Inter-island differentiation and contrasting patterns of diversity in the iconic Canary Island sub-alpine endemic Echium wildpretii (Boraginaceae)

Journal:	*Systematics and Biodiversity*
Manuscript ID	TSAB-2020-0115.R1
Manuscript Type:	Original Research Article
Keywords:	Echium, Boraginaceae, Canary Islands, Island biogeography, Microsatellites, Population genetics, Taxonomy

Inter-island differentiation and contrasting patterns of diversity in the iconic Canary Island sub-alpine endemic *Echium wildpretii* (Boraginaceae)

RACHAEL E. GRAHAM[1,2], J. ALFREDO REYES-BETANCORT[3], MARK A. CHAPMAN[2], 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] Jardín de Aclimatación de La Orotava, Instituto Canario de Investigaciones Agrarias (ICIA), C/ Retama 2 38400 Puerto de la Cruz, Tenerife, Spain.

* m.carine@nhm.ac.uk

Abstract

The sub-alpine zones of oceanic islands are unique and dynamic ecosystems with high levels of endemism, making them particularly suitable model systems in which to investigate evolutionary and biogeographic processes. The sub-alpine flora of the Canary Islands is restricted to the islands of Tenerife and La Palma. Its origins are poorly understood. *Echium wildpretii* Hook.f. is an iconic species of the subalpine zones of these two islands, with distinct subspecies recognised on each island. This study examines patterns of genetic and morphological diversity in *E. wildpretii* to investigate the diversity and evolution of the lineage. Nine microsatellite markers were designed and used to investigate population genetic structure and patterns of gene flow within and between islands and populations. Morphological characters were assessed to test the distinctiveness of the two subspecies recognised. Strong genetic differentiation was observed between islands with higher genetic diversity on the younger island of La Palma than on Tenerife. Very low levels of inter-island gene flow were observed indicating that these taxa are reproductively isolated and evolving independently. Morphological analysis confirmed the distinctiveness of plants from the two islands. Given their genetic and morphological

distinctiveness the taxa on Tenerife and La Palma merit recognition as distinct species. Higher genetic diversity in the La Palma species is consistent with an origin of the lineage on this island via upslope colonisation, followed by dispersal to Tenerife where the plants show lower genetic diversity.

Key words: *Echium*, Boraginaceae, Canary Islands, island biogeography, microsatellites, population genetics, taxonomy

Introduction

The sub-alpine floras of oceanic islands offer excellent study systems in which to investigate evolutionary processes and ecological adaptation. The dynamic nature of the sub-alpine ecosystem on oceanic islands make them a hotspot for speciation and endemism where it is possible to study evolution in action (Steinbauer et al., 2012, 2016; Fernández-Palacios et al., 2014), but it also means they are disproportionately at risk from future climate change (Dirnböck et al., 2011).

The sub-alpine scrub zone in the Canary Islands, a volcanic oceanic archipelago of the north-east Atlantic, is confined to high elevation on the two highest islands of Tenerife and La Palma (Fig. 1). This zone is found above the treeline from 2000 m above sea level (a.s.l.) to the highest points on the islands (3718 m at Pico del Teide, Tenerife and 2425 m at Roque de los Muchachos, La Palma (Fernández-Palacios et al., 2014)). The area of sub-alpine vegetation on Tenerife and La Palma occupies 145 km^2 and 15 km^2 respectively (Fernández-Palacios et al., 2014). It experiences harsh climatic conditions and provides challenging conditions for plants. Average temperatures are generally low, with mean annual temperature in the sub-alpine zone ranging from 3.5 °C at Pico del Teide (Del-Arco et al., 2006) to 13.6 °C at 2345 m a.s.l. in Las Cañadas on Tenerife (Gieger and Leuschner, 2004), and with an annual mean of around 8 ˚C at the highest parts of La Palma (von Suchodoletz et al., 2013). Additionally, as the sub-alpine zones are above the trade wind inversion there is little cloud cover and consequently very little precipitation (Gieger and Leuschner, 2004). The lack of cloud cover also results in high levels of solar radiation during the day, which causes rapid heating of the air and leaf surfaces meaning that, as well as the cold, the flora of the sub-alpine zone has to acclimate to a large diurnal variation in temperature (Smith and Young, 1987).

The harsh abiotic conditions of the Canarian sub-alpine zone, small area of extent, and its geographic isolation from other alpine areas (the nearest being the High Atlas mountains of Morocco, over 700 km away) means that its flora is relatively species poor but highly specialised, with a large component of endemic species that display specific adaptations to high altitude (Steinbauer et al., 2012; Fernández-Palacios et al., 2014; Irl et al., 2015). Indeed, of the major vegetation zones in the Canary Islands, the sub-alpine zones exhibit the highest levels of endemism with 31.7% and 34.9% of sub-alpine plant species being Canary Island endemics on Tenerife and La Palma respectively (Steinbauer et al., 2012). ~~Phylogenetic evidence suggests that the origins of this sub-alpine endemic flora are complex, with taxa having colonised via multiple different routes.~~

Echium is one of the largest evolutionary radiations in Macaronesia with 29 endemic species (Bramwell, 1972; Santos-Guerra, 1983; Carvalho and Pontes, 2010). *Echium wildpretii* Hook.f. ~~is~~is a biennial species present in the sub-alpine zone of Tenerife and La Palma as subsp. *wildpretii* on Tenerife and subsp. *trichosiphon* (Svent.) Bramwell on La Palma. The subspecies have similar morphology but differ in characters including flower colour (Fig. 1). Bramwell (1972) also suggested that the two differ in the shape of the inflorescence (broadest at the middle in subsp. *trichosiphon* and tapering evenly to the base in subsp. *wildpretii*), with subsp. *trichosiphon* also exhibiting a more densely hispid calyx and corolla, and a more deeply bifid style than subsp. *wildpretii*. As a sub-alpine endemic with a narrow climatic niche and a very restricted distribution, *E. wildpretii* is likely to be negatively impacted by future climate change (Dirnböck et al., 2011). The Red List assessment of the Spanish flora assessed *E. wildpretii* subsp. *trichosiphon* as "Vulnerable" D2 under the IUCN criteria, but made no assessment of subsp. *wildpretii* (Moreno, 2008).

The aim of this study is to investigate the diversity and evolution of *E. wildpretii.* We utilise simple sequence repeat markers (SSRs, also known as microsatellites) to investigate patterns of genetic diversity and structure in *E. wildpretii*. SSR markers are highly polymorphic, have a simple Mendelian mode of inheritance, and are abundant in the genome (Selkoe and Toonen, 2006). They are also a cost-effective method for generating data for a large number of individuals if existing genomic resources such as a transcriptome are available (Ellis and Burke, 2007; Hodel et al., 2016). SSRs have proved informative for population level analysis of other Macaronesian plants, including *Argyranthemum* (Asteraceae; White et al., 2018), *Brachypodium* (Poaceae; Shiposha et al., 2016), *Bencomia* (Rosaceae; González-Pérez et al., 2009), *Sambucus* (Sambucaceae; Sosa et al., 2010), *Pinus* (Pinaceae; Navascués et al., 2006) and *Micromeria* (Lamiaceae; Puppo

et al., 2016). In this paper, we quantify levels of genetic diversity in *E. wildpretii* and assess how this diversity is partitioned between islands and populations. We also analyse population genetic structure and patterns of gene flow within and between islands to test the distinctiveness of the two subspecies currently recognised. Given our findings, we review morphological variation in *E. wildpretii*, and reconsider the evolution, taxonomic status and conservation of its two subspecies.

Materials and methods

Study site

Fieldwork to collect samples was carried out in the Canary Islands during May 2016. Plant material was collected under permits from the Cabildo de Tenerife (numbers 22835 and 24339), the Cabildo de La Palma (number 2016005709), and the Gobierno de Canarias (number 671303). For each of the two sub-species of *E. wildpretii* four populations were sampled, with the aim of representing as much of the geographic and altitudinal range of the taxa as possible (Fig. 1 and Appendix S1; see Supplemental Data). Voucher specimens for each population were deposited at the Natural History Museum, London (BM) and at Jardín de Aclimatación de la Orotava (ORT). Fresh leaf material was collected from multiple individuals of each population, spaced at least 2 metres apart, and was dried in silica gel (Chase and Hills, 1991). A total of 100 individuals were sampled for this study.

Primer design

SSR primer design was carried out by utilising an *E. wildpretii* transcriptome (White et al., 2016). The transcriptome sequence was searched for repeat regions using misa.pl (http://pgrc.ipk-gatersleben.de/misa/), with the minimum number of di-, tri- and tetra-repeats set to 6, 4 and 4 respectively. A total of 1266 regions were identified. After excluding sequences in which the SSR occurred within the first or last 50 bp of the transcript, or was in a compound formation, a total of 780 SSRs were retained. As longer SSRs are more likely to show length variation and therefore to be informative for analysis, the longest 50 SSRs were selected for consideration. The transcriptome assembly pipeline that was used (see White et al., 2016) utilised Trinity (Haas et al., 2013) which assembles reads into isoforms which are considered alternative forms of each gene; we therefore reduced our dataset to only contain one isoform per gene.

Sequences of the 50 transcripts were used in a BLAST search (Altschul et al., 1990) against the tomato (*Solanum lycopersicum* L.) genome, as this is the model plant organism most closely related to *Echium*. The BLAST search was optimised for dissimilar sequences (discontiguous megablast). One transcript was shown to potentially contain an intron which could have made the PCR product too large for analysis, therefore this locus was excluded. Primers were designed for the remaining 49 SSRs using Primer3 (Untergasser et al., 2012) with the product size parameter set to 120-400 bp. 24 primer pairs were selected across the range of product sizes to allow for efficient multiplexing of samples for downstream genotyping.

Primer screening

Primers were screened across a set of eight *Echium* DNA samples representing individuals from two populations of *E. wildpretii* subsp. *wildpretii*, three populations of *E. wildpretii* subsp. *trichosiphon*, two populations of *E. pininana* Webb & Berthel. (the likely sister species of *E. wildpretii*), and one individual of the more distantly related Madeiran species *E. candicans* L.f.. The choice of both close and distant relatives of *E. wildpretii* allowed us to assess the universality of the primers across the Macaronesian *Echium* clade.

DNA was extracted from silica-dried leaf samples using either a modified Sodium dodecyl sulfate (SDS) extraction protocol (Page and Minocha, 2005) or the BioSprint 96 DNA Plant Kit (Qiagen, Manchester, UK). PCR was carried out using a three-primer method with the universal fluorescently labelled primers (either TET or FAM) which anneal to the 5' end of the forward primer (Schuelke, 2000). Each PCR reaction contained 10 mM Tris-HCl (pH 8.8), 50 mM KCl, 0.01% Tween 20, 1.5 mM MgCl2, 0.2 mM dNTPs, 0.04 μM forward primer, 0.2 μM reverse primer, 0.2 μM fluorescent primer, 1 unit of Taq DNA polymerase, 1 ng DNA, and was made up to 15 μL with water. A touchdown PCR programme was used, consisting of an initial denaturation at 94 °C for 3 minutes, then 10 cycles of 94 °C for 30 seconds, 65 °C for 30 seconds (decreasing by 1 °C per cycle), and 72 °C for 1 minute, then 30 cycles of 94 °C for 30 seconds, 55 °C for 30 seconds, and 72 °C for 1 minute, and then a final elongation step of 72 °C for 7 minutes. PCR products were run on a 1 % agarose gel stained with GelRed (Biotium, Hayward, California, USA) and visually inspected to assess amplification success.

Of the 24 primer pairs tested, 14 were discarded as they either failed to amplify or produced multiple/indistinct bands on the gel. The remaining 10 primer pairs were retained for genotyping and further analysis as they amplified successfully and produced what

appeared to be a single PCR product. To prepare samples for genotyping, PCR products were diluted 1:30 with sterile distilled water and multiplexed in three sets so that all products could be differentiated by size and fluorescent label. Genotyping was performed on an ABI3730xl DNA analyser (Applied Biosystems, Carlsbad, California, USA) at the Department of Zoology, University of Oxford, United Kingdom. Alleles were scored from the raw traces using GeneMarker 2.6.7 (SoftGenetics, State College, Pennsylvania, USA). One locus produced indistinct peaks and showed very low levels of polymorphism so was excluded from further analysis. The remaining nine loci (Table 1) were amplified across the full set of *E. wildpretii* samples using the PCR conditions described previously and were multiplexed in two groups. Genotyping and scoring were carried out as above.

Genetic diversity and differentiation

Statistics for each of the populations were calculated using GenAlEx v6.503 (Peakall and Smouse, 2012). Samples with greater than 50% missing data were excluded. Genetic diversity was estimated for each of the eight *E. wildpretii* populations by calculating the number of alleles per locus (Na), number of private alleles (PA), expected and observed heterozygosity (H_e and H_o), the Shannon's Information Index (I) and the fixation index (F). Further, departure from Hardy-Weinberg Equilibrium (HWE) and estimates of null allele frequencies were calculated in Cervus v3.0.7 (Kalinowski et al., 2007). Tests for linkage disequilibrium amongst loci were implemented in Genepop v4.7.5(Raymond and Rousset, 1995; Rousset, 2008).

Genetic differentiation between populations was estimated by calculating pairwise F_{ST} and Nei distances between sampling localities. Analysis of molecular variance (AMOVA) was carried out using F_{ST} as the measure of variance to assess the distribution of genetic variance in *E. wildpretii*. This analysis was carried out with three different groupings to investigate the distribution of genetic variance: (1) within Tenerife, (2) within La Palma, and (3) between the two islands. A Principal Coordinates Analysis (PCoA) was performed in GenAlEx, and the first two axes plotted to visualise the genetic distance between individuals. Model-based clustering analysis was carried out on the PCoA results using the mclust package (Scrucca et al., 2016) implemented in RStudio v3.4.3 (RStudio Team, 2016) to identify the most likely number of clusters in the dataset. Tests for recent population bottlenecks were implemented BOTTLENECK v1.2.02 (Cornuet and Luikart, 1996), using two different microsatellite mutation models: infinite allele model (IAM) and stepwise mutation model (SMM).

Network analysis

An unrooted network of all individuals was constructed using the APE package (Paradis et al., 2004) implemented in RStudio v3.4.3. The function "aboot" was used to perform a neighbour-joining analysis based on pairwise Nei genetic distances, with 100 bootstrap replicates. The resulting tree file was exported into FigTree v1.4.3 (Rambaut, 2016) for editing.

