38.

Fig. S 20 Time calibrated species tree of *Argyranthemum*.

Fig. S 21 Convergence of node date times across two independent MCMCTree analyses.

Fig. S 22 Ancestral range estimation for geography (left) and habitat (right) performed in BioGeoBEARS using DIVALIKE and BAYAREALIKE models respectively.

Fig. S 23 Morphology of *A. callichrysum*, *A. broussonetii* subsp. *gomerensis* and *A. broussonetii* subsp. *broussonetii*.

Supporting Information Tables

Table S 1 Summary of ipyrad parameters used for each assembly.

Table S 2 Cluster classification for each dataset identified by mclust.

Table S 3 AICc ranked substitution models for each dataset.

Table S 4 Summary statistics of MrBayes chain convergence.

Table S 5 Potential vegetation codes identified across georeferenced localities.

Table S 6 Best supported biogeographic model with varying thresholds for filtering potentially erroneous data.

Table S 7 Biogeographic areas identified for all taxa in *Argyranthemum*.

Table S 8 Summary of the 183 partitioned D-statistics performed.

Table S 9 Summary statistics for the samples used in our GBS assemblies generated using ipyrad.

Table S 10 Summary statistics for each ipyrad assembly.

Table S 11 Proportion of missing data across samples used in the phylogenetic analysis.

Supporting Information Methods

Methods S 1 Genetic clustering using PCA and STRUCTURE.

Methods S 2 Jupyter notebook of D-statistics performed.

Tables

Table 1 Leaf sampling for genotyping-by-sequencing (GBS) with taxa, collection reference, location, leaf and representative voucher specimen barcodes. Barcodes are for leaf and voucher specimens deposited at the Natural History Museum London (BM).

Taxa	Collection	Location	Leaf	Voucher
G. coronaria	White *et al.* 797	Andalucía, province of Cádiz	BM001092800	BM013407815
G. segetum	White *et al.* 796	Andalucía, province of Cádiz	BM001092799	BM013407814
A. adauctum subsp. *adauctum*	White *et al.* 120	Between La Laguna and Las Cañadas del Teide	BM010765622	BM000828632
A. adauctum subsp. *adauctum*	White *et al.* 135	Valle de la Orotova, TF-21	BM010765636	BM000828618
A. adauctum subsp. *canariense*	White *et al.* 363	GC-600 South of Presa de la Siberia	BM010765864	BM000828531
A. adauctum subsp. *canariense*	White *et al.* 366	South East of Presa de la Siberia	BM010765867	BM000828528
A. adauctum subsp. *dugourii*	White *et al.* 163	Off-road track North East of Vilaflor	BM010765664	BM000828592
A. adauctum subsp. *dugourii*	White *et al.* 166	Off-road track North East of Vilaflor	BM010765667	BM000828589
A. adauctum subsp. *erythrocarpon*	White.*et al.* 43	HI-1 between El Brezal and El Salvador	BM010765545	BM000828709
A. adauctum subsp. *erythrocarpon*	White *et al.* 44	HI-1 between El Brezal and El Salvador	BM010765546	BM000828708
A. adauctum subsp. *gracile*	White *et al.* 356	GC-60 between La Plata and Agualatente	BM010765857	BM000828534
A. adauctum subsp. *gracile*	White.*et al.* 360	GC-60 South of Risco las Candelillas	BM010765861	BM000828533*
A. adauctum subsp. *jacobaeifolium*	White *et al.* 368	South of Valsendero	BM010765869	BM000828526
A. adauctum subsp. *jacobaeifolium*[1]	White *et al.* 370	Near to La Laguna	BM010765871	BM000828524
A. adauctum subsp. *palmensis*	White *et al.* 58	Walk from Los Tilos to Marcos y Corderos	BM010765560	BM000828694
A. adauctum subsp. *palmensis*	White *et al.* 62	Walk from Los Tilos to Marcos y Corderos	BM010765564	BM000828690
A. broussonetii subsp. *broussonetii*	White *et al.* 157	Roques del Fraile	BM010765658	BM000828598
A. broussonetii subsp. *broussonetii*	White *et al.* 494	La Cumbrilla, Anaga	BM010765995	BM000828683*
A. broussonetii subsp. *broussonetii*	White *et al.* 552	Path to Mesa del Sabinal, Anaga	BM010766053	BM000828674*
A. broussonetii subsp. *broussonetii*	White *et al.* 664	Chamorga, Anaga	BM010766164	BM000828483*
A. broussonetii subsp. *broussonetii*	White *et al.* 674	Chamorga, Anaga	BM010766174	BM000828482*