Isolation by distance

To investigate the role of isolation-by-distance (IBD) in our dataset, Mantel tests were performed in GenAlEx to test for correlation between matrices of pairwise population F_{ST} and geographic distance between sampling localities. As this study sampled multiple populations separated by short geographic distances within islands, which are then separated by a much greater distance between islands, any subtle patterns of within-island local adaptation could be swamped by a strong geographic signal. Therefore, in addition to a Mantel test carried out with a full dataset of all eight *E. wildpretii* sampling localities, we also analysed datasets corresponding to the four Tenerife and four La Palma localities respectively.

Genetic structure

Genetic clustering among individuals was analysed using Structure v2.3.4 (Pritchard et al., 2000), which uses a Bayesian clustering algorithm to assign individuals to genetic clusters/populations across a range of numbers of populations (*K*). To determine the most likely number of clusters, values of *K* from one to eight were tested with 10 iterations of 50,000 MCMC generations and a burn-in of 20,000 generations for each value of *K*. The most likely value of *K* was determined from the rate of change in posterior probability between successive values of *K* (delta K) using the method of Evanno et al. (2005) using STRUCTURE HARVESTER v0.6.94 (Earl and vonHoldt, 2012). The ten iterations of each *K* were combined using the greedy algorithm in CLUMPP (Jakobsson and Rosenberg, 2007). The CLUMPAK online server (Kopelman et al., 2015) was used to implement CLUMPP and to visualise the results as plots displaying individuals as vertical bars, coloured to represent membership of different genetic clusters.

Gene flow

The effective number of migrants per generation (N_m) was estimated from pairwise F_{ST} values in GenAlEx. N_m values were calculated between sampling localities within each island, as well as between the two islands. It is generally assumed that if Nm ≥ 1 (i.e. one or more migrants per generation) the level of gene flow is sufficient to prevent populations from diverging due to genetic drift (Ellstrand and Elam, 1993).

Rates of recent migration (last 1-2 generations) were estimated using BayesAss v3.0.4 (Wilson et al., 2003), which uses a Bayesian Inference framework to estimate the rate of migration between populations over the last two generations. This can also be used to calculate the probability that an individual originated from its source population or is a first- or second-generation immigrant from another population. The MCMC was run for 100 million iterations, sampling every 100 iterations, and with the first ten million iterations discarded as burn-in. The mixing parameters for allele frequencies and inbreeding coefficients were set to 0.30 and 0.50 respectively, as these values are expected to produce a suitable rate of accepted changes (Wilson and Rannala, n.d.). To ensure that these mixing parameter values were appropriate for the data being analysed, the acceptance rates were checked to ensure that they fell within the optimal range (20-60%) which usually maximises the log likelihood. The trace file generated by the MCMC run was analysed in Tracer v1.6 (Rambaut and Drummond, 2013) to ensure convergence. A second run was performed using the same parameters but a different seed number for comparison, with both runs producing near identical migration rate estimates.

Morphology

Material cultivated from wild-collected seed, herbarium specimens at BM and ORT and field observations were used to assess the morphological distinctiveness of plants from Tenerife and La Palma. Through this work we tested whether the characters previously described as diagnostic for the two subspecies recognised (Bramwell, 1972) are robust, and to identify any additional characters which distinguish the taxa.

Variation in leaf morphology was assessed from herbarium specimens (listed in the Taxonomic Treatment) and from cultivated plants grown from seed collected in Tenerife and La Palma. Herbarium specimens were examined under a dissecting microscope to assess variation in pubescence of the calyx and corolla, and division of the style. For some specimens, floral dissections were made following softening of the flowers. Style

measurements were taken from three flowers per specimen, with a graticule used to measure the length of the stigma lobes.

Results

Primer amplification

Of the 24 primers screened in *E. wildpretii*, nine successfully amplified, produced a single PCR product matching the expected size, and were sufficiently polymorphic for genotyping and further analysis (Table 1). These nine primers also amplified and produced a single PCR product in both *E. pininana* and *E. candicans*, indicating that these markers could be applied to other Macaronesian *Echium* species. Of the 100 *E. wildpretii* individuals initially sampled, 97 were retained for analysis after excluding those with >50% missing data.

Genetic diversity and differentiation

Mean number of alleles per locus (N_a) varies from 1.222 to 3.778 per population, with an average of 2.889 across all populations (Table 2). The Tenerife populations of *E. wildpretii* have a significantly lower number of alleles per locus than the La Palma populations (Welch Two Sample t-test, t = -3.611, df = 3.353, p-value = 0.015). The range of N_a amongst the La Palma populations is relatively small (0.334), whereas there is a large range amongst the Tenerife populations (1.445), due to population T1 which has a very low number of alleles (1.222) and zero observed heterozygosity. The mean number of private alleles per locus is relatively low across all populations but shows a wider range on Tenerife (0-0.667) than La Palma (0.222-0.333). Relatively low estimated frequencies of null alleles were found in the La Palma populations (average < 20%; similar to average estimates across 233 studies reviewed by Dakin and Avise (2004) and no loci showed departure from HWE. However, estimated frequencies of null alleles were higher in the Tenerife populations and there was also departure from HWE for two loci (Appendix 2; see Supplemental Data). This could be associated with the reduced genetic diversity in this latter taxon and the monomorphic population causing an apparent excess of null alleles (Chapuis and Estoup, 2007). Tests for linkage disequilibrium (LD) found two loci pairs in the La Palma populations with significant LD (Appendix 3; see Supplemental Data), however these same loci pairs were non-significant in the Tenerife populations. Amongst

the Tenerife populations several loci pairs had significant LD. This likely reflects high levels of population structure and inbreeding amongst *E. wildpretii* on Tenerife (discussed below), as these factors are known to create LD (Slatkin, 2008).

The infinite allele model (IAM) and stepwise mutation model (SMM) were applied to test for population bottlenecks on Tenerife and La Palma. In a population at mutation-drift equilibrium the probability that a locus shows either an excess or deficit of heterozygosity is approximately equal, whereas populations that have experienced a recent bottleneck event (i.e. a reduction in their effective population size) are expected to show an excess of heterozygosity relative to the observed number of alleles (Cornuet and Luikart, 1996). Under the SMM model, which is considered to be more appropriate for microsatellites, both the Tenerife and La Palma populations show a deficit of heterozygosity which inconsistent with a population bottleneck (Appendix 4; see Supplemental Data). The analysis also found the allele frequency distribution for both populations to be approximately L-shaped, as expected under mutation-drift equilibrium rather than a recent bottleneck. It should be noted however, that bottleneck tests have limited statistical power to detect reductions in population size, especially for smaller sample sizes (Zachariah Peery et al., 2012).

The Shannon's Information Index (I), a measure of diversity that considers the evenness of allele frequencies, is also higher across the La Palma populations (0.885) than the Tenerife populations (0.406) (Table 2). Again, population T1 has a very low Index (I = 0.064), compared to the species average (I = 0.646). Observed (H_o) and expected (H_e) heterozygosity values follow a similar pattern (Table 2). Thus, expected heterozygosity is significantly lower (t = -3.348, df = 3.719, p-value = 0.016) in all the Tenerife populations (0.00-0.176) than the La Palma populations (0.262-0.407). It is notable that population T1 has $H_o = 0$, meaning that it is completely homozygous. Observed heterozygosity is also lower than the expected value for all populations of *E. wildpretii*, indicating some level of inbreeding throughout the species. The fixation index or inbreeding coefficient (F) averages 0.360 across all populations and is lower on La Palma (0.207) than Tenerife (0.567), but there is considerable variation between populations on each island. For example, on Tenerife population T4 has moderate levels of inbreeding (F=0.410), whereas population T1 has an F value of 1, indicating that it is completely inbred.

Results of the AMOVAs (Fig. 2) show that there is some genetic differentiation between Tenerife and La Palma, but there is more molecular variation within (62%) than between

(38%) the islands. When the two subspecies are analysed separately, there is again much more variation within than between populations, with the among group variance considerably higher on Tenerife (21%) than on La Palma (6%). Pairwise F_{ST} ranged from 0.037 to 0.225 within islands, whereas inter-island F_{ST} values are higher and range from 0.235 to 0.514 (Table 3). The overall F_{ST} between Tenerife and La Palma is 0.381 (Fig. 2). On La Palma all pairwise F_{ST} values are ≤0.082, indicating low levels of population structure, whereas on Tenerife pairwise F_{ST} varies from 0.037 to 0.225.

The PCoA (Fig. 3a) separates *E. wildpretii* into two broad clusters along coordinate 1, which explains a large proportion of the variation (32.25%). Individuals of *E. wildpretii* at the lower end of coordinate 2 are quite strongly separated by island, whereas there is some overlap at the higher end of this axis where the points are more dispersed. The plot shows relatively weak clustering by sampling locality. Most individuals of *E. wildpretii* subsp. *trichosiphon* are quite densely clustered together in a single group and show no obvious separation between localities. For subsp. *wildpretii*, there is some separation of T1 and T2 from T4. The PCoA plot of coordinates 1 and 3 (Appendix ~~S2a~~S5a; see Supplemental Data) also shows separation of individuals into two broad clusters corresponding to Tenerife and La Palma, whereas the plot of coordinates 2 and 3 (Appendix ~~S2b~~S5b; see Supplemental Data) shows no obvious clusters. Model-based clustering analysis of coordinates 1 and 2 identified three clusters, based on the Bayesian Information Criterion (Appendix ~~S3a~~S6a; see Supplemental Data). The 41 individuals from La Palma were resolved as a single cluster, while those from Tenerife were separated into two clusters comprising a large dispersed group of 39 individuals and a small tight group of 17 individuals (Appendix ~~S3b~~S6b; see Supplemental Data).

Network analysis

The unrooted network (Fig. 3b) shows that individuals are strongly grouped by island, but individuals from the same population within an island do not often cluster together on the same branch. The network also shows that individuals from La Palma tend to be placed on longer branches than those from Tenerife.

Isolation by distance

A Mantel test of the full *E. wildpretii* dataset indicates that genetic differentiation between populations is consistent with a pattern of IBD ($R^2 = 0.78$, $P = 0.01$). The high R^2 indicates that geographic distance explains a large proportion of the variation in genetic

differentiation. This reflects the large genetic and geographic distance between populations on different islands. The effect of IBD was not statistically significant for the separate Tenerife ($R^2 = 0.36$, $P = 0.18$) or La Palma ($R^2 = 0.44$, $P = 0.12$) datasets, although both showed weak positive correlation between genetic and geographic distance.

Genetic structure

The results of the Structure analysis (Fig. 3c) are congruent with the PCoA results, with the optimal number of clusters (K) being 2 (Appendix ~~S4~~S7; see Supplemental Data) and separating individuals into two clusters which correspond to the islands of Tenerife and La Palma, although some individuals show genetic components from both clusters. Using a cut-off of <90 % ancestry from a single cluster to represent individuals that are admixed, five individuals from Tenerife (representing 8.93 % of the Tenerife sample) and two from La Palma (4.88 %) appear to be admixed. K = 3 receives much weaker support but does reveal some within-island differentiation, with the presence of two genetic clusters within Tenerife (Fig. 3c). Individuals from localities T1 and T4 are mostly assigned to different genetic clusters, while individuals from localities T2 and T3 show a mixture of association to both Tenerife genetic clusters. Based on the 90 % cut-off used above, seven individuals from locality T2 (38.89 %) and three individuals from T3 (20.00 %) appear to be admixed. In contrast, the K = 3 plot shows no genetic differentiation among sampling localities from La Palma. Values of K > 3 show a similar pattern of strong inter-island differentiation with some population structure within Tenerife but not La Palma, although these clustering values are very poorly supported (Appendix ~~S5~~S8; see Supplemental Data).

Gene flow

The estimated number of migrants exchanged per generation (Nm) between sampling localities within each island is significantly higher on La Palma than on Tenerife (Welch Two Sample t-test, t = -4.582, df = 7.100, p-value = 0.001). ~~It is generally assumed that if Nm ≥ 1 (i.e. one or more migrants per generation) the rate of gene flow is sufficient to prevent populations from diverging due to genetic drift (Ellstrand and Elam, 1993).~~ For all pairs of sampling localities within La Palma, Nm is much greater than 1, indicating high levels of gene flow. This suggests that these collection sites do not represent distinct populations and are instead subsamples of a single inter-breeding La Palma population. In contrast, Nm values between sampling localities on Tenerife vary from 0.360 to 2.554, indicating that gene flow is restricted between some populations, (i.e. between T1, T2 and

T4), but is more frequent between others, (i.e. between T3 and T1, T2 and T4). The estimated Nm between Tenerife and La Palma is 0.406, which is sufficiently low as to allow the two island populations to diverge due to drift.

Migration estimates generated using BayesAss are congruent with the Nm estimates and show that rates of contemporary gene flow between the two subspecies of *E. wildpretii* are low, with less than 1% of individuals derived from immigration per generation. Of the Tenerife individuals 0.58% (± 0.57) per generation are estimated to be derived from recent immigration, with an equivalent figure of 0.83% (± 0.81) on La Palma. Individual ancestry estimates show that none of the individuals sampled are likely to have been derived from recent inter-island migration. The probability that an individual is a first- or second-generation immigrant is ≤ 0.061 for all 97 individuals.

Morphology

Field observations reveal a consistent difference in flower colour between plants on the two islands with individuals on Tenerife having red flowers, while those on La Palma have pink flowers (Fig. 1, 4). In both taxa, the flowers turn blue upon drying. Inflorescence shape varies between individuals, but no consistent differences were seen between plants on the two islands.

Differences in leaf morphology between taxa were apparent from herbarium specimens and from cultivated plants. Plants from Tenerife have narrowly oblanceolate to almost linear leaves with a distinct petiole at the base, whereas plants from La Palma have broader leaves which are oblanceolate to narrowly elliptic in shape, with the lamina extending almost to the base (Fig. ~~1,~~ 4).

Examination of floral characters revealed no consistent differences in floral traits. Trichomes were present on the calyx and corolla of flowers of plants from both islands, but there was no obvious difference in the density of pubescence. The length of the stigmatic lobes (potentially a diagnostic character; Bramwell, 1972) ranged from 0.38-0.96 mm for Tenerife (mean = 0.70 ± 0.17 SD) and 0.58-1.30 (mean = 0.85 ± 0.21 SD) for La Palma (Appendix ~~S6~~S9; see Supplemental Data).

Discussion

Inter-island genetic structure

The SSR analysis supports the differentiation of populations of *E. wildpretii* from La Palma and Tenerife ~~and their recognition as distinct taxa~~. In the structure analysis, K = 2 is more strongly supported than any higher values of K, and the clusters defined correspond to the two different taxa recognised. A similar finding is apparent from the PCoA. In the analyses of gene flow and migration it is evident that there is little or no gene flow between Tenerife and La Palma: the estimated number of migrants exchanged per generation (Nm) between Tenerife and La Palma (0.406) is low and consistent with divergence of the two island populations due to drift. Migration estimates indicate that none of the individuals sampled are likely to have been derived from recent inter-island migration. The Mantel test suggests that there is a significant pattern of isolation-by-distance in *E. wildpretii*, with sampling localities on different islands separated by a much greater genetic distance than those on the same island. Geographic distance was able to explain a large proportion of the variation in genetic distance ($R^2 = 0.78$) between localities. This geographic isolation is unsurprising, since the sub-alpine zones of the islands are effectively "islands within islands", separated by 133 km, with no areas of suitable habitat in between, and it is congruent with estimates of inter-island gene flow. In contrast, the closest coastal regions of Tenerife and La Palma are separated by only 85 km. If rates of dispersal are assumed to decline with increasing geographic distance, then higher elevation island habitats should experience lower rates of inter-island migration than low elevation habitats (Steinbauer et al., 2012, 2013). All else being equal, greater genetic distance might therefore be expected between sub-alpine taxa on different islands, than between taxa that occur in low elevation habitats separated by smaller geographic distances. However, there is no empirical evidence comparing genetic structure in low and high elevation species in the Canaries to support the effect of elevation on gene flow.

In addition to the significant geographic barrier, *E. wildpretii* also lacks any obvious means of exchanging genetic material over long distances. The plants produce fruit in the form of a dry four-seeded nutlet (Bramwell, 1972), which does not have any obvious features to facilitate dispersal by animals (e.g. fleshy fruit, or hooked seeds) or by wind (e.g. light dust-like seeds, or winged seeds) (Howe and Smallwood, 1982). *Echium wildpretii* is also

primarily bee-pollinated (Dupont et al., 2004). Most species of bees have a maximum foraging range of up to a few kilometres, with honeybees (*Apis mellifera*) able to travel up to 14 km (Zurbuchen et al., 2010), so it is unlikely that they could significantly facilitate inter-island gene flow. *Echium wildpretii* is partially bird-pollinated on Tenerife (Dupont et al., 2004), which could provide a mechanism for long distance pollen dispersal between islands, but the lack of evidence of bird pollination on La Palma (Valido et al., 2002) makes this unlikely.

Reproductive incompatibility between taxa is another mechanism that could limit gene flow between the islands, however many Canary Island endemic plant species have weak reproductive barriers and will hybridize freely when brought into contact (van Hengstum et al., 2012). This seems to be true of *Echium*, with SSR markers indicating that *E. wildpretii*, *E. pininana* and *E. simplex* hybridise with each other when grown in cultivation (Maunder, 1997). As *E. wildpretii* has sufficiently weak reproductive barriers to allow hybridisation with other *Echium* species (for example a natural hybrid with *E. auberianum* is known; see Schönfelder et al., 1993), it seems unlikely that any degree of intrinsic reproductive incompatibility could have developed between Tenerife and La Palma plants since they diverged.

Limited gene flow due to geographic distance therefore seems to be the major driver of reproductive isolation between the islands. However, it is possible that the subspecies are differentially adapted, at least with respect to flower colour. *Echium wildpretii* has evolved from a primarily insect-pollinated lineage and is supported as sister to the blue-flowered *E. pininana* (Böhle et al., 1996; Valido et al., 2002; Graham et al, in prep.). The origin of red flowers in *E. wildpretii* subsp. *wildpretii*, a classic feature of the bird pollination syndrome (Cronk and Ojeda, 2008), appears to be associated with a shift to partial bird pollination on Tenerife (Dupont et al., 2004). There is no evidence that the pink-flowered *E. wildpretii* subsp. *trichosiphon* on La Palma is visited by nectar-feeding birds (Dupont et al., 2004). Field experiments demonstrate that pollinator visitation significantly increases seed set in *E. wildpretii* subsp. *wildpretii* (Sedlacek et al., 2012), but it is not entirely clear what, if any, selective advantage is provided by birds over native insect species as pollen vectors in *Echium* (Jaca et al., 2018). The pollination biology of *E. wildpretii* subsp. *trichosiphon* has been much less studied.

Patterns of genetic diversity within islands

The results also revealed differences between islands in the partitioning of genetic variation within islands. Firstly, the La Palma populations are more genetically diverse, exhibiting higher allele diversity and Shannon's Information Index than the Tenerife subspecies (Table 2). The network analysis (Fig. 3) also points to the greater diversity in La Palma wherein branch lengths among the La Palma individuals are longer than among the individuals from Tenerife.

The higher level of genetic diversity observed in La Palma compared to Tenerife plausibly reflect the evolutionary origins of this species. Phylogenetic evidence supports *E. pininana*, a laurel forest endemic species from La Palma, as the sister species of the *E. wildpretii* lineage (Böhle et al, 1996; Garcia-Maroto et al, 2009; Graham et al, in prep). As the *E. pininana* + *E. wildpretii* lineage is nested within a Macaronesian clade of mostly low-mid altitude species, the most parsimonious evolutionary scenario is upslope colonisation and ecological speciation from the laurel forest to the sub-alpine zone on La Palma, with subsequent dispersal of *E. wildpretii* from La Palma to Tenerife. Such a scenario ~~would~~could account for the lower genetic diversity observed in the Tenerife subspecies of *E. wildpretii*, since the colonisation of Tenerife from La Palma is likely to have involved a ~~significant~~ genetic bottleneck effect~~.~~ (although we did not find a significant bottleneck when we tested for this). A La Palma origin for the *E. wildpretii* lineage is also consistent with morphological observations, as the La Palma plants are more similar to *E. pininana*, which has pale blue-lilac flowers and lanceolate leaves than are the Tenerife plants.

The contrasting geological histories of the sub-alpine zone of the two islands may also have impacted the diversity levels observed. On Tenerife, the sub-alpine zone has undergone a number of major geological events, with the current Teide-Pico Viejo volcano forming ca. 30 K years ago and the final eruption inside the caldera occurring as recently as 1798 (Carracedo et al., 2007). This may have either made the sub-alpine zone of Tenerife unsuitable for establishment until relatively recently, or if *E. wildpretii* colonised early it will likely have experienced several reductions in population size as areas of habitat were destroyed by eruptions and landslides, both of which could account for the impoverished levels of genetic diversity seen in *E. wildpretii* on Tenerife. In contrast, whilst La Palma is considered a younger island than Tenerife, all recent eruptive activity on La Palma has taken place in the Southern Cumbre Vieja region rather than the central caldera (where *E. wildpretii* is found), which was formed by a massive landslide approx.

560 K years ago (Carracedo et al., 1999). This geological history means that the sub-alpine zone of La Palma may have been suitable for plants relatively early compared to Tenerife, and the lack of recent eruptive activity means that major bottleneck events are less likely to have occurred.

Very few species endemic to the sub-alpine zones of the Canaries have undergone population genetic analysis. However, *Bencomia exstipulata* Svent., a rare species endemic to sub-alpine zones of Tenerife and La Palma and a member of a monophyletic Macaronesian radiation (Helfgott et al., 2000) shows a similar pattern. Only one small natural population remains on each island, although reintroduction efforts have established several new populations. Microsatellite analysis revealed high genetic diversity in the natural population on La Palma, but much lower diversity on Tenerife (González-Pérez et al., 2009). There is evidence that the species was previously more abundant on Tenerife~~ and it is~~. It has been hypothesised that the lower heterozygosity observed in the Tenerife population was due to a bottleneck effect caused by volcanic activity, which likely killed many individuals and reduced the area of suitable habitat (González-Pérez et al., 2009; Marrero et al., 2019~~) but~~). However, it is notable that in the *E. wildpretii* lineage we also observe a similar pattern with heterozygosity significantly lower in all the Tenerife populations than the La Palma populations and colonisation of Tenerife from La Palma ~~would~~could also ~~be consistent with~~explain the patterns observed in ~~this~~both species.

We also see differences between Tenerife and La Palma in the genetic structure between populations of *E. wildpretii*. Pairwise $F_{ST} \leq 0.082$ on La Palma indicates low levels of population structure. On Tenerife pairwise F_{ST} varies from 0.037 to 0.225, indicating that the degree of genetic differentiation between populations is much more variable, and in one case (between populations T1 and T4) is almost as high as the differentiation seen between some inter-island pairs. This pattern is supported by model-based clustering analysis of the PCoA results which resolves the La Palma individuals as a single cluster, whereas the Tenerife individuals are separated into two clusters. The splitting of individuals from Tenerife into two clusters appears to be due to a tight cluster of points in the PCoA, which corresponds to several genetically similar individuals from population T1. Only three distinct multilocus genotypes are observed amongst the eleven individuals sampled in population T1, and nine of these individuals share an identical genotype. Compared to La Palma, partitioning of genetic diversity among populations on Tenerife is highly unequal. Locality T1 is exceptional for its very low genetic diversity and extremely high rate of inbreeding (Fixation index = 1). The very low level of allelic diversity and

heterozygosity observed in this locality suggests that it has undergone a dramatic bottleneck event. This population may have been established from a single individual, as each inflorescence can produce many seeds.

Nm values >1 are found in the gene flow analysis for La Palma, indicating high levels of gene flow between sampling sites and suggesting that they are subsamples of a single inter-breeding La Palma population. In contrast, the variable and in some cases low Nm values on Tenerife indicate restricted gene flow between some populations, and could be sufficiently low to produce genetically differentiated populations. The PCoA plot similarly indicates that there is no separation of populations in La Palma but some structuring in Tenerife (Fig. 3). This is also found in the AMOVA wherein some genetic differentiation among the Tenerife populations is found, but very little among those on La Palma. Although a test for isolation-by-distance was not significant for the Tenerife dataset, the observed pattern of differentiation does seem to be consistent with a reduction in gene flow correlated with increasing geographic distance. The apparently stronger effect of geographic distance on populations in Tenerife than La Palma may be due to the difference in size of the sub-alpine zone of each island. The area occupied on Tenerife is larger, with the most distant sampling localities separated by approximately 13.5 km (distance between localities T1 and T4), compared to 5 km on La Palma. Weaker population structure on La Palma may also result from efforts to conserve *E. wildpretii* subsp. *trichosiphon* through seed sowing and the re-introduction of cultivated plants (A. Palomares, pers. com.)

There are some similarities in population genetic structure between *E. wildpretii* and the sub-alpine endemic *Viola cheiranthifolia* Bonpl. from Tenerife, in which there is strong genetic differentiation between the two populations - those close to Teide volcano and those occurring along the Las Cañadas wall (Rodríguez-Rodríguez et al., 2019). Indeed, the two *Viola* populations in the Cañadas region have recently been separated as distinct species (Marrero Gómez et al., 2020). However, *V. cheiranthifolia* is restricted to very high elevations, with individuals on Teide and the caldera wall separated by a large expanse of inhospitable Caldera floor. It is therefore hypothesised that gene flow between these areas is limited by the ability of pollinating insects to travel long distances (Rodríguez-Rodríguez et al., 2019). There may well be different factors driving population genetic structure in these two taxa, as *E. wildpretii* has a more contiguous distribution across the caldera floor and is not solely reliant on insects for pollination (Dupont et al., 2004).

Taxonomic and conservation implications

The results presented here suggest that the plants from Tenerife and La Palma represent two independently evolving ~~taxa~~lineages. In addition to the strong genetic differentiation and limited gene flow between the two taxa, they are also morphologically distinct. The distinctiveness of these taxa has long been recognised but their taxonomic status has changed over time, with the La Palma taxon variously recognised as either a species, subspecies, or a variety.

The Tenerife taxon was first referred to as *E. bourgaeanum* by Bourgeau in a series of specimen labels in the mid-19th century. This name was validly published by Coincy in 1903, but the name *E. wildpretii* was published for the same taxon a year earlier by J. D. Hooker (1902) and therefore takes priority. The La Palma taxon was first described by Sprague (1914) under the species name *E. perezii*. He considered that it differed from the Tenerife taxon in its lax inflorescence , decurrent leaf lamina base, longer style arms, and paler corolla. The La Palma taxon was subsequently reduced to varietal level as *E. bourgaeanum* var. *trichosiphon* by E. Sventenius in Ceballos & Ortuño (1951). In his revision of Macaronesian *Echium* (1972) Bramwell recognised it at subspecies rank and made the new combination *Echium wildpretii* subsp. *trichosiphon.*

Bramwell (1972) differentiated the two taxa using the shape of the inflorescence, pubescence of the calyx and corolla, width of the corolla lobes, and the length of stigmatic lobes. However, his taxonomic revision makes no mention of the differences in leaf morphology or corolla colour previously described by Sprague (1914). Bramwell (1972) describes the stigmatic lobes of plants from La Palma as being “at least twice as long” as those in plants from Tenerife. Morphological measurements presented here do not support Bramwell’s observation, since, whilst there is a significant difference in the mean stigma lobe length of material from La Palma and Tenerife (0.85 versus 0.70 mm respectively), there is substantial variation in lobe length within and between individuals from the same island (Appendix ~~S3~~S9; see Supplemental Data). From our morphological study, there is no support for the differentiation of La Palma and Tenerife plants based on inflorescence shape or the morphology of floral parts. Nevertheless, they can be clearly distinguished by flower colour and leaf shape. Specifically, plants on La Palma have pink flowers and broader leaves with the leaf lamina extending to near the base, whereas plants on Tenerife have red flowers and narrower leaves with a distinct petiole (see **Error! Reference source not found.**).

Based on the genetic and morphological differentiation demonstrated in this study, we propose recognising plants from Tenerife and La Palma as distinct species. Published names at species rank already exist for both taxa: *E. wildpretii* is used for plants from Tenerife and *E. perezii* for plants from La Palma.

Most individuals of these species occur in designated protected areas (either Cañadas del Teide or Caldera de Taburiente National Parks), so their habitat is at low risk from land use change. There have also been ~~some~~active conservation efforts ~~to actively conserve *E. wildpretii*~~ on both islands by installing fencing and by planting out cultivated plants (A. Palomares, pers. com.). However, ~~*E. wildpretii*~~both species may still be at risk from some forms of human activity. For example, there is a long tradition of beekeepers taking their hives of honeybees to Las Cañadas on Tenerife to utilise the nectar resource provided by *E. wildpretii* and other flowering plants during the summer months (Dupont et al., 2004). The non-native honeybees compete with native bee and bird species for nectar resources and their different foraging behaviour may alter patterns of gene flow. Honeybees tend to visit fewer individual plants than native bees (Dupont et al., 2004) and may be contributing to high rates of inbreeding observed in *E. wildpretii* on Tenerife.

Inbreeding depression in *E. wildpretii* in Tenerife is more severe when plants are under drought stress (Sedlacek et al., 2012) and climate simulations for the Canary Islands predict drier conditions at high elevations (Sperling et al., 2004). Future conservation efforts should therefore focus on maintaining genetic diversity and avoiding inbreeding, with the aim of maximising its genetic potential to adapt to future environmental change (Pauls et al., 2013) and avoid the synergistic effects of climate change and inbreeding depression (Sedlacek et al., 2012).