A. broussonetii subsp. *broussonetii*	White *et al.* 719	Las Casas de la Cumbre, Anaga	BM010766218	BM000828476*
A. broussonetii subsp. *broussonetii*	White *et al.* 749	Barranco de Valle Crispín, Anaga	BM010766248	BM000828668*
A. broussonetii subsp. *broussonetii*	White *et al.* 768	Barranco de Valle Crispín, Anaga	BM010766267	BM000828667*
A. broussonetii subsp. *gomerensis*	White *et al.* 110	CV-5 between Las Rosas and La Palmita	BM010765612	BM000828642
A. broussonetii subsp. *gomerensis*	White *et al.* 112	South of La Palmita	BM010765614	BM000828640
A. callichrysum	White *et al.* 95	Valley below TF-713 in Barranco de la Guancha	BM010765597	BM000828657
A. callichrysum	White *et al.* 97	Roque de Agando	BM010765599	BM000828655
A. coronopifolium	White *et al.* 80	Chinamada, Anaga	BM010765582	BM000828672
A. coronopifolium	Graham *et al.* 107b	Anaga, Afur Roque Marrubial	BM001092356	BM000828856
A. dissectum	Graham *et al.* 13	Fajã dos Padres above cable car station	BM001092072	BM000828763
A. dissectum	Graham *et al.* 19	By tunnel entrance near Fajã da Ovelha	BM001092082	BM000828769
A. escarrei	White *et al.* 335	GC-200 between La Playa and Tirma	BM010765836	BM000828542
A. escarrei	White *et al.* 338	GC-200 near Degollada de la Aldea	BM010765839	BM000828542*
A. filifolium	White *et al.* 344	GC-200 North of Mogán	BM010765845	BM000828540
A. filifolium	White *et al.* 346	Barranco de Fataga	BM010765847	BM000828538
A. foeniculaceum	White *et al.* 142	TF-436 between Las Portelas and Masca	BM010765643	BM000828611
A. foeniculaceum	White *et al.* 144	TF-436 between Masca and Santiago del Teide	BM010765645	BM000828609
A. frutescens subsp. *canariae*	White *et al.* 319	North of La Atalaya	NA	BM000828553
A. frutescens subsp. *canariae*	White *et al.* 320	North of La Atalaya	NA	BM000828552
A. frutescens subsp. *foeniculaceum*	White *et al.* 107	TF-712 through Barranco del Valle	BM010765609	BM000828645
A. frutescens subsp. *foeniculaceum*	White *et al.* 116	Agulo	BM010765618	BM000828636
A. frutescens subsp. *frutescens*	White *et al.* 567	Maria Jiménez, Anaga	BM010766068	BM000828558*
A. frutescens subsp. *frutescens*	White *et al.* 585	Maria Jiménez, Anaga	BM010766086	BM000828557*
A. frutescens subsp. *frutescens*	White *et al.* 611	Barranco del Cercado de Andrés, Anaga	BM010766112	BM000828514*
A. frutescens subsp. *frutescens*	White *et al.* 620	Barranco del Cercado de Andrés, Anaga	BM010766120	BM000828513*
A. frutescens subsp. *gracilescens*	White *et al.* 177	TF-625 above Porís de Abona	BM010765678	BM000828578
A. frutescens subsp. *gracilescens*	White *et al.* 179	Road near to Arafo	BM010765680	BM000828576
A. frutescens subsp. *parviflorum*	White *et al.* 101	Calle la Lajita North of Aeropuerto de GO	BM010765603	BM000828651

A. frutescens subsp. *parviflorum*	White *et al.* 92	Barranco del Revolcadero above San Sebastián	BM010765594	BM000828660
A. frutescens subsp. *pumilum*[1]	White *et al.* 326	Alongside GC-200, North of Laja del Risco	BM010765827	BM000828551
A. frutescens subsp. *pumilum*[1]	White *et al.* 329	Overlooking Laja del Risco	BM010765830	BM000828548
A. frutescens subsp. *succulentum*	White *et al.* 229	Between Almáciga and Roque Bermejo, Anaga	BM010765730	BM000828575*
A. frutescens subsp. *succulentum*	White *et al.* 234	Between Almáciga and Roque Bermejo, Anaga	BM010765735	BM000828574*
A. frutescens subsp. *succulentum*[1]	White *et al.* 242	Roque Bermejo	BM010765743	BM000828732*
A. frutescens subsp. *succulentum*	White *et al.* 244	Roque Bermejo	BM010765745	BM000828731*
A. gracile	White *et al.* 169	TF-38 above Chío and Guía de Isora	BM010765670	BM000828586
A. gracile	White *et al.* 172	TF-82 above Tijoco Bajo	BM010765673	BM000828583
A. haematomma	Graham *et al.* 15	Path between Prazeres and Paul do Mar	BM000828765	BM001092077
A. haouarytheum	White *et al.* 56	LP-2 approximately two km North of El Charco	BM010765558	BM000828696
A. haouarytheum	White *et al.* 57	Walk below Volcán de San Antonio	BM010765559	BM000828695
A. hierrense	White *et al.* 38	HI-50 East of Sabinosa	BM010765540	BM000828714
A. hierrense	White *et al.* 47	HI-15 approx. 1km North of Villa de Valverde	BM010765549	BM000828705
A. lidii	White *et al.* 321	Amagro	BM010765822	BM000828547*
A. lidii	White *et al.* 325	Amagro	BM010765826	BM000828546*
A. maderense	White *et al.* 775	Haría, above Barranco de Temisa	BM010766274	BM000828473
A. maderense	White *et al.* 777	Haría, Vueltas de Malpaso	BM010766276	BM000828471
A. pinnatifidum subsp. *montanum*	Graham *et al.* 25	On path from Pico do Arieiro to Pico Ruivo	BM001092094	BM000828775
A. pinnatifidum subsp. *pinnatifidum*	Graham *et al.* 38	West of Encumeada Just after third road tunnel	BM001092116	BM000828788
A. pinnatifidum subsp. *succulentum*	Graham *et al.* 3	Ponta de São Lourenço	BM001092048	BM000828753
A. sventenii	Graham *et al.* 119a	By main road to Restinga	BM001092381	BM000828870
A. sventenii	Graham *et al.* 119b	By main road to Restinga	BM001092382	BM000828870
A. tenerifae[1]	White *et al.* 131	TF-24 below to Observatorio del Teide	BM010765632	BM000828622
A. tenerifae	White *et al.* 159	TF-24, Cañadas del Teide	BM010765660	BM000828596
A. tenerifae	White *et al.* 160	TF-24, Cañadas del Teide	BM010765661	BM000828595
A. tenerifae[1]	White *et al.* 564	Cañadas - walk to Refuge	BM010766065	BM000828596*
A. thalassophilum	Filipe Silva	Selvagem Pequena	NA	UMad s/n

A. vincentii	White *et al.* 123	Barranco de la Gota near to TF-523, above Arafo	BM010765625	BM000828629
A. vincentii	White *et al.* 125	Barranco de la Gota near to TF-523, above Arafo	BM010765627	BM000828627
A. webbii	White *et al.* 49	road near Lomo los Machines	BM010765551	BM000828703
A. webbii	White *et al.* 50	LP-1 between Llano Negro and Roque del Faro	BM010765552	BM000828702
A. winteri[1]	White *et al.* 794	Pájara, Jandía, Pico de La Zarza	BM001092793	NA
A. winteri	White *et al.* 795	Pájara, Jandía, Pico de La Zarza	BM001092794	NA

[1] Samples were removed for having less than 500,000 filtered GBS reads

* If a voucher for this leaf sample is not available a representative voucher from the same population is given

White *et al.* refers to collection made by O. White, M. Carine, A. Reyes-Betancort A. Santos-Guerra, G. Torre and M. Olangua-Corral

Graham *et al.* refers to collections made by R. Graham, M. Carine, M. Menezes de Sequeira

Table 2 Comparison of ancestral range estimation models DEC, DIVALIKE and BAYAREALIKE as implemented in BioGeoBEARS for geography and habitats. The best scoring models based on log-likelihood (Lnl) and corrected Akaike information criterion weight (AICc wt) are highlighted in bold. Other parameters include rate of dispersal (d) and rate of extinction (e).

		LnL	numparams	d	e	AICc	AICc_wt
Geography	DEC	-71.52	2	0.021	0.034	147.4	1.20×10^{-10}
	DIVALIKE	**-64.34**	**2**	**0.022**	**1.00×10^{-12}**	**133.0**	**1.50×10^{-07}**
	BAYAREALIKE	-90.39	2	0.022	0.440	185.1	7.50×10^{-19}
Habitats	DEC	-128.50	2	0.170	0.043	261.3	6.20×10^{-8}
	DIVALIKE	-129.00	2	0.180	1.00×10^{-12}	262.2	3.80×10^{-8}
	BAYAREALIKE	**-112.10**	**2**	**0.025**	**0.430**	**228.6**	**0.76**

Table 3 Summary of biogeographic events identified using BioGeoBEARS across 50 biogeographic stochastic maps for geography (DIVALIKE) and habitats (BAYAREALIKE). Mean counts for each event type are shown along with standard deviations and percentage of overall events. In the context of the oceanic islands, we interpret vicariance as inter-island allopatric speciation for geography and habitat shifts for habitat data.

		Geography			Habitat		
		Mean	sd	pct.	Mean	sd	pct.
Anagenetic	Range expansion	10.06	0.24	20.51%	7.76	2.41	7.30%
	Range contraction	0.00	0.00	0.00%	59.50	4.86	55.99%
Cladogenetic	Intra-range; narrow	27.78	0.51	56.62%	39.00	0.00	36.70%
	Intra-range; subset	0.00	0.00	0.00%	0.00	0.00	0.00%
	Vicariance	11.22	0.51	22.87%	0.00	0.00	0.00%
Total		**49.06**		**100.00%**	**106.26**		**100.00%**

Table 4 Summary of D-statistics performed between clades from the same island (tests 1-29) and between clades of non-monophyletic multi-island endemic taxa (tests 30-33). For each test performed, the taxa at positions P1, P2, and P3 are shown, together with the D-statistic, mean bootstrap value, bootstrap standard deviation, Z score, ABBA and BABA frequencies and number of loci used in the test. D-statistics significant at the 0.01 level are indicated with an asterisk. D-statistics that were also significant based on partitioned D-statistics are highlighted in bold.

n	P1	P2	P3	D	bootmean	bootstd	Z	ABBA	BABA	nloci
1	***A. callichrysum,*** ***A. broussonetii* subsp. *gomerensis*** **(La Gomera)**	***A. sventenii,*** ***A. hierrense*** **(El Hierro)**	***A. adauctum* subsp. *erythrocarpon*** **(El Hierro)**	**0.33 ***	**0.33**	**0.05**	**6.92**	**237.34**	**120.35**	**3100**
2	*A. callichrysum,* *A. broussonetii* subsp. *gomerensis* (La Gomera)	*A. haouarytheum* (La Palma)	*A. adauctum* subsp. *palmensis* (La Palma)	0.09	0.09	0.05	1.73	165.34	137.25	2877
3	***A. callichrysum,*** ***A. broussonetii* subsp. *gomerensis*** **(La Gomera)**	***A. webbii*** **(La Palma)**	***A. adauctum* subsp. *palmensis*** **(La Palma)**	**0.20 ***	**0.20**	**0.05**	**4.01**	**169.81**	**114.18**	**3038**
4	*A. callichrysum,* *A. broussonetii* subsp. *gomerensis* (La Gomera)	*A. webbii* (La Palma)	*A. haouarytheum* (La Palma)	0.20 *	0.21	0.05	3.94	148.08	98.02	2720
5	*A. sventenii,* *A. hierrense* (El Hierro)	*A. callichrysum,* *A. broussonetii* subsp. *gomerensis* (La Gomera)	*A. frutescens* subsp. *parviforum,* *A. frutescens* subsp. *foeniculaceum* (La Gomera)	0.02	0.02	0.06	0.36	104.75	100.49	3195
6	***A. frutescens* subsp. *parviforum,*** ***A. frutescens* subsp. *foeniculaceum*** **(La Gomera)**	***A. frutescens* subsp. *succulentum*** **(Tenerife)**	***A. broussonetii* subsp. *broussonetii*** **(Tenerife)**	**0.29 ***	**0.29**	**0.06**	**4.57**	**114.54**	**62.85**	**2905**
7	*A. frutescens* subsp. *parviforum,* *A. frutescens* subsp. *foeniculaceum* (La Gomera)	*A. vincentii* (Tenerife)	*A. broussonetii* subsp. *broussonetii* (Tenerife)	0.05	0.04	0.07	0.66	60.87	55.49	2389
8	*A. frutescens* subsp. *parviforum,* *A. frutescens* subsp. *foeniculaceum* (La Gomera)	*A. gracile* (Tenerife)	*A. broussonetii* subsp. *broussonetii* (Tenerife)	0.07	0.07	0.06	1.18	105.62	92.53	3554
9	***A. frutescens* subsp. *parviforum,*** ***A. frutescens* subsp. *foeniculaceum*** **(La Gomera)**	***A. frutescens* subsp. *frutescens*** **(Tenerife)**	***A. broussonetii* subsp. *broussonetii*** **(Tenerife)**	**0.20 ***	**0.20**	**0.04**	**4.58**	**132.42**	**88.74**	**3700**