Based on the IUCN Red List criteria (IUCN, 2012) both *E. wildpretii* and *E. perezii* should be assessed as "Vulnerable" under criterion D.1: "Population with a very restricted area of occupancy (typically less than 20 km^2) or number of locations (typically five or fewer) such that it is prone to the effects of human activities or stochastic events within a very short time period in an uncertain future, and is thus capable of becoming Critically Endangered or even Extinct in a very short time period." There is no evidence for ongoing population decline in either species, so they do not qualify for classification in the more severe risk categories. However, their very restricted areas of occupation and small number of locations means that *E. wildpretii* and *E. perezii* could quickly become threatened with extinction by stochastic events such as wildfires, landslides or volcanic activity. In the

longer term the level of threat to both species may increase as climate change reduces the area of suitable habitat.

Conclusions

The results presented in this study represent the first detailed population genetic analysis of a member of the Macaronesian *Echium* clade and provide new insights into the evolution of the Canarian sub-alpine flora. Microsatellite analysis reveals that the plants on Tenerife and La Palma (*E. wildpretii* subsp. *wildpretii* and subsp. *trichosiphon*, as currently recognised) represent two genetically distinct taxa which are reproductively isolated by a geographic barrier. The genetic and morphological distinctiveness of the two taxa supports their recognition as separate species, as proposed here. Higher genetic diversity in the La Palma species is consistent with an origin of the lineage on this island via upslope ecological speciation, followed by dispersal to the sub-alpine zone of Tenerife where the plants show reduced genetic diversity.

Understanding the genetic status of sub-alpine taxa such as in the *E. wildpretii* lineage is also important for their conservation, as they are predicted to be extremely vulnerable to future climate change, and their level of genetic diversity may give an indication of their capacity to adapt and persist.

Taxonomic treatment

Key to species:

Leaves linear to narrowly oblanceolate, with a distinct petiole at base; corolla red……*Echium wildpretii*

Leaves oblanceolate to narrowly elliptic, with lamina extending almost to base; corolla pink.....*Echium perezii*

Echium wildpretii Hook.f., Bot. Mag. 128: t. 7847 (1902).

Type: Figure 7847 of *Curtis's Botanical Magazine* (1902) which accompanies the original description (lecto designated by Bramwell, *Lagascalia*, 2 (1), 37-115 (1972)).

=*Echium bourgaeanum* Webb ex Coincy, Bull. Herb. Boissier Ser. 2, iii. 275 (1903).

Type: Tenerife, *Bourgeau* Pl. Can. Exsicc. No 1436, 1856 (lectotype P00571938 image!; designated here; isolectotypes: P00571928 image!, P00571933 image!, P00571934 image!, P00571935 image!, P00571936 image!, P00571937 image!, P00571939 image!)

Other specimens seen:

Canary Islands. Tenerife: Las Cañadas, 14 May 2016, *Graham & Carine* 52 (BM000828802*; ORT); Las Cañadas, Mirador de San Jose, 14 May 2016, *Graham & Carine* 56 (BM000828806*; ORT): Las Cañadas, Los Azulejos, Mirador Llano de Ucanca, 14 May 2016, *Graham & Carine* 58 (BM000828808*; ORT); Las Cañadas, southern rim of caldera, South of main road from Parador towards Vilaflor, 18 May 2016, *Graham, Carine & White* 104 (BM000828853*; ORT); cultivated, s.d., *Pérez*, 275 (BM, P00571931 image!); seeds from Teide, cultivated at Reading, 12 Jun 2003*, s.col.* RDG120603 (RNG, 2 sheets); Below Mt. Teide between Roques de Garcia and LLanos de Ucanca, 8 Aug 2001, *De Silva, Priestley & Santos* RDG080801 (RNG); Las Cañadas, Llano de Ucanca, 1965, *Lems* 2613 (RNG); Cañadas del Teide, near the Parador National, 6th July 1984, *G. van Buggenhout* 86486 (P00571941 image!); Ad rupes montis de la Fontaleza Cañadas del Teyde, 5th Sept 1845, *E. Bourgeau* 895 (P00571932 image!); Puerto Orotava (cultivated), s.d., *G.V. Pérez*, s.n. (P00571929 image!, P00571930 image!); Chasna, Santa Verde de Tenerife, 27th June 1855, *H. de la Perraudière*, s.n. (P00571940 image!). Cultivated, s.d., *G. Hibon*, s.n. (P00571941 image!). Gran Canaria (cultivated), 26 April 1982, *B. de Retz* 82882 (P00571944 image!).

Notes:

The designation *E. bourgaeanum* was first used in the mid-19th century on the labels of a series of exsiccatae distributed by Bourgeau but it was not effectively published. It was subsequently (and erroneously) listed in synonymy with *E. auberianum* in the account of Christ (1887). The name was validly published by Coincy (1903) citing two collections by Bourgeau: no. 895 collected in 1845 and no. 1436 collected in 1855. Since Coincy worked

at Paris we have selected P00571938, a duplicate of *Bourgeau* 1436 as lectotype. It is a specimen that was annotated by Coincy in 1902.

The name *E. wildpretii* was published by J. D. Hooker in 1902 and therefore takes priority over *E. bourgaeanum*. Bramwell indicated that the type of *E. wildpretii* was the illustration in Curtis's Botanical Magazine (1902) which accompanies the original description which we consider to be a lectotypification.

Echium perezii Sprague, Bull. Misc. Inform. Kew. 210 (1914)

Type: Tenerife (in cult.): cultivated by Dr. Pérez in his garden at Villa Orotava, Tenerife. s. coll., s.n. 4 June 1913. (K! lectotype designated here).

=*E. bourgaeanum* var. *trichosiphon* Svent. In Ceballos & Ortuño, *Veg. Fl. Forest. Canar. Occ.*: 409 (1951)

Type: La Palma, sobre El Paso, June 1950, *F. Ortuño* (ORT1952!)

≡*Echium wildpretii* subsp. *trichosiphon* (Svent.) Bramwell, Lagascalia 2(1): 78 (1972).

Other specimens seen:

Canary Islands. La Palma: By road up to Roque de los Muchachos from Garafia, 24 May 2016, *Graham & White* 138 (BM000828889* ORT); Caldera de Taburiente, on path from Roque de los Muchachos to Torre el Time, 24 May 2016, *Graham & White* 139 (BM000828890; ORT); Roque de los Muchachos, by road up to car park, 24 May 2016, *Graham & White* 146 (BM000828897*; ORT); Near Roque de los Muchachos, below observatory station, 26 May 2016, *Graham & White* 154 (BM000828909*; ORT); Caldera de Taburiente below Roque de Las Muchachos, 3 Aug 2001, *De Silva, Culham, Pitman, Dyga & Priestley* RDG030801 (RNG, 2 sheets).

Notes:

In the protologue of *E. perezii*, Sprague refers to (i) material cultivated in Tenerife by Pérez which flowered in 1913, and (ii) to material cultivated at Kew from seeds sent from Tenerife by Pérez which flowered in 1914. The specimen cultivated at Kew is extremely limited, comprising a few thyrses and an inflorescence bract. The material cultivated by Pérez is more extensive. We have selected it here as lectotype, as it is clear that Sprague examined this material for the protologue.

Whilst the herbarium specimens are clearly referable to this species, there is some ambiguity in the original information and circumscription of this taxon. The locality given as the original site of collection (Punta Llana, Barranco del Agua) is typical of *E. pininana;* it is not where *E. perezii* would be found. Furthermore, an account of *E. perezii* by Sprague (1914) in *Kew Misc. Bull.* includes a photograph of a specimen growing in Pérez's garden that seems likely to be a hybrid between *E. pininana* and *E. perezii*. An illustration of this species in Figure 8617 of *Curtis's Bot. Mag.* (1915) is largely consistent with the taxon as circumscribed here but includes a sketch of the habit that is not referable to *E. perezii* and also shows little resemblance to the plant photographed in 1914. This may have been a naïve interpretation drawn without sight of living plants.

E. coeleste Stapf, *Bot. Mag.* 148: t. 8977 (1923) was treated in synonymy of *E. wildpretii* subsp. *trichosiphon* by Bramwell in 1972. Described from cultivation at Kew by Stapf (1923) the only specimen at Kew is a single dissected flower with a note to indicate that the plant died before a specimen could be made (Cult. Hort. Kew, s.d., *s.col.*, s.n., K!). The plate accompanying the protologue shows a plant with blue flowers, broad leaves and bracts, and a distinct leafless base to the stem that clearly place it in *E. pininana*.

Specimens marked with an asterisk (*) were sampled for flower dissection (see the Taxonomic Treatment for details).

Acknowledgements

The authors thank Arnoldo Santos-Guerra and Oliver White for their assistance in carrying out fieldwork. The authors would also like to thank Sandra Nogué, Stephen Harris and members of the Chapman lab for helpful comments on this manuscript.

Author contributions

R. Graham and M. Carine undertook fieldwork, and A. Reyes-Betancort assisted seed collection on Tenerife. R. Graham undertook the lab work and all data analysis. R. Graham drafted the manuscript. M. Chapman, A. Reyes-Betancort and M. Carine commented on the manuscript and contributed to revisions.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Supplemental data

Supplemental data for this article can be accessed here:

Funding

This work was supported by the Natural Environmental Research Council through a SPITFIRE DTP PhD studentship to REG [grant number NE/L002531/1].

References

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology* 215: 403—10.

Böhle, U. R., Hilger, H. H. & Martin. W. F. (1996). Island colonization and evolution of the insular woody habit in *Echium* L.(Boraginaceae). *Proceedings of the National Academy of Sciences* 93: 11740—11745.

Bramwell, D. 1972. A revision of the genus *Echium* in Macaronesia. *Lagascalia* 2: 37—115.

Carracedo, J. C., Day, S. J., Guillou, H. & Gravestock, P. (1999). Later stages of volcanic evolution of La Palma, Canary Islands: Rift evolution, giant landslides, and the genesis of the Caldera de Taburiente. *Bulletin of the Geological Society of America* 111: 755—768.

Carracedo, J. C., Badiola, E. R., Guillou, H., Paterne, M., Scaillet, S., Torrado, F. P., Paris, R., Fra-Paleo, U. & Hansen, A. (2007). Eruptive and structural history of Teide Volcano and rift zones of Tenerife, Canary Islands. *Geological Society of America Bulletin* 119: 1027—1051.

Carvalho, J. A. & Pontes, T. (2010). A new species from the island of Porto Santo (Madeira Archipielago). *Anales del Jardín Botánico de Madrid* 67: 87—96.

Chapuis, M. P., & Estoup, A. (2007). Microsatellite null alleles and estimation of population differentiation. *Molecular Biology and Evolution* 24: 621—631.

Chase, M. W. & Hills, H. H. (1991). Silica Gel: An Ideal Material for Field Preservation of Leaf Samples for DNA Studies. *Taxon* 40: 215—220.

Cornuet, J. M., & Luikart, G. (1996). Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics* 144: 2001—2014.

Cronk, Q. C. B. & Ojeda, I. (2008). Bird-pollinated flowers in an evolutionary and molecular context. *Journal of Experimental Botany* 59: 715—727.

Dakin, E. E., & Avise, J. C. (2004). Microsatellite null alleles in parentage analysis. *Heredity*

93: 504—509.

Del-Arco, M., Pérez-de-Paz, P. L., Acebes, J. R., González-Mancebo, J. M., Reyes-Betancort, J. A., Bermejo, J. A., De-Armas, S. & González-González, R. (2006). Bioclimatology and climatophilous vegetation of Tenerife (Canary Islands). *Annales Botanici Fennici* 43: 167—192.

Dirnböck, T., Essl, F. & Rabitsch, W. (2011). Disproportional risk for habitat loss of high-altitude endemic species under climate change. *Global Change Biology* 17: 990—996.

Dupont, Y. L., Skov, C., Hansen, D. M., Valido, A. & Olesen, J. M. (2004). Impact of introduced honey bees on native pollination interactions of the endemic *Echium wildpretii* (Boraginaceae) on Tenerife, Canary Islands. *Biological Conservation* 118: 301—311.

Earl, D. A. & von Holdt, B. M. (2012). STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* 4: 359—361.

Ellis, J. R. & Burke, J. M. (2007). EST-SSRs as a resource for population genetic analyses. *Heredity* 99: 125—132.

Ellstrand, N. C. & Elam, D. R. (1993). Population Genetic Consequences of Small Population Size: Implications for Plant Conservation. *Annual Review of Ecology and Systematics* 24: 217—242.

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Molecular Ecology* 14: 2611—2620.

Fernández-Palacios, J. M., Otto, R., Thebaud, C. & Price, J. (2014). Overview of habitat history in subtropical oceanic island summit ecosystems. *Arctic, Antarctic, and Alpine Research* 46: 801—809.

García-Maroto, F., Mañas-Fernández, A., Garrido-Cárdenas, J. A., Alonso, D. L., Guil-Guerrero, J. L., Guzmán, B. & Vargas, P. (2009). Δ6-Desaturase sequence evidence for explosive Pliocene radiations within the adaptive radiation of Macaronesian *Echium* (Boraginaceae). *Molecular Phylogenetics and Evolution* 52: 563—574.

Gieger, T. & Leuschner, C. (2004). Altitudinal change in needle water relations of *Pinus canariensis* and possible evidence of a drought-induced alpine timberline on Mt. Teide, Tenerife. *Flora - Morphology, Distribution, Functional Ecology of Plants* 199: 100—109.

Gómez Campo, C. (1996). *Libro rojo de especies vegetales amenazadas de las Islas Canarias*. Viceconsejería de Medio Ambiente. Consejería de Política Territorial. Gobierno de Canarias.

González-Pérez, M. A., Lledo, M. D., Lexer, C., Fay, M., Marrero, M., Banares-Baudet, A., Carque, E. & Sosa, P. A. (2009). Genetic diversity and differentiation in natural and reintroduced populations of *Bencomia exstipulata* and comparisons with *B. caudata* (Rosaceae) in the Canary Islands: an analysis using microsatellites. *Botanical Journal of the Linnean Society* 160: 429—441.

Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden, J., Couger, M. B., Eccles, D., Li, B., Lieber, M. & MacManes, M. D. (2013). *De novo* transcript sequence reconstruction from RNA-Seq: reference generation and analysis with Trinity. *Nature Protocols* 8: 1494—1512.