10	***A. callichrysum,* *A. broussonetii* subsp. *gomerensis* (La Gomera)**	***A. tenerifae,* *A. adauctum* subsp. *adauctum,* *A. adauctum* subsp. *dugorii* (Tenerife)**	***A. broussonetii* subsp. *broussonetii* (Tenerife)**	**0.21 ***	**0.21**	**0.04**	**4.65**	**190.82**	**125.04**	**3818**
11	***A. callichrysum,* *A. broussonetii* subsp. *gomerensis* (La Gomera)**	***A. foeniculaceum,* *A. coronopifolium* (Tenerife)**	***A. broussonetii* subsp. *broussonetii* (Tenerife)**	**0.23 ***	**0.23**	**0.05**	**4.81**	**187.39**	**116.85**	**3619**
12	*A. frutescens* subsp. *parviforum,* *A. frutescens* subsp. *foeniculaceum* (La Gomera)	*A. vincentii* (Tenerife)	*A. frutescens* subsp. *succulentum* (Tenerife)	-0.03	-0.04	0.06	0.55	89.58	96.05	2107
13	*A. frutescens* subsp. *parviforum,* *A. frutescens* subsp. *foeniculaceum* (La Gomera)	*A. gracile* (Tenerife)	*A. frutescens* subsp. *succulentum* (Tenerife)	-0.04	-0.04	0.05	0.79	133.53	143.57	2920
14	*A. frutescens* subsp. *parviforum,* *A. frutescens* subsp. *foeniculaceum* (La Gomera)	*A. frutescens* subsp. *frutescens* (Tenerife)	*A. frutescens* subsp. *succulentum* (Tenerife)	-0.05	-0.05	0.04	1.27	138.19	152.26	3010
15	***A. callichrysum,* *A. broussonetii* subsp. *gomerensis* (La Gomera)**	***A. tenerifae,* *A. adauctum* subsp. *adauctum,* *A. adauctum* subsp. *dugorii* (Tenerife)**	***A. frutescens* subsp. *succulentum* (Tenerife)**	**0.33 ***	**0.33**	**0.05**	**6.89**	**169.65**	**85.76**	**2956**
16	***A. callichrysum,* *A. broussonetii* subsp. *gomerensis* (La Gomera)**	***A. foeniculaceum,* *A. coronopifolium* (Tenerife)**	***A. frutescens* subsp. *succulentum* (Tenerife)**	**0.40 ***	**0.40**	**0.05**	**8.29**	**175.03**	**74.25**	**2805**
17	*A. frutescens* subsp. *parviforum,* *A. frutescens* subsp. *foeniculaceum* (La Gomera)	*A. gracile* (Tenerife)	*A. vincentii* (Tenerife)	0.06	0.05	0.05	1.04	104.32	93.39	2414
18	*A. frutescens* subsp. *parviforum,* *A. frutescens* subsp. *foeniculaceum* (La Gomera)	*A. frutescens* subsp. *frutescens* (Tenerife)	*A. vincentii* (Tenerife)	0.08	0.08	0.04	1.96	126.28	106.86	2493
19	***A. callichrysum,* *A. broussonetii* subsp. *gomerensis* (La Gomera)**	***A. tenerifae,* *A. adauctum* subsp. *adauctum,* *A. adauctum* subsp. *dugorii* (Tenerife)**	***A. vincentii* (Tenerife)**	**0.35 ***	**0.35**	**0.05**	**6.86**	**131.82**	**63.08**	**2406**

20	*A. callichrysum,* *A. broussonetii* subsp. *gomerensis* (La Gomera)	*A. foeniculaceum,* *A. coronopifolium* (Tenerife)	*A. vincentii* (Tenerife)	0.23	0.23	0.06	3.60	102.71	64.17	2300
21	*A. frutescens* subsp. *parviforum,* *A. frutescens* subsp. *foeniculaceum* (La Gomera)	*A. frutescens* subsp. *frutescens* (Tenerife)	*A. gracile* (Tenerife)	-0.04	-0.04	0.03	1.31	199.51	217.19	3698
22	***A. callichrysum,*** ***A. broussonetii* subsp. *gomerensis*** **(La Gomera)**	***A. tenerifae,*** ***A. adauctum* subsp. *adauctum,*** ***A. adauctum* subsp. *dugorii*** **(Tenerife)**	***A. gracile*** **(Tenerife)**	**0.33 ***	**0.32**	**0.04**	**7.90**	**229.48**	**116.68**	**3663**
23	***A. callichrysum,*** ***A. broussonetii* subsp. *gomerensis*** **(La Gomera)**	***A. foeniculaceum,*** ***A. coronopifolium*** **(Tenerife)**	***A. gracile*** **(Tenerife)**	**0.22 ***	**0.22**	**0.04**	**4.99**	**177.87**	**112.93**	**3462**
24	***A. callichrysum,*** ***A. broussonetii* subsp. *gomerensis*** **(La Gomera)**	***A. tenerifae,*** ***A. adauctum* subsp. *adauctum,*** ***A. adauctum* subsp. *dugorii*** **(Tenerife)**	***A. frutescens* subsp. *frutescens*** **(Tenerife)**	**0.31 ***	**0.31**	**0.04**	**8.70**	**229.20**	**120.12**	**3855**
25	***A. callichrysum,*** ***A. broussonetii* subsp. *gomerensis*** **(La Gomera)**	***A. foeniculaceum,*** ***A. coronopifolium*** **(Tenerife)**	***A. frutescens* subsp. *frutescens*** **(Tenerife)**	**0.24 ***	**0.24**	**0.04**	**5.58**	**182.17**	**112.17**	**3646**
26	***A. callichrysum,*** ***A. broussonetii* subsp. *gomerensis*** **(La Gomera)**	***A. foeniculaceum,*** ***A. coronopifolium*** **(Tenerife)**	***A. tenerifae,*** ***A. adauctum* subsp. *adauctum,*** ***A. adauctum* subsp. *dugorii*** **(Tenerife)**	**0.19 ***	**0.19**	**0.04**	**4.63**	**233.85**	**158.98**	**3754**
27	*A. callichrysum,* *A. broussonetii* subsp. *gomerensis* (La Gomera)	*A. adauctum* subsp. *canariense,* *A. adauctum* subsp. *gracile,* *A. adauctum* subsp. *jacobaefolium* (Gran Canaria)	*A. frutescens* subsp. *canariense* (Gran Canaria)	0.32 *	0.32	0.06	5.49	108.38	55.25	2425
28	*A. callichrysum,* *A. broussonetii* subsp. *gomerensis* (La Gomera)	*A. filifolium,* *A. escarrei,* *A. lidii* (Gran Canaria)	*A. frutescens* subsp. *canariense* (Gran Canaria)	0.28 *	0.28	0.06	4.56	114.55	64.73	2397