Helfgott, D. M., Francisco-Ortega, J., Santos-Guerra, A., Jansen, R. K. & Simpson, B. B. (2000). Biogeography and Breeding System Evolution of the Woody *Bencomia* Alliance (Rosaceae) in Macaronesia Based on ITS Sequence Data. *Systematic Botany* 25: 82—97.

van Hengstum, T., Lachmuth, S., Oostermeijer, J. G. B., den Nijs, J. C. M., Meirmans, P. G., & van Tienderen, P. H. (2012). Human-induced hybridization among congeneric endemic plants on Tenerife, Canary Islands. *Plant Systematics and Evolution* 298: 1119—1131.

Hodel, R. G., Gitzendanner, M. A., Germain-Aubrey, C. C., Liu, X., Crowl, A. A., Sun, M., Landis, J. B., Segovia-Salcedo, M. C., Douglas, N. A., Chen, S. & Soltis, D. E. (2016). A new resource for the development of SSR markers: Millions of loci from a thousand plant transcriptomes. *Applications in Plant Sciences* 4: 1600024.

Howe, H. F. & Smallwood, J. (1982). Ecology of Seed Dispersal. *Annual Review of Ecology and Systematics* 13: 201—228.

Irl, S. D. H., Harter, D. E. V., Steinbauer, M. J., Gallego Puyol, D., Fernández-Palacios, J. M., Jentsch, A. & Beierkuhnlein, C. (2015). Climate vs. topography - spatial patterns of plant species diversity and endemism on a high-elevation island. *Journal of Ecology* 103: 1621—1633.

IUCN. (2012). *IUCN Red List Categories and Criteria*. IUCN, Gland, Switzerland.

Jaca, J., Nogales, M. & Traveset, A. (2018). Reproductive success of the Canarian *Echium simplex* (Boraginaceae) mediated by vertebrates and insects. *Plant Biology* 21: 216—226.

Jakobsson, M. & Rosenberg, N. A. (2007). CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* 23: 1801—1806.

Kalinowski, S. T., Taper, M. L., & Marshall, T. C. (2007). Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Molecular Ecology* 16: 1099—1106.

Kopelman, N. M., Mayzel, J., Jakobsson, M., Rosenberg, N. A. & Mayrose, I. (2015). Clumpak: A program for identifying clustering modes and packaging population structure inferences across K. *Molecular Ecology Resources* 15: 1179—1191.

Marrero, M. V., Oostermeijer, G., Nogales, M., van Hengstum, T., Saro, I., Carqué, E., Sosa, P. A., & Bañares Á. (2019). Comprehensive population viability study of a rare endemic shrub from the high mountain zone of the Canary Islands and its conservation implications. *Journal for Nature Conservation* 47: 65—76.

Marrero Gómez, M. V., Martín Esquivel, J. L., Docoito Díaz, J. R. & Suárez Izquierdo, M. (2020). *Viola guaxarensis* (Violaceae): a new Viola from Tenerife, Canary Islands, Spain. *Willdenowia* 50: 13–21

Maunder, M. (1997). *Botanic garden response to the biodiversity crisis: implications for threatened species management.* Unpublished doctoral thesis, University of Reading.

Moreno, J. C. (2008). *Lista Roja de Flora Vascular Española.* Dirección General de Medio Natural y Política Forestal (Ministerio de Medio Ambiente, y Medio Rural y Marino, y Sociedad Española de Biología de la Conservación de Plantas), Madrid.

Navascués, M., Vaxevanidou, Z., González-Martínez, S. C., Climent, J., Gil, L. & Emerson, B. C. (2006). Chloroplast microsatellites reveal colonization and metapopulation dynamics in the Canary Island pine. *Molecular Ecology* 15: 2691—2698.

Page, A. F. & Minocha, S. C. (2005). Analysis of gene expression in transgenic plants. *Methods in Molecular Biology* 286: 291—311.

Paradis, E., Claude, J. & Strimmer, K. (2004). APE: Analyses of phylogenetics and evolution in R language. *Bioinformatics* 20: 289—290.

Pauls, S. U., Nowak, C., Bálint, M. & Pfenninger, M. (2013). The impact of global climate change on genetic diversity within populations and species. *Molecular Ecology* 22: 925—946.

Peakall, R. & Smouse, P. E. (2012). GenALEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics* 28: 2537—2539.

Peery, M. Z., Kirby, R., Reid, B. N., Stoelting, R., Doucet-Bëer, E. L. E. N. A., Robinson, S., ... & Palsbøll, P. J. (2012). Reliability of genetic bottleneck tests for detecting recent population declines. *Molecular Ecology* 21: 3403—3418.

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

Puppo, P., Curto, M., & Meimberg, H. (2016). Genetic structure of *Micromeria* (Lamiaceae) in Tenerife, the imprint of geological history and hybridization on within-island diversification. *Ecology and Evolution* 6: 3443—3460.

Rambaut, A. (2016). FigTree v1.4.3. Retrieved from: http://tree.bio.ed.ac.uk/software/figtree/ (accessed: 30 January 2018).

Rambaut, A. & Drummond, A. J. (2013). Tracer v1.6. Retrieved from: http://tree.bio.ed.ac.uk/software/tracer/. (accessed: 14 September 2019).

Raymond, M. (1995). GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity* 86: 248—249.

Rodríguez-Rodríguez, P., de Castro, A., Segui, J., Traveset, A., & Sosa, P. A. (2019). Alpine species in dynamic insular ecosystems through time: conservation genetics and niche

shift estimates of the endemic and vulnerable *Viola cheiranthifolia*. *Annals of Botany* 123: 505—519.

Rousset, F. (2008). genepop'007: a complete re-implementation of the genepop software for Windows and Linux. *Molecular Ecology Resources* 8: 103—106.

RStudio Team. (2016). RStudio: Integrated Development for R. Version 1.0. 136. RStudio, Inc., Boston, MA. Retrieved from: http://www.rstudio.com. (accessed: 30 January 2018).

Santos-Guerra, A. (1983). *Vegetación y flora de La Palma.* Santa Cruz de Tenerife: Editorial Interinsular Canaria.

Schuelke, M. (2000). An economic method for the fluorescent labeling of PCR fragments. *Nature Biotechnology* 18: 233–234.

Schönfelder, P., León Arencibia, M. C. & Wildpret, W. (1993). Catálogo de la Flora vascular de la isla de Tenerife (Islas Canarias). *Itinera Geobotanica* 7: 375-404.

Scrucca, L., Fop, M., Murphy, T. B., & Raftery, A. E. (2016). Mclust 5: Clustering, classification and density estimation using Gaussian finite mixture models. *R Journal* 8: 289.

Sedlacek, J., Schmid, B., Matthies, D., & Albrecht, M. (2012). Inbreeding depression under drought stress in the rare endemic *Echium wildpretii* (Boraginaceae) on Tenerife, Canary islands. *PLoS ONE* 7: e47415.

Selkoe, K. A. & Toonen, R. J. (2006). Microsatellites for ecologists: A practical guide to using and evaluating microsatellite markers. *Ecology Letters* 9: 615—629.

Shiposha, V., Catalán, P., Olonova, M., & Marques, I. (2016). Genetic structure and diversity of the selfing model grass *Brachypodium stacei* (Poaceae) in Western Mediterranean: out of the Iberian Peninsula and into the islands. *PeerJ* 4: e2407.

Slatkin, M. (2008). Linkage disequilibrium—understanding the evolutionary past and mapping the medical future. *Nature Reviews Genetics* 9: 477—485.

Smith, A. P. & Young, T. P. (1987). Tropical Alpine Plant Ecology. *Annual Review of Ecology and Systematics* 18: 137—158.

Sosa, P. A., González-Pérez, M. Á., Moreno, C., & Clarke, J. B. (2010). Conservation genetics of the endangered endemic *Sambucus palmensis* Link (Sambucaceae) from the Canary Islands. *Conservation Genetics* 11: 2357—2368.

Sperling, F. N., Washington, R., & Whittaker, R. J. (2004). Future climate change of the subtropical North Atlantic: Implications for the cloud forests of Tenerife. *Climatic Change* 65: 103–123.

Sprague, T. A. (1914). *Echium perezii*. *Bulletin of Miscellaneous Information, Royal Botanic Garden, Kew* 1914: 210.

Steinbauer, M. J., Field, R., Grytnes, J. A., Trigas, P., Ah-Peng, C., Attorre, F., Birks, H. J. B., Borges, P. A., Cardoso, P., Chou, C. H. & De Sanctis, M. (2016). Topography-driven isolation, speciation and a global increase of endemism with elevation. *Global Ecology and Biogeography* 25: 1097—1107.

Steinbauer, M. J., Irl, S. D. H. & Beierkuhnlein, C. (2013). Elevation-driven ecological isolation promotes diversification on Mediterranean islands. *Acta Oecologica* 47: 52—56.

Steinbauer, M. J., Otto, R., Naranjo-Cigala, A., Beierkuhnlein, C., & Fernández-Palacios, J. M. (2012). Increase of island endemism with altitude - speciation processes on oceanic islands. *Ecography* 35: 23—32.

Suchodoletz, H. von, Glaser, B., Thrippleton, T., Broder, T., Zang, U., Eigenmann, R., Kopp, B., Reichert, M. & Ludwig, Z. (2013). The influence of Saharan dust deposits on La Palma soil properties (Canary Islands, Spain). *Catena* 103: 44—52.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B. C., Remm, M. & Rozen, S. G. (2012). Primer3-new capabilities and interfaces. *Nucleic Acids Research* 40: e115.

Valido, A., Dupont, Y. L. & Hansen, D. M. (2002). Native birds and insects, and introduced honey bees visiting *Echium wildpretii* (Boraginaceae) in the Canary Islands. *Acta Oecologica* 23: 413—419.

White, O. W., Doo, B., Carine, M. A. & Chapman, M. A. (2016). Transcriptome sequencing and simple sequence repeat marker development for three Macaronesian endemic plant species. *Applications in plant sciences* 4: 1600050.

White, O. W., Reyes-Betancort, A., Chapman, M. A. & Carine, M. A. (2018). Independent homoploid hybrid speciation events in the Macaronesian endemic genus *Argyranthemum*. *Molecular Ecology* 27: 4856—4874.

Wilson, G. A., and B. Rannala. Documentation for BayesAss 1.3. Retrieved from: http://rannala.org/docs/BayesAss.1.3.pdf. (accessed: 14 September 2018).

Wilson, G. A. & Rannala, B. (2003). Bayesian inference of recent migration rates using multilocus genotypes. *Genetics* 163: 1177—1191.

Yockteng, R., Ballard, H. E., Mansion, G., Dajoz, I. & Nadot, S. (2003). Relationships among pansies (*Viola* section *Melanium*) investigated using ITS and ISSR markers. *Plant Systematics and Evolution* 241: 153—170.

Zurbuchen, A., Landert, L., Klaiber, J., Müller, A., Hein, S., & Dorn, S. (2010). Maximum foraging ranges in solitary bees: only few individuals have the capability to cover long foraging distances. *Biological Conservation* 143: 669—676.

FIGURE AND TABLE LEGENDS

Table 1. Details of microsatellite primers used in the investigation.

Table 2. Genetic diversity statistics for *E. wildpretii*. Localities refer to sampling sites in Fig. 1. Diversity estimates are shown for each sampling locality, as well as mean values for each island and a total mean for all individuals where appropriate. Standard error values are shown in parentheses. N = average sample size; Na = average number of alleles per locus; PA = average number of private alleles per locus; I = Shannon's information index; Ho = observed heterozygosity; He = expected heterozygosity; F = fixation index.

Table 3. Pairwise genetic distances between sampling localities, calculated using FST. Locality names refer to those listed in Fig. 1. Genetic distances between localities on different islands are shown in bold.

Table 4. Estimated numbers of migrants exchanged per generation (Nm) between sampling localities within Tenerife and La Palma, calculated using FST. Locality names refer to those listed in Fig. 1.

Fig. 1. Location of Tenerife and La Palma (a) and sampling locations of *E. wildpretii* in the sub-alpine zones of Tenerife (b) and La Palma (c), showing images of subsp. *wildpretii* (d) and subsp. *trichosiphon* (e). In (b) and (c), locality codes used throughout are displayed above, with the number of individuals sampled shown in parentheses and coordinates shown below in decimal degrees.

Fig. 2. AMOVA analyses of three different groupings based on FST values

Fig. 3. Genetic variation and clustering in *E. wildpretii* based on (a) Principal coordinate analysis (PCoA), (b) Neighbour-Joining Network and (c) Structure. Sampling locality names

refer to those in Fig. 1. (a) PCoA based on genetic distance between individuals. Data points are displayed as coloured squares (Tenerife) or circles (La Palma), with the colours indicating the sampling locality of each individual. The percentage of variation explained by each coordinate is shown in parentheses. (b) Unrooted neighbour-joining network with branch lengths proportional to Nei's genetic distance. Branches are coloured according the legend in (a) to represent the sampling locality of individuals at the tips. (c) Structure analysis of E. wildpretii showing K=2 and K=3. Individuals are represented as vertical bars.

Fig. 4. Flower and leaf morphology of *Echium wildpretii*. (a) Flowers from Tenerife (b) Flowers from La Palma, (c) Mature rosette leaves from cultivated plants grown from seeds collected in Tenerife and La Palma. Each leaf is sampled from a different individual.