29	***A. tenerifae,*** ***A. adauctum* subsp. *adauctum,*** ***A. adauctum* subsp. *dugorii*** **(Tenerife)**	***A. adauctum* subsp. *canariense,*** ***A. adauctum* subsp. *gracile,*** ***A. adauctum* subsp. *jacobaefolium*** **(Gran Canaria)**	***A. filifolium,*** ***A. escarrei,*** ***A. lidii*** **(Gran Canaria)**	**0.39 ***	**0.39**	**0.03**	**14.64**	**289.28**	**126.54**	**3978**
30	*A. callichrysum* (La Gomera)	*A. broussonetii* subsp. *gomerensis* (La Gomera)	*A. broussonetii* subsp. *broussonetii* (Tenerife)	0.05	0.05	0.06	0.89	81.07	72.65	3347
31	***A. filifolium,*** ***A. escarrei,*** ***A. lidii*** **(Gran Canaria)**	***A. adauctum* subsp. *canariense,*** ***A. adauctum* subsp. *gracile,*** ***A. adauctum* subsp. *jacobaefolium*** **(Gran Canaria)**	***A. adauctum* subsp. *adauctum,*** ***A. adauctum* subsp. *dugorii*** **(Tenerife)**	**0.36 ***	**0.36**	**0.03**	**12.30**	**267.43**	**126.36**	**3934**
32	*A. callichrysum,* *A. broussonetii* subsp. *gomerensis* (La Gomera)	*A. adauctum* subsp. *erythrocarpon,* *A. adauctum* subsp. *palmensis* (El Hierro)	*A. adauctum* subsp. *adauctum,* *A. adauctum* subsp. *dugorii* (Tenerife)	0.14	0.14	0.04	3.20	192.39	145.92	3720
33	*A. callichrysum,* *A. broussonetii* subsp. *gomerensis* (La Gomera)	*A. adauctum* subsp. *erythrocarpon,* *A. adauctum* subsp. *palmensis* (El Hierro)	*A. adauctum* subsp. *canariense,* *A. adauctum* subsp. *gracile,* *A. adauctum* subsp. *jacobaefolium* (Gran Canaria)	0.08	0.08	0.04	1.81	168.37	144.59	3733

Figures

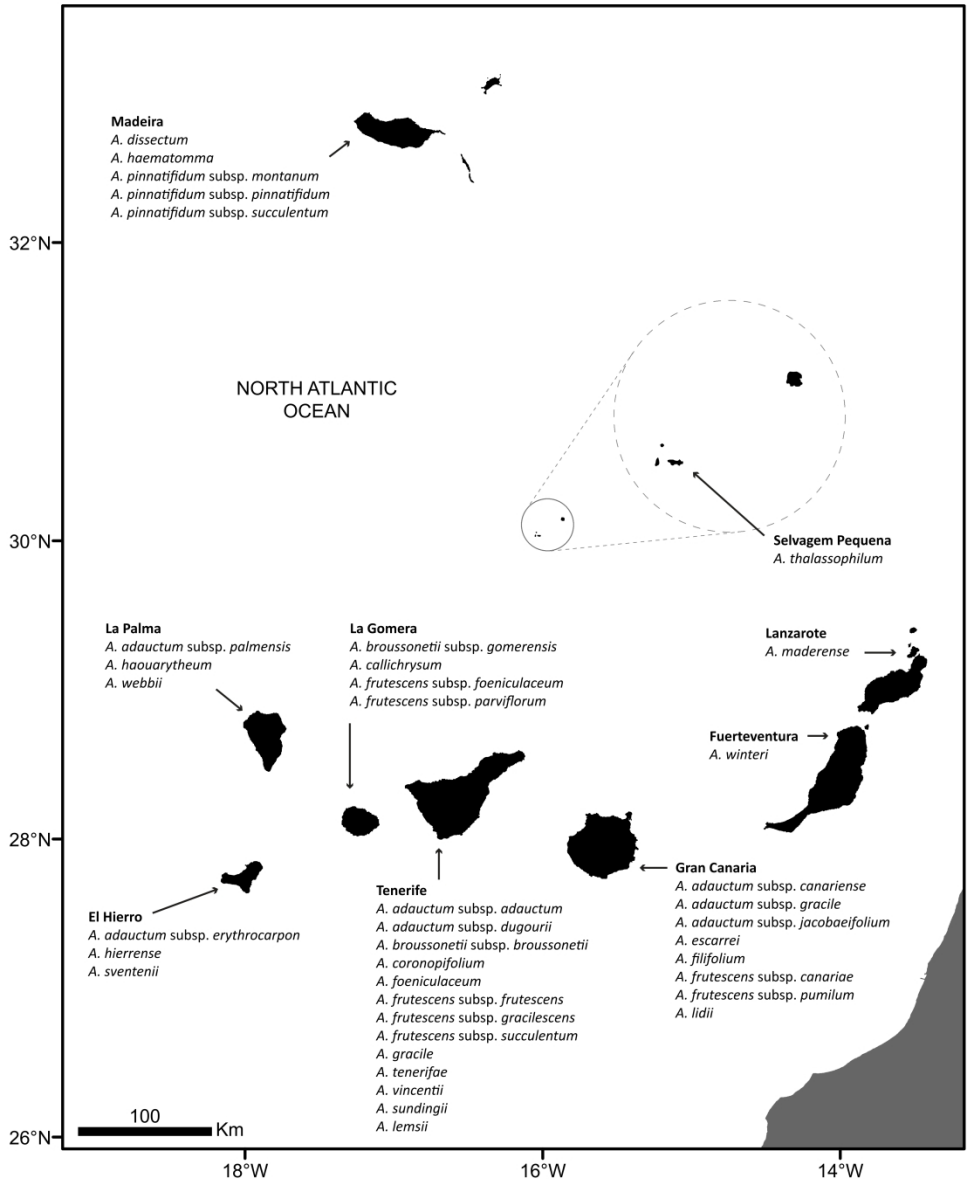
Fig. 1 Madeira, Selvagem Pequena and Canary Islands in the North Atlantic Ocean and the taxa in *Argyranthemum* occurring on each island that were sampled for this study.

Fig. 2 (a) Diagrams of the events identified in the biogeographical stochastic mapping analyses adapted from Matzke (2020). Anagenetic events include range expansion and contraction. Cladogenetic events include intra-range speciation (narrow and subset) and vicariance. In the context of the oceanic islands we interpret vicariance inter-island allopatric speciation for geography and habitat shifts for habitat data respectively. (b-e) Four taxon pectinate trees for D-statistics showing (b) ABBA and (c) BABA allele distributions where red arrows indicate hybridisation events between P3 and P2 or P1 and D-statistics testing for (d) hybridisation between lineages from the same island and (e) between multiple island endemic lineages.

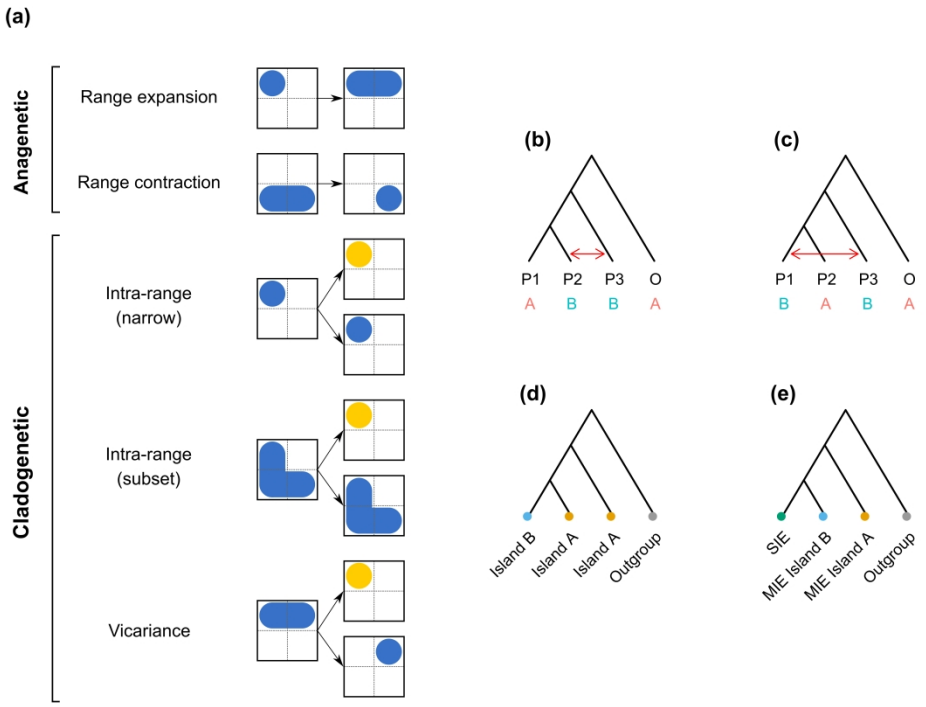
Fig. 3 Maximum likelihood tree of *Argyranthemum* generated using RAxML-NG for the dataset based on a clustering threshold of 90% and minimum sample number of 30. Branch lengths are shown except for outgroup taxa from *Glebionis* which were truncated to equal the longest branch length in *Argyranthemum*. Bootstrap values ≥ 70 are shown above the branches and posterior probabilities ≥0.95 from MrBayes are shown below the branches. Tips are coloured by island and clades A-G are discussed in the Results section (see phylogenetic reconstruction). A scale bar proportional to branch length is also shown.

Fig. 4 Ancestral range estimation in *Argyranthemum* for geography (left) and habitat (right) performed in BioGeoBEARS using DIVALIKE and BAYAREALIKE models respectively. The annotation on each node shows the most likely ancestral range for each model. Geographical states are abbreviated as: M (Madeira), S (Selvagem Pequena), L (Lanzarote), F (Fuerteventura), C (Gran Canaria), T (Tenerife), P (La Palma), G (La Gomera) and H (El Hierro). Habitat states are abbreviated as: E (*Euphorbia* scrubland), T (thermo-sclerophyllous woodland), L (laurel forest), P (pine forest) and S (subalpine zone). Clades A-G are the same as those in figure 3 and are discussed in the Results section (see phylogenetic reconstruction).