Table 1

Sequence ID[1]	Microsatellite Sequence	Fwd Primer Sequence	Rev Primer Sequence	Product size	N alleles
TR7148_c0_g1_i1	(TA)9	TGAGATTGTGACAAACAAACA	ACCATCATCATCATTCATTCA	125-137	5
TR17762_c0_g1_i1	(AGAA)3	AACAGGAGGTGGAAAACAG	TAGAGCATCAGCTTCCATATT	127-162	16
TR13997_c0_g1_i1	(ATC)8	GATCAAGCCAGATGTTGTCT	AAGGGTCTGTGTACCATGAG	150-162	5
TR20207_c3_g3_i1	(AGA)8	CCACACATTATTAGCAGTCCT	GAGATTTCGCTGACTTCATT	172-199	12
TR13020_c1_g2_i1	(CAC)8	AATAGAGATGAGCCCAATACA	ATGCTGTTTAAAGGGTTAAGG	202-225	5
TR19717_c0_g2_i1	(CTC)12	AACCAGACCAACAAGATGAC	CAGCAGGTGTGTTGGAAG	210-248	13
TR20168_c0_g1_i1	(AGA)8	AGCTGAAGAAGACGAAGAAGT	AAGATCCAAGCTACCCTCAC	258-270	6
TR20766_c0_g10_i1	(TGG)8	AATAACAGGTCCCTTCTTGAG	AACTGCATTGTAAATTCTGGA	270-285	6
TR15051_c0_g2_i1	(ACACCG)3	TTCCTCTGCCGCCCCTGCT	ACTCTTCTTATCAAACCACTCC	349-366	11

[1] From White et al. (2016)

Table 2

Locality	N	N_a	PA	I	H_o	H_e	F
T1	11.444 (0.242)	1.222 (0.147)	0.222 (0.147)	0.064 (0.042)	0.000 (0.000)	0.034 (0.022)	1.000 (0.000)
T2	17.556 (0.176)	2.556 (0.412)	0.667 (0.373)	0.443 (0.134)	0.101 (0.042)	0.243 (0.078)	0.550 (0.109)
T3	12.667 (0.441)	2.556 (0.338)	0.000 (0.000)	0.518 (0.123)	0.115 (0.068)	0.289 (0.073)	0.662 (0.127)
T4	10.111 (0.261)	2.667 (0.167)	0.333 (0.167)	0.602 (0.095)	0.176 (0.037)	0.353 (0.064)	0.410 (0.135)
subsp. *wildpretii* mean	12.944 (0.496)	2.250 (0.171)		0.406 (0.061)	0.098 (0.024)	0.230 (0.036)	0.567 (0.064)
LP1	8.444 (0.242)	3.444 (0.530)	0.222 (0.147)	0.891 (0.176)	0.262 (0.060)	0.476 (0.085)	0.388 (0.139)
LP2	9.444 (0.242)	3.778 (0.619)	0.333 (0.167)	1.002 (0.159)	0.402 (0.066)	0.539 (0.075)	0.248 (0.071)
LP3	8.667 (0.373)	3.444 (0.338)	0.333 (0.167)	0.839 (0.126)	0.407 (0.074)	0.459 (0.066)	0.089 (0.130)
LP4	11.667 (0.167)	3.444 (0.530)	0.222 (0.147)	0.808 (0.188)	0.386 (0.096)	0.424 (0.099)	0.075 (0.070)
subsp. *trichosiphon* mean	9.556 (0.250)	3.528 (0.247)		0.885 (0.079)	0.364 (0.037)	0.474 (0.040)	0.207 (0.056)
Total	11.250 (0.341)	2.889 (0.167)		0.646 (0.057)	0.231 (0.027)	0.352 (0.031)	0.360 (0.047)

Table 3

subsp. *wildpretii*				**subsp. *trichosiphon***				
T1	**T2**	**T3**	**T4**	**LP1**	**LP2**	**LP3**	**LP4**	
								T1
0.121								**T2**
0.109	0.037							**T3**
0.225	0.145	0.144						**T4**
0.450	**0.311**	**0.314**	**0.271**					**LP1**
0.475	**0.330**	**0.320**	**0.235**	0.082				**LP2**
0.479	**0.328**	**0.321**	**0.255**	0.053	0.046			**LP3**
0.514	**0.359**	**0.357**	**0.267**	0.064	0.068	0.049		**LP4**

Table 4

subsp. *wildpretii*					**subsp. *trichosiphon***			
	T1	T2	T3	T4		LP1	LP2	LP3
T1					LP1			
T2	0.756				LP2	2.952		
T3	1.283	2.554			LP3	5.884	6.796	
T4	0.360	0.778	1.196		LP4	2.846	3.663	4.936

Appendix S1: Details of locations sampled for population genetic analysis

Appendix S2: ~~**Additional Principal Coordinate Analysis (PCoA) plots**~~Per-locus per-taxon population genetic parameters

Appendix S3: Linkage disequilibrium amongst microsatellite loci

Appendix S4: Results of bottleneck testing under IAM and SMM models

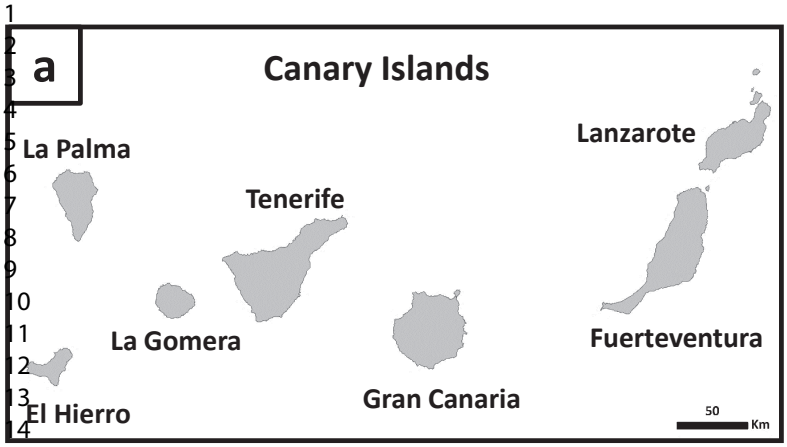
Appendix S5: Additional Principal Coordinate Analysis (PCoA) plots

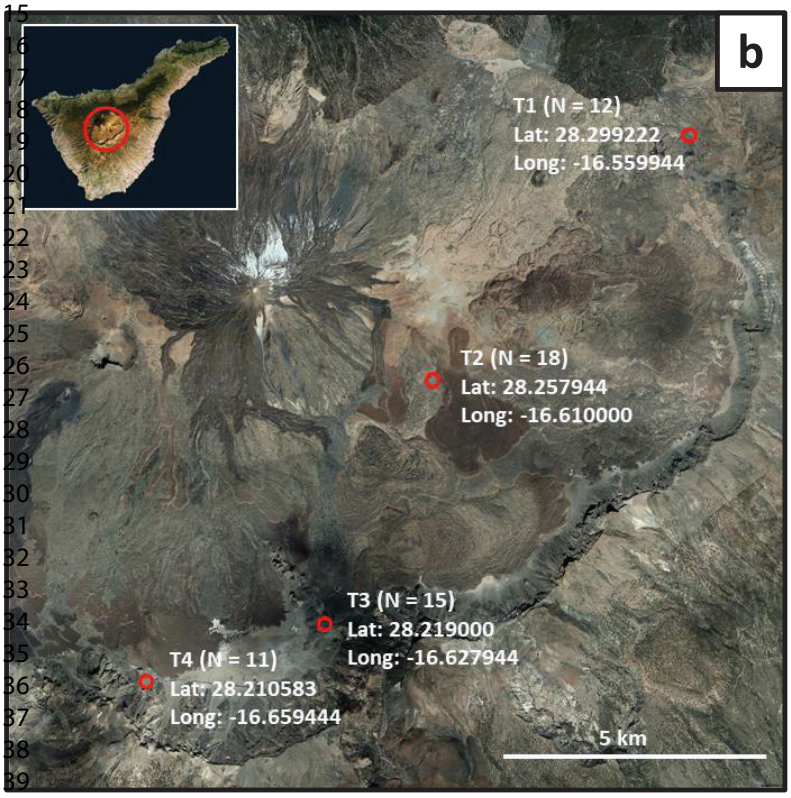
Appendix S6: Results of mclust model-based clustering analysis

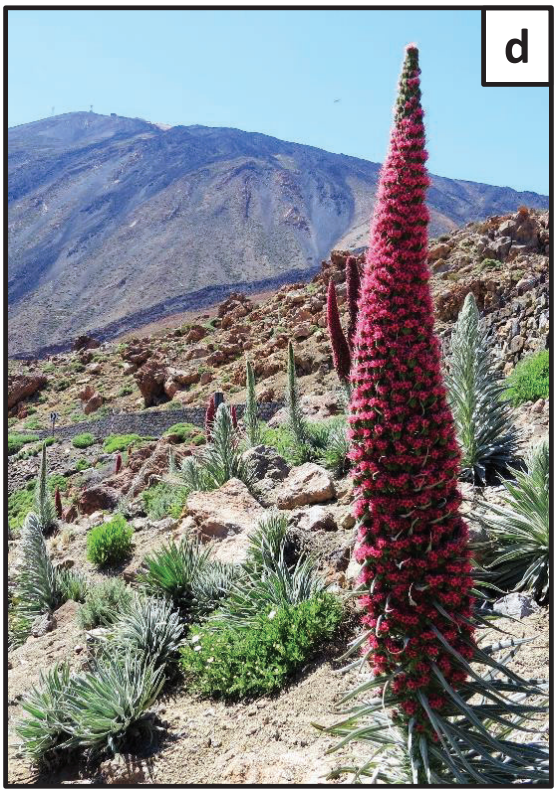
Appendix ~~**S4**~~S7: Evanno's DeltaK for Structure analysis of *E. wildpretii*

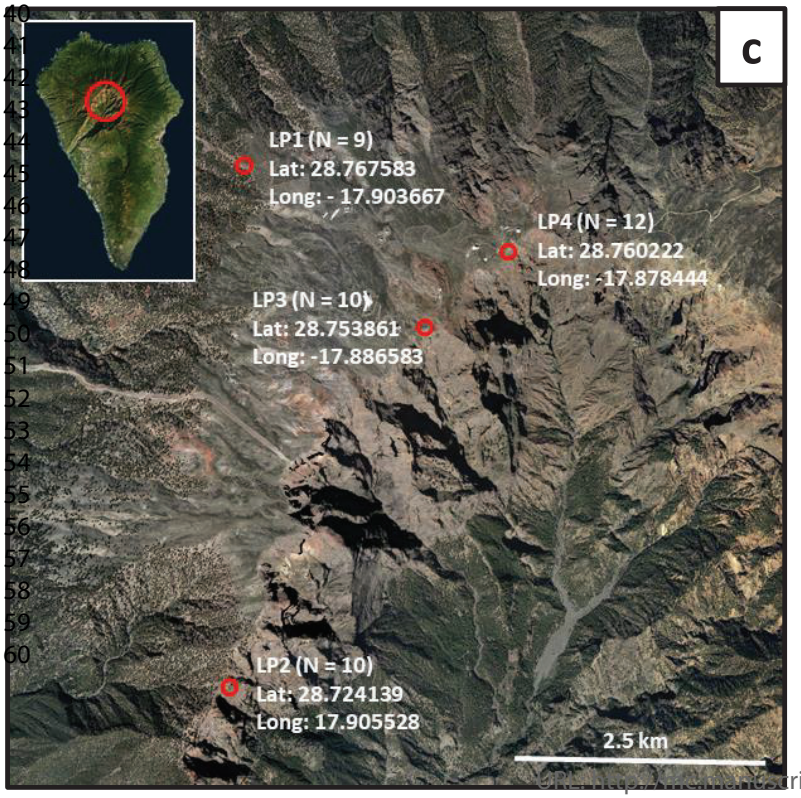
Appendix ~~**S5**~~S8: Structure analysis of *E. wildpretii* showing K=2-8

Appendix ~~**S6**~~S9: Measurements of stigma lobe length from herbarium specimens at BM


a
Canary Islands
La Palma
Lanzarote
Tenerife
La Gomera
Fuerteventura
El Hierro
Gran Canaria
50
Km


b
T1 (N = 12)
Lat: 28.299222
Long: -16.559944
T2 (N = 18)
Lat: 28.257944
Long: -16.610000
T3 (N = 15)
Lat: 28.219000
Long: -16.627944
T4 (N = 11)
Lat: 28.210583
Long: -16.659444
5 km


d

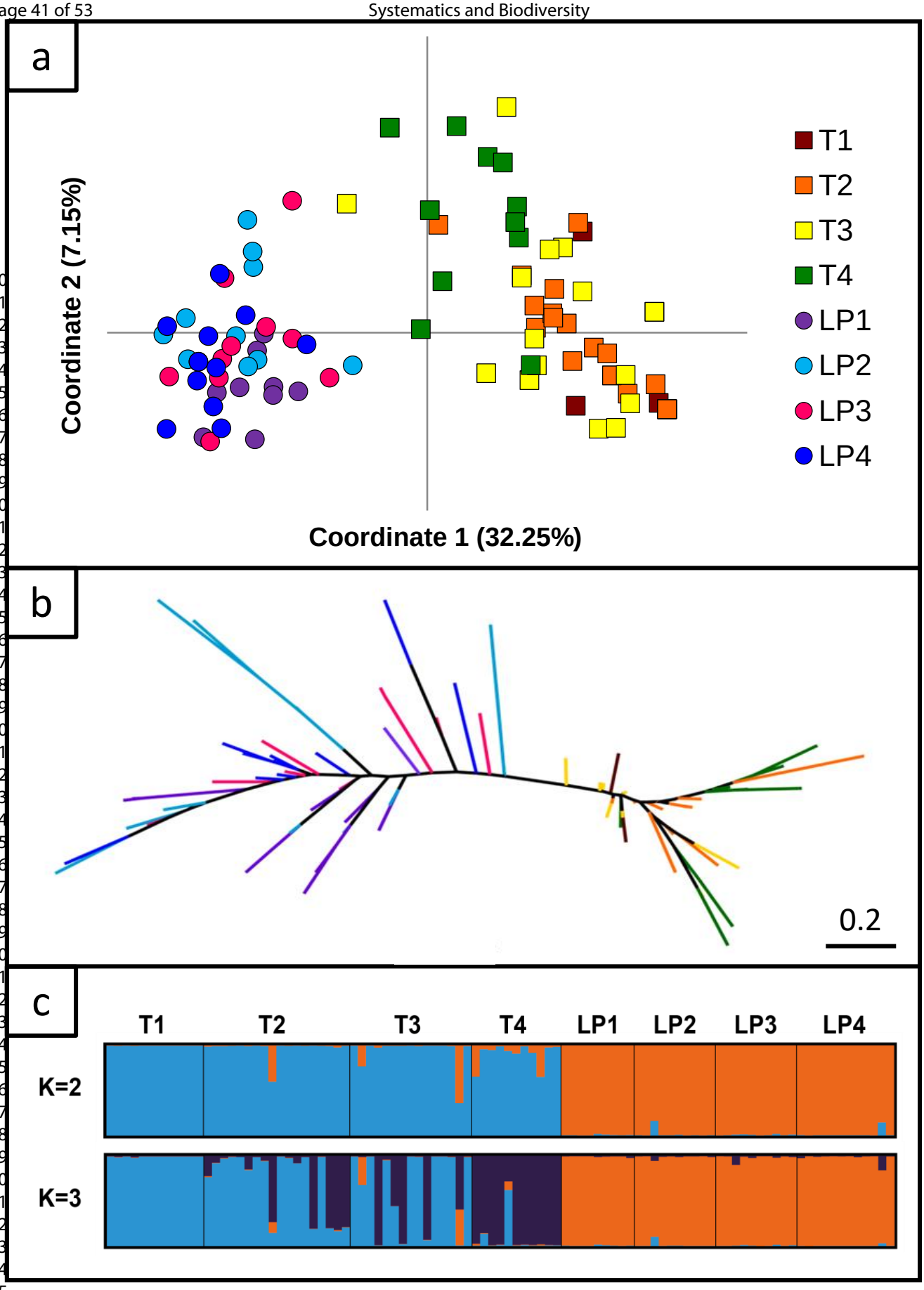

c
LP1 (N = 9)
Lat: 28.767583
Long: - 17.903667
LP4 (N = 12)
Lat: 28.760222
Long: -17.878444
LP3 (N = 10)
Lat: 28.753861
Long: -17.886583
LP2 (N = 10)
Lat: 28.724139
Long: 17.905528
2.5 km