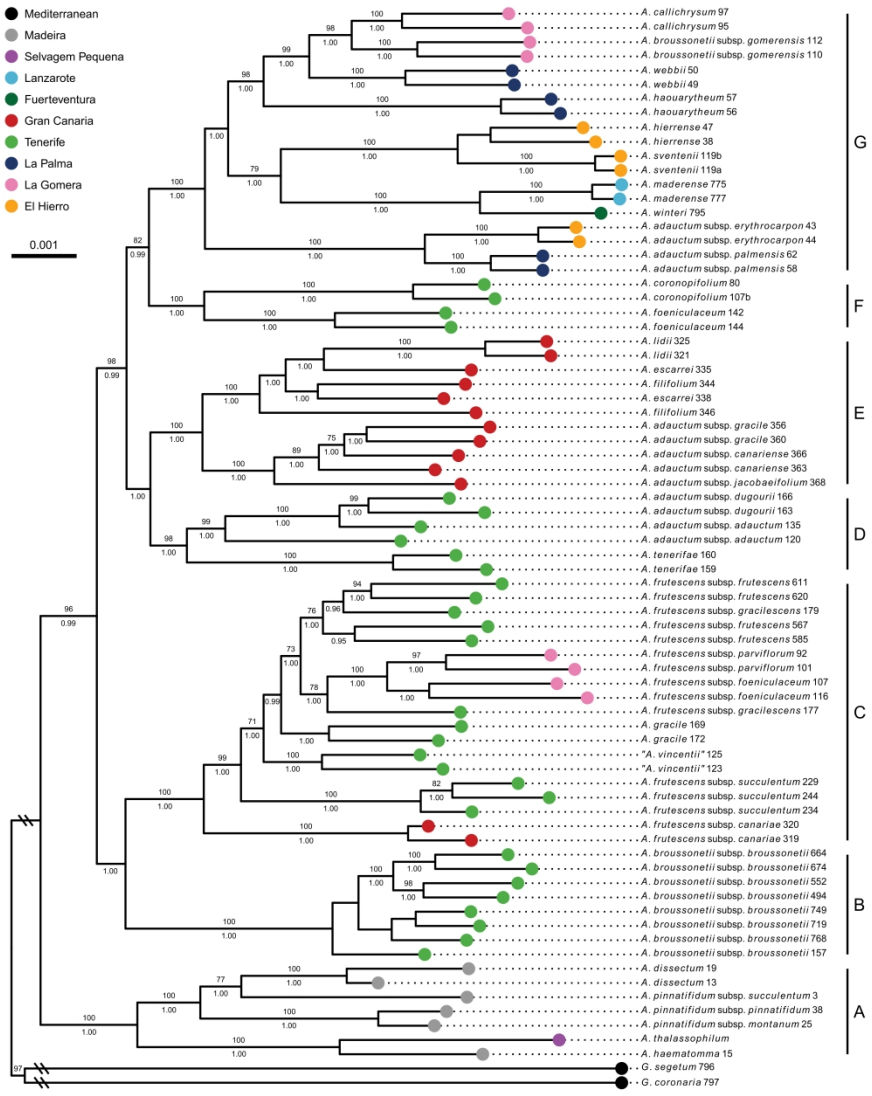
Fig. 5 Number of dispersal events estimated in *Argyranthemum* for (a) geography and (b) habitat. Counts of dispersal events were averaged across the 50 biogeographic stochastic models. The frequency of event is denoted by colour, with warmer colours indicating a higher frequency. The direction of dispersal is from the row state to the column state. The sum and percentages of events involving each area, either as a source (rows) or sink (columns) are given on the margins. Geographical states are abbreviated as: M (Madeira), S (Selvagem Pequena), L (Lanzarote), F (Fuerteventura), C (Gran Canaria), T (Tenerife), P (La Palma), G (La Gomera) and H (El Hierro). Habitat states are abbreviated as: E (*Euphorbia* scrubland), T (thermo-sclerophyllous woodland), L (laurel forest), P (pine forest) and S (subalpine zone).



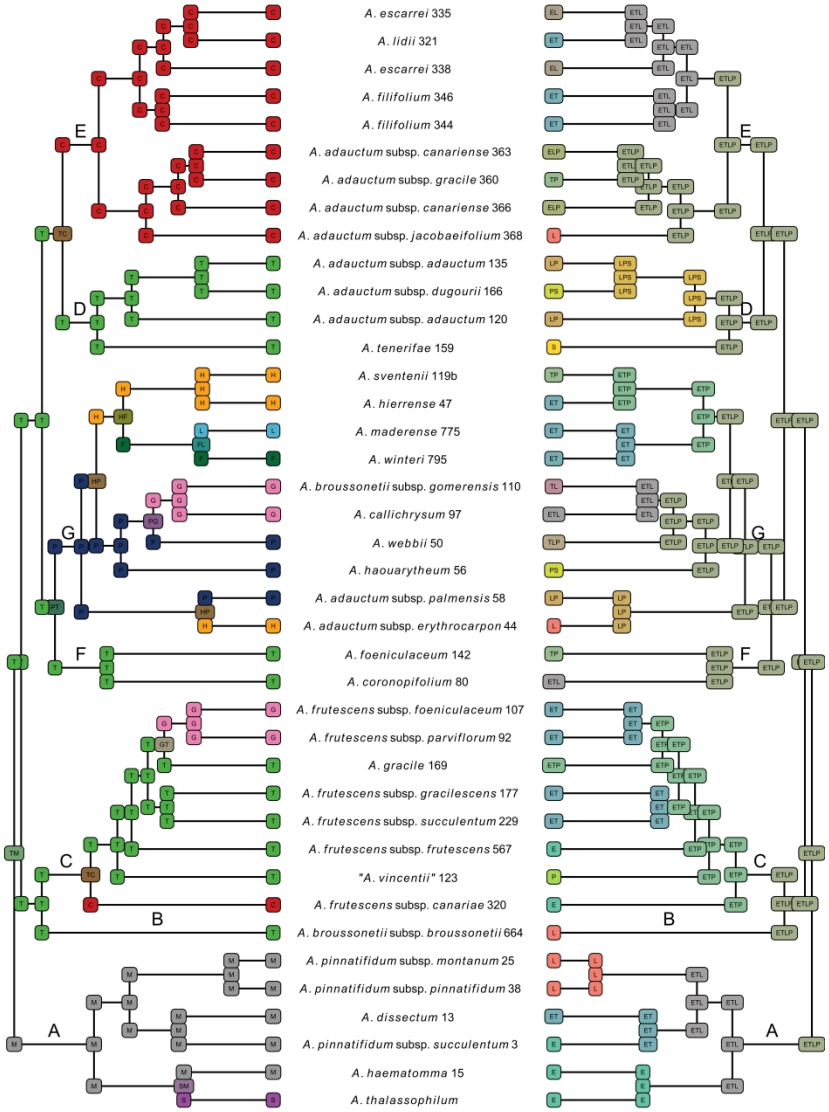
Madeira, Selvagem Pequena and Canary Islands in the North Atlantic Ocean and the taxa in *Argyranthemum* occurring on each island that were sampled for this study.



(a) Diagrams of the events identified in the biogeographical stochastic mapping analyses adapted from Matzke (2020). Anagenetic events include range expansion and contraction. Cladogenetic events include intra-range speciation (narrow and subset) and vicariance. In the context of the oceanic islands we interpret vicariance inter-island allopatric speciation for geography and habitat shifts for habitat data respectively. (b-e) Four taxon pectinate trees for D-statistics showing (b) ABBA and (c) BABA allele distributions where red arrows indicate hybridisation events between P3 and P2 or P1 and D-statistics testing for (d) hybridisation between lineages from the same island and (e) between multiple island endemic lineages.



Maximum likelihood tree of *Argyranthemum* generated using RAxML-NG for the dataset based on a clustering threshold of 90% and minimum sample number of 30. Branch lengths are shown except for outgroup taxa from *Glebionis* which were truncated to equal the longest branch length in *Argyranthemum*. Bootstrap values ≥ 70 are shown above the branches and posterior probabilities ≥0.95 from MrBayes are shown below the branches. Tips are coloured by island and clades A-G are discussed in the Results section (see phylogenetic reconstruction). A scale bar proportional to branch length is also shown.



Ancestral range estimation in *Argyranthemum* for geography (left) and habitat (right) performed in BioGeoBEARS using DIVALIKE and BAYAREALIKE models respectively. The annotation on each node shows the most likely ancestral range for each model. Geographical states are abbreviated as: M (Madeira), S (Selvagem Pequena), L (Lanzarote), F (Fuerteventura), C (Gran Canaria), T (Tenerife), P (La Palma), G (La Gomera) and H (El Hierro). Habitat states are abbreviated as: E (Euphorbia scrubland), T (thermo-sclerophyllous woodland), L (laurel forest), P (pine forest) and S (subalpine zone). Clades A-G are the same as those in figure 3 and are discussed in the Results section (see phylogenetic reconstruction).

(a)

	H	P	G	T	C	F	L	S	M	
H	-	0.70 (0.95)	0.02 (0.14)	0.00 (0.00)	0.00 (0.00)	0.28 (0.45)	0.24 (0.43)	0.00 (0.00)	0.00 (0.00)	1.24 12.33%
P	0.90 (0.79)	-	0.96 (0.20)	0.00 (0.00)	0.00 (0.00)	0.28 (0.45)	0.24 (0.43)	0.00 (0.00)	0.00 (0.00)	2.38 23.66%
G	0.00 (0.00)	0.06 (0.31)	-	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.06 0.60%
T	0.36 (0.48)	0.64 (0.48)	1.00 (0.00)	-	2.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	4.00 39.76%
C	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	-	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 0.00%
F	0.20 (0.40)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	-	0.52 (0.50)	0.00 (0.00)	0.00 (0.00)	0.72 7.16%
L	0.18 (0.39)	0.00 (0.00)	0.02 (0.14)	0.00 (0.00)	0.00 (0.00)	0.44 (0.50)	-	0.00 (0.00)	0.00 (0.00)	0.64 6.36%
S	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	-	0.18 (0.39)	0.18 1.79%
M	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.84 (0.37)	-	0.84 8.35%
	1.64 16.30%	1.40 13.92%	2.00 19.88%	0.00 0.00%	2.00 19.88%	1.00 9.94%	1.00 9.94%	0.84 8.35%	0.18 1.79%	10.06 100.00%

(b)

	E	T	L	P	S	
E	-	0.64 (0.72)	0.34 (0.59)	0.6 (0.83)	0.76 (0.72)	2.34 30.15%
T	0.24 (0.56)	-	0.24 (0.52)	0.78 (0.86)	0.70 (0.76)	1.96 25.26%
L	0.24 (0.52)	0.3 (0.51)	-	0.30 (0.51)	0.76 (0.77)	1.6 20.62%
P	0.14 (0.35)	0.22 (0.58)	0.32 (0.55)	-	1.04 (0.88)	1.72 22.16%
S	0.00 (0.00)	0.04 (0.20)	0.06 (0.24)	0.04 (0.20)	-	0.14 1.80%
	0.62 7.99%	1.2 15.46%	0.96 12.37%	1.72 22.16%	3.26 42.01%	7.76 100.00%

Number of dispersal events estimated in *Argyranthemum* for (a) geography and (b) habitat. Counts of dispersal events were averaged across the 50 biogeographic stochastic models. The frequency of event is denoted by colour, with warmer colours indicating a higher frequency. The direction of dispersal is from the row state to the column state. The sum and percentages of events involving each area, either as a source (rows) or sink (columns) are given on the margins. Geographical states are abbreviated as: M (Madeira), S (Selvagem Pequena), L (Lanzarote), F (Fuerteventura), C (Gran Canaria), T (Tenerife), P (La Palma), G (La Gomera) and H (El Hierro). Habitat states are abbreviated as: E (Euphorbia scrubland), T (thermo-sclerophyllous woodland), L (laurel forest), P (pine forest) and S (subalpine zone).