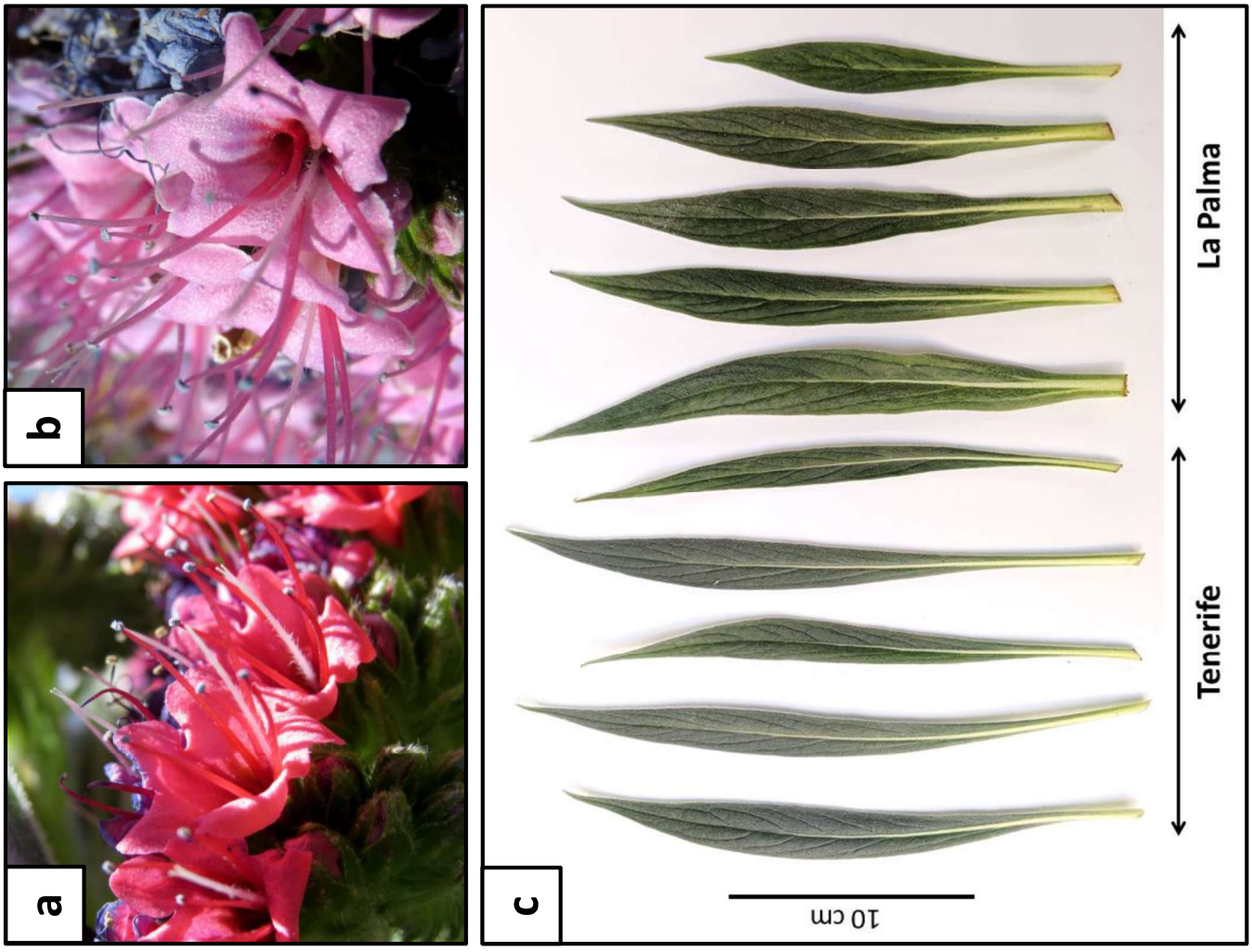

e

No. of groups	2
N	97
F_{ST}	0.381
P	0.001

No. of groups	4
N	56
F_{ST}	0.207
P	0.001

No. of groups	4
N	41
F_{ST}	0.058
P	0.001


a
Coordinate 2 (7.15%)
Coordinate 1 (32.25%)
T1
T2
T3
T4
LP1
LP2
LP3
LP4
b
0.2
c
T1
T2
T3
T4
LP1
LP2
LP3
LP4
K=2
K=3


b
a
c
La Palma
Tenerife
10 cm

Supplementary Material

Appendix S1: Details of locations sampled for population genetic analysis

For full details of voucher collections see the Taxonomic Treatment.

Island	Taxon	Locality	N	Latitude	Longitude	Elevation/ m a.s.l.	Voucher collection (herbarium)	Voucher barcode
Tenerife	*Echium wildpretii* subsp. *wildpretii*	T1	12	N 28˚ 17' 57.2"	W 16˚ 33' 35.8"	2127	Graham & Carine 52 (BM/ORT)	BM000828802
		T2	18	N 27˚ 14' 27.6"	W 15˚ 35' 35"	2325	Graham & Carine 56 (BM/ORT)	BM000828806
		T3	15	N 28˚ 13' 8.4"	W 16˚ 37' 40.6"	2111	Graham & Carine 58 (BM/ORT)	BM000828808
		T4	11	N 28˚ 12' 38.1"	W 16˚ 39' 34"	2035	Graham, White and Carine 104 (BM/ORT)	BM000828853
La Palma	*Echium wildpretii* subsp. *trichosiphon*	LP1	9	N 28˚ 46' 3.3"	W 17˚ 54' 13.2"	2017	Graham & White 138 (BM/ORT)	BM000828889
		LP2	10	N 28˚ 43' 26.9"	W 17˚ 54' 19.9"	1983	Graham & White 139 (BM/ORT)	BM000828890
		LP3	10	N 28˚ 45' 13.9"	W 17˚ 53' 11.7"	2395	Graham & White 146 (BM/ORT)	BM000828897
		LP4	12	N 28˚ 45' 36.8"	W 17˚ 52' 42.4"	2349	Graham & White 154 (BM/ORT)	BM000828909

Appendix S2: Per-locus per-taxon population genetic parameters

E. wildpretii

Locus	k	N	HObs	HExp	PIC	HW	F(Null)
7148	3	53	0.057	0.056	0.055	ND	0.0062
20207	4	53	0.075	0.211	0.201	ND	0.4557
20766	3	56	0.161	0.649	0.57	***	0.6019
13997	3	54	0.037	0.037	0.036	ND	0.0029
13020	2	53	0.019	0.157	0.143	ND	0.712
17762	6	51	0.059	0.533	0.491	***	0.7974
20168	5	48	0.083	0.326	0.301	ND	0.5851
19717	7	48	0.271	0.454	0.4	ND	0.249
15051	7	50	0.1	0.337	0.321	ND	0.5285
Average	4.44	51.78			0.28		0.44

E. perezii

Locus	k	N	HObs	HExp	PIC	HW	F(Null)
7148	3	38	0.158	0.316	0.282	ND	0.3213
20207	8	39	0.385	0.758	0.716	NS	0.3139
20766	4	40	0.075	0.166	0.159	ND	0.3608
13997	3	41	0.268	0.369	0.331	ND	0.1599
13020	3	37	0.270	0.289	0.266	ND	0.0755
17762	10	39	0.564	0.748	0.706	NS	0.1307
20168	5	40	0.525	0.651	0.595	NS	0.1014
19717	6	34	0.559	0.729	0.675	NS	0.1161
15051	7	36	0.528	0.705	0.644	NS	0.1264
Average	5.44	38.22			0.64		0.19

Appendix S3: Linkage disequilibrium amongst microsatellite loci

Statistically significant results (P < 0.05) are highlighted in bold.

E. wildpretii

Locus 1	Locus 2	P-Value	S.E.	Switches
7148	20207	1.000	0.000	2367
7148	20766	0.185	0.007	9367
20207	20766	**0.006**	0.002	8964
7148	13997	1.000	0.000	2211
20207	13997	1.000	0.000	2550
20766	13997	0.645	0.011	6475
7148	13020	1.000	0.000	4058
20207	13020	0.090	0.006	5386
20766	13020	**0.005**	0.001	13473
13997	13020	**0.044**	0.003	3631
7148	17762	0.224	0.012	4879
20207	17762	0.172	0.017	4283
20766	17762	**0.004**	0.001	11193
13997	17762	0.172	0.012	3235
13020	17762	**0.007**	0.001	9058
7148	20168	No contingency table		
20207	20168	0.058	0.010	3623
20766	20168	**0.006**	0.002	8719
13997	20168	0.125	0.011	2875
13020	20168	**0.005**	0.001	6342
17762	20168	**0.011**	0.003	4377
7148	19717	1.000	0.000	3383
20207	19717	0.088	0.010	5508
20766	19717	0.269	0.015	10802
13997	19717	0.063	0.004	4229
13020	19717	0.139	0.008	7041
17762	19717	0.532	0.024	5045
20168	19717	**0.037**	0.008	4870
7148	15051	**0.010**	0.001	4953
20207	15051	0.230	0.016	2801
20766	15051	0.155	0.017	5595
13997	15051	0.186	0.013	2177
13020	15051	0.488	0.018	3742
17762	15051	0.192	0.023	2712
20168	15051	**0.000**	0.000	7436
19717	15051	**0.031**	0.006	4499

E. perezii

Locus 1	Locus 2	P-Value	S.E.	Switches
7148	20207	0.682	0.023	4143
7148	20766	0.225	0.020	3080
20207	20766	0.508	0.032	1399
7148	13997	0.666	0.012	8146
20207	13997	**0.000**	0.000	5487
20766	13997	0.267	0.020	4035
7148	13020	0.643	0.013	11653
20207	13020	0.506	0.019	6846
20766	13020	0.280	0.016	5135
13997	13020	0.379	0.011	15110
7148	17762	0.328	0.025	4139
20207	17762	**0.030**	0.009	1911
20766	17762	0.121	0.018	1552
13997	17762	0.563	0.027	5379
13020	17762	0.710	0.018	6376
7148	20168	0.091	0.011	5114
20207	20168	0.123	0.021	2834
20766	20168	0.329	0.028	1978
13997	20168	0.313	0.020	7694
13020	20168	0.522	0.019	8832
17762	20168	0.098	0.020	2484
7148	19717	0.379	0.019	7439
20207	19717	0.168	0.024	2679
20766	19717	0.500	0.020	2852
13997	19717	0.213	0.017	6688
13020	19717	0.886	0.011	8057
17762	19717	0.769	0.028	1892
20168	19717	0.293	0.028	2916
7148	15051	0.752	0.019	5533
20207	15051	0.117	0.020	3370
20766	15051	0.596	0.022	2318
13997	15051	0.666	0.019	6890
13020	15051	0.263	0.014	8867
17762	15051	0.225	0.030	2379
20168	15051	0.092	0.015	3533
19717	15051	0.600	0.028	3081

Appendix S4: Results of bottleneck testing under IAM and SMM models

E. wildpretii

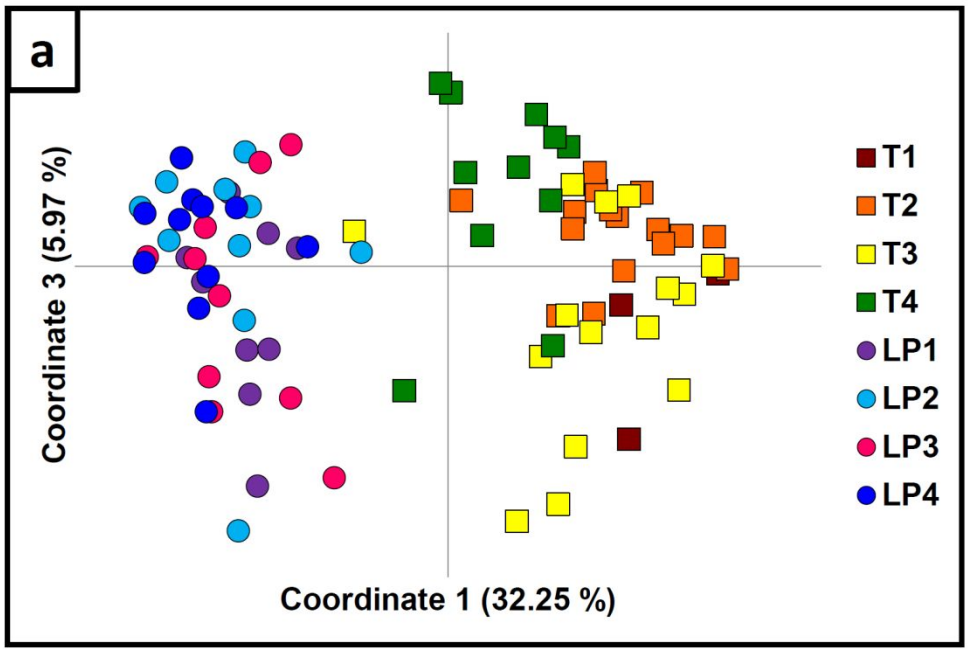
	Observed			Under the IAM				Under the SMM			
locus	n	ko	He	Heq	S.D.	DH/sd	Prob	Heq	S.D.	DH/sd	Prob
7148	106	3	0.056	0.315	0.186	-1.397	0.091	0.461	0.128	-3.161	0.002
20207	106	4	0.211	0.411	0.179	-1.116	0.194	0.582	0.111	-3.351	0.010
20766	112	3	0.649	0.320	0.190	1.734	0.014	0.445	0.146	1.403	0.029
13997	108	3	0.037	0.318	0.177	-1.585	0.046	0.454	0.138	-3.020	0.001
13020	106	2	0.157	0.182	0.167	-0.151	0.440	0.230	0.170	-0.429	0.437
17762	102	6	0.533	0.560	0.154	-0.179	0.362	0.727	0.070	-2.786	0.025
20168	96	5	0.326	0.511	0.161	-1.149	0.153	0.670	0.086	-3.976	0.003
19717	96	7	0.454	0.627	0.131	-1.321	0.124	0.771	0.053	-5.941	0.000
15051	100	7	0.337	0.625	0.132	-2.179	0.044	0.770	0.051	-8.456	0.000

E. perezii

	Observed			Under the IAM				Under the SMM			
locus	n	ko	He	Heq	S.D.	DH/sd	Prob	Heq	S.D.	DH/sd	Prob
7148	76	3	0.316	0.340	0.179	-0.132	0.453	0.466	0.138	-1.082	0.158
20207	78	8	0.758	0.687	0.111	0.636	0.323	0.805	0.045	-1.043	0.130
20766	80	4	0.166	0.437	0.172	-1.577	0.098	0.585	0.110	-3.797	0.006
13997	82	3	0.369	0.329	0.182	0.218	0.440	0.473	0.132	-0.794	0.196
13020	74	3	0.289	0.341	0.176	-0.295	0.415	0.473	0.135	-1.363	0.117
17762	78	10	0.748	0.755	0.085	-0.087	0.370	0.848	0.032	-3.156	0.016
20168	80	5	0.651	0.524	0.159	0.794	0.230	0.676	0.081	-0.309	0.304
19717	68	6	0.729	0.603	0.137	0.919	0.172	0.734	0.067	-0.077	0.369
15051	72	7	0.705	0.657	0.120	0.397	0.431	0.775	0.053	-1.324	0.090

Population	Test	IAM	SMM
Tenerife	Sign test: No. of loci with heterozygosity excess (probability)	1 (p = 0.011)	1 (p = 0.005)
	Wilcoxon test (Probability of heterozygosity excess)	0.981	0.998
La Palma	Sign test: No. of loci with heterozygosity excess (probability)	5 (p = 0.602)	0 (p = 0.000)
	Wilcoxon test (Probability of heterozygosity excess)	0.248	1

Appendix S5: Additional Principal Coordinate Analysis (PCoA) plots



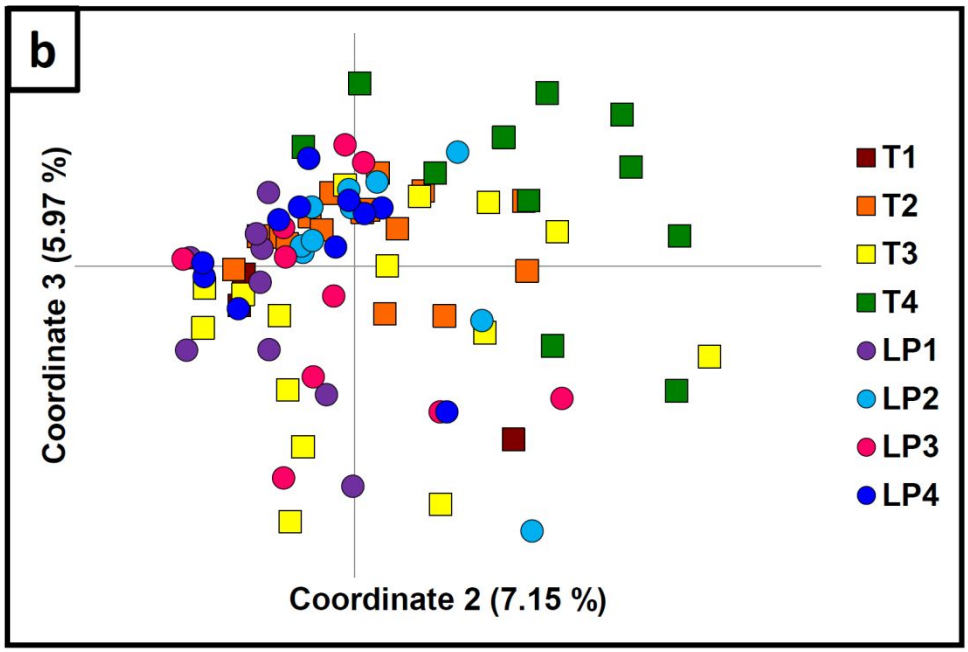


Figure shows a) PCoA plot of coordinates 1 and 3; b) PCoA plot of coordinates 2 and 3. Sampling locality names refer to those listed in **Error! Reference source not found.**.a. Data points are displayed as coloured squares (Tenerife) or circles (La Palma), with the colours indicating the sampling locality of each individual. The percentage of variation explained by each coordinate is shown in parentheses.

Appendix S6: Results of mclust model-based clustering analysis

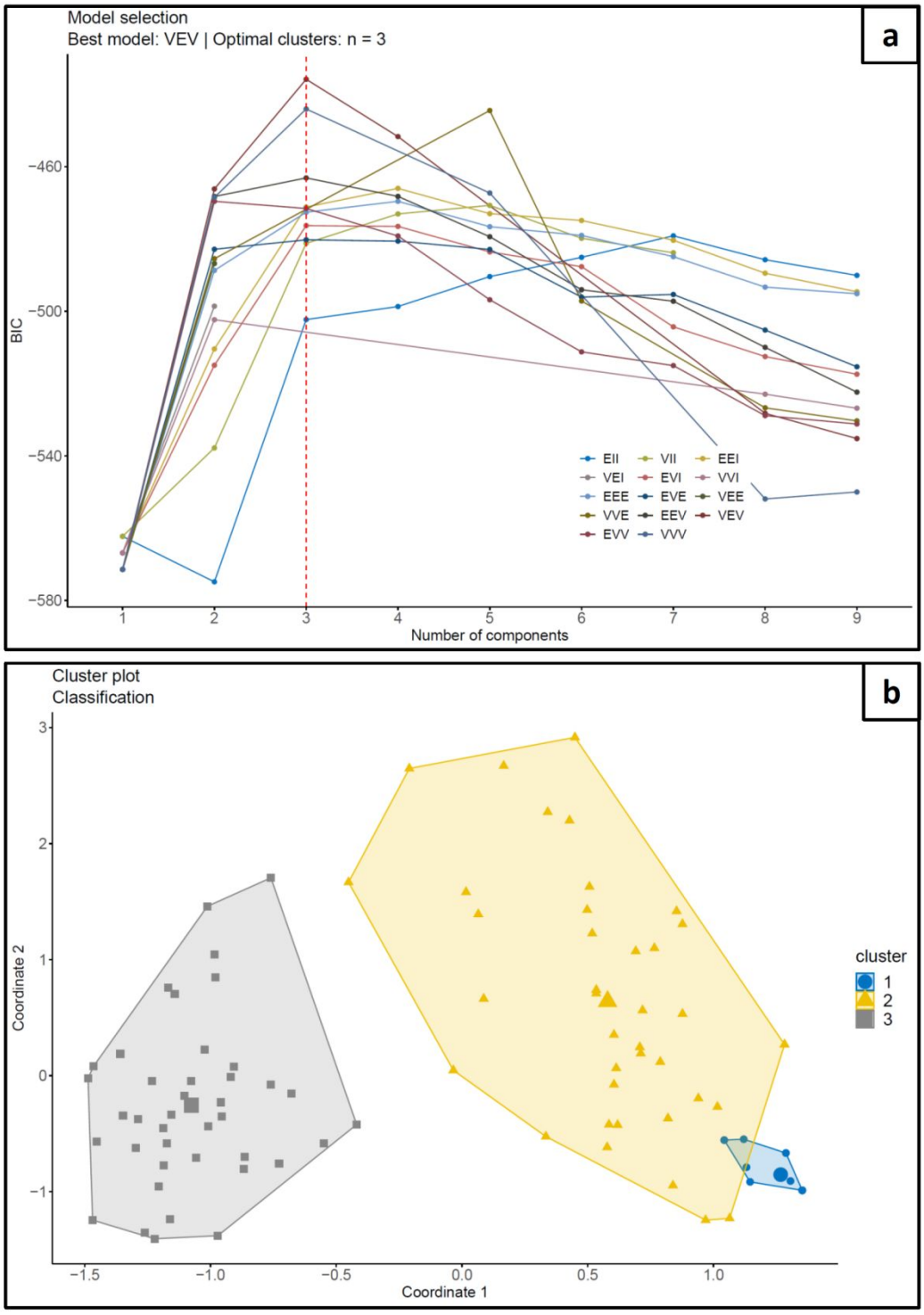
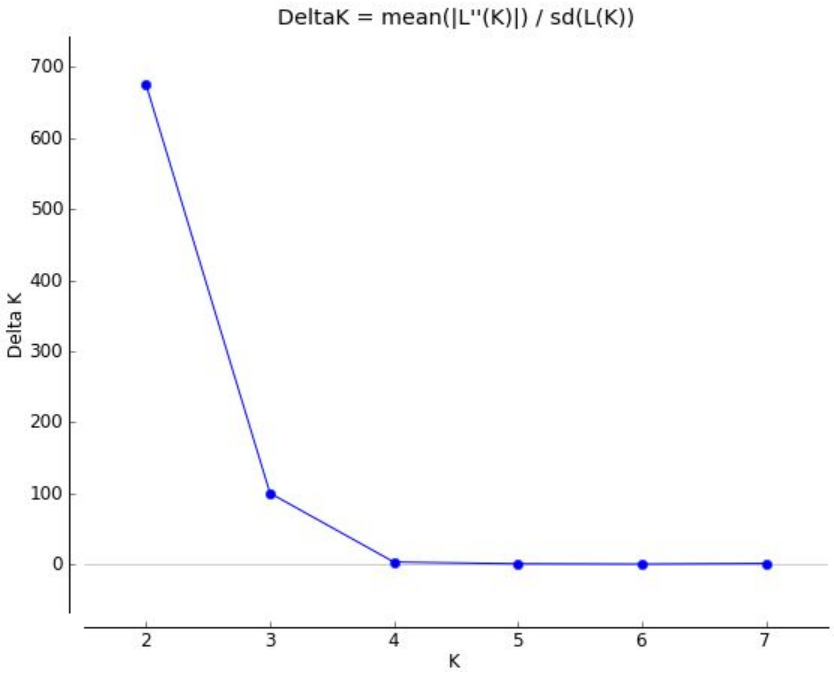
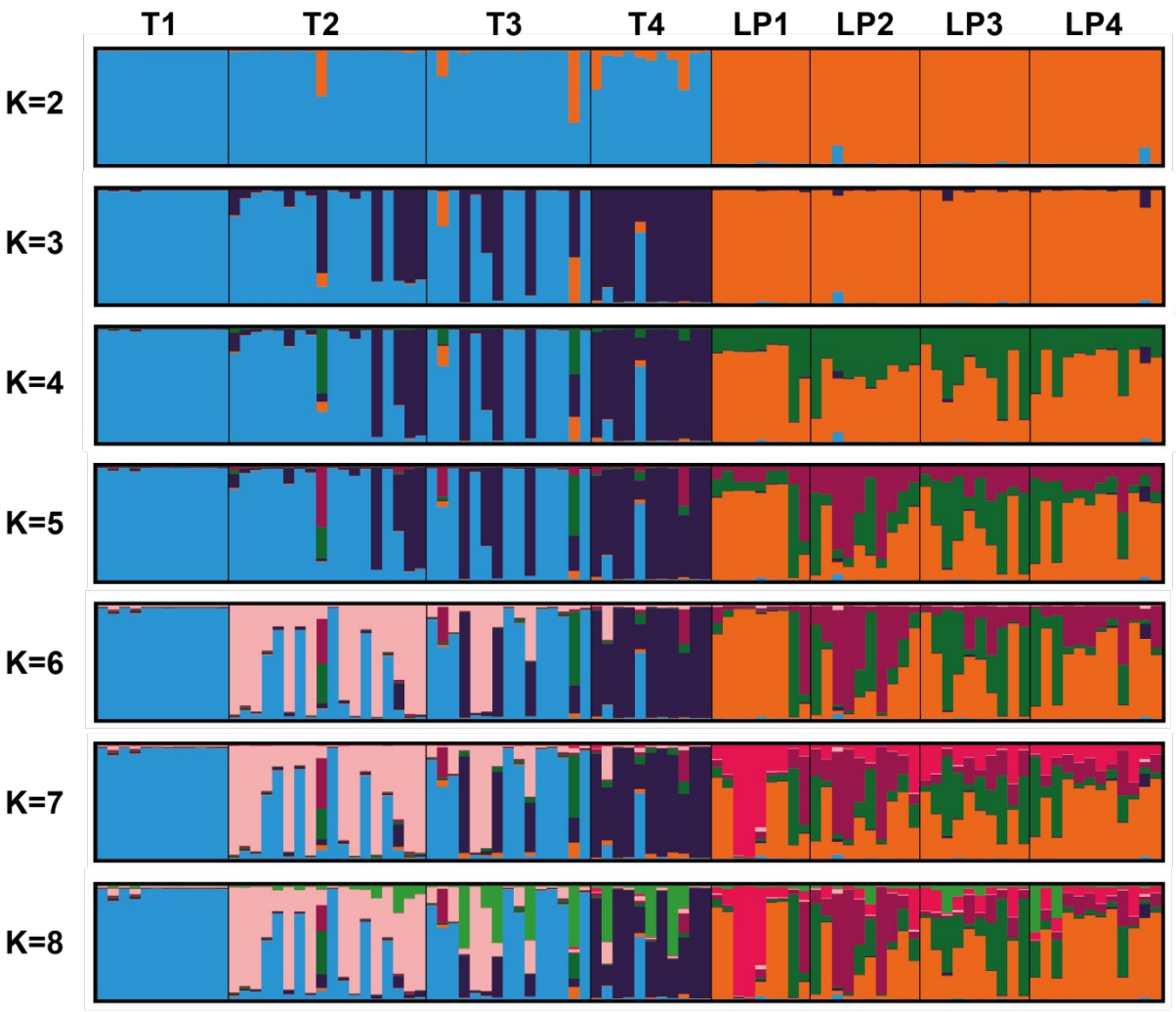


Figure shows a) the selection of the optimum model and number of clusters using the Bayesian Information Criterion (BIC); b) PCoA plot showing the classification of individuals into the three clusters identified by mclust.

Appendix S7: Evanno's DeltaK for Structure analysis of *E. wildpretii*



Appendix S8: Structure analysis of *E. wildpretii* showing *K*=2-8



Appendix S9: Measurements of stigma lobe length from herbarium specimens at BM

Taxon	Specimen	Length of stigmatic lobe (mm)			Mean length (mm ± standard deviation)	
E. wildpretii subsp. *wildpretii*	Graham & Carine 56 (BM)	0.84	0.88	0.62	0.78	0.70 (± 0.17)
	Graham, Carine & White 104 (BM)	0.40	0.62	0.38	0.47	
	Graham & Carine 52 (BM)	0.78	0.72	0.64	0.71	
	Graham, Carine 58 (BM)	0.80	0.72	0.96	0.83	
	Pérez, 275 (BM)	0.78	0.84	0.56	0.73	
E. wildpretii subsp. *trichosiphon*	Graham & White 138 (BM)	0.84	0.90	0.74	0.83	0.85 (± 0.21)
	Graham & White 154 (BM)	1.30	0.62	0.90	0.94	
	Graham & White 146 (BM)	0.58	0.94	0.86	0.79	