









Research Article

Tracing the biogeographic history of the world's most isolated insular floras

Ángela Aguado-Lara^{1,2*} , Isabel Sanmartín² , Johannes J. Le Roux³ , Carlos García-Verdugo⁴ , Sonia Molino^{1,5,6} , Peter Convey^{7,8,9,10,11} , Bettine Jansen van Vuuren⁷ , and Mario Mairal^{1*} 

¹Departamento de Biodiversidad, Ecología y Evolución, Universidad Complutense de Madrid, Madrid, Spain

²Real Jardín Botánico (RJB), CSIC, Madrid, Spain

³School of Natural Sciences, Macquarie University, Sydney, New South Wales, Australia

⁴Department of Botany, Universidad de Granada, Granada, Spain

⁵Department of Biology, Universidad Autónoma de Madrid, Madrid, Spain

⁶Department of Biosciences, Faculty of Biomedical and Health Sciences, Universidad Europea de Madrid, Madrid, Spain

⁷Department of Zoology, Centre for Ecological Genomics and Wildlife Conservation, University of Johannesburg, Auckland Park, South Africa

⁸British Antarctic Survey, NERC, Cambridge, United Kingdom

⁹Millennium Institute—Biodiversity of Antarctic and Sub-Antarctic Ecosystems (BASE), Santiago, Chile

¹⁰Cape Horn International Center (CHIC), Puerto Williams, Chile

¹¹School of Biosciences, University of Birmingham, Edgbaston, United Kingdom

*Authors for correspondence. Ángela Aguado-Lara. E-mail: aaguado@rjb.csic.es; anaguado@ucm.es; Mario Mairal. E-mail: mariomai@ucm.es; mariomairal@gmail.com

Received 5 August 2024; Accepted 9 January 2025

Abstract Inferring general biogeographic patterns in the sub-Antarctic region has been challenging due to the disparate geological origins of its islands and archipelagos—ranging from Gondwanan fragments to uplifted seafloor and more recently formed volcanic islands—and the remoteness of these island systems, spread around the austral continental landmasses. Here, we conduct phylogenetic reconstruction, divergence time estimation, and Bayesian Island Biogeographic analyses to reconstruct the spatio-temporal colonization histories of seven vascular plant lineages, which are either widespread across the sub-Antarctic region (*Acaena magellanica*, *Austroblechnum penna-marina*, *Azorella selago*, *Notogrammitis crassior*) or restricted to an extremely remote sub-Antarctic province (*Colobanthus kerguelensis*, *Polystichum marionense*, *Pringlea antiscorbutica*). Our results reveal high biological connectivity within the sub-Antarctic region, with southern landmasses (Australia, New Zealand, South America) as key sources of sub-Antarctic plant diversity since the Miocene, supporting long-distance dispersal as the primary colonization mechanism rather than tectonic vicariance. Despite the geographic isolation of the sub-Antarctic islands, eastward and westward colonization events have maintained this connectivity, likely facilitated by eastward-moving marine and wind currents, short-term weather systems, and/or dispersal by birds. Divergence time estimates indicate that most species diverged within the Plio-Pleistocene, with crown ages predating the Last Glacial Maximum, suggesting that sub-Antarctic archipelagos acted as refuges for biodiversity. Our findings highlight the role of one of the most remote sub-Antarctic archipelagos as both a refugium and a source of (re)colonization for continental regions. These results underscore the urgent need for establishing priority conservation plans in the sub-Antarctic, particularly in the face of climate change.

Key words: austral biogeography, island colonization, long-distance dispersal, phylogenetics, sub-Antarctic islands.

1 Introduction

The sub-Antarctic islands harbor a rich diversity of endemic species, which have been the subject of significant debate over decades (Hooker & Fitch, 1844; Convey, 2007). Central to this debate are disagreements concerning the roles of dispersal (time, direction, mode) versus geotectonic vicar-

iance in shaping the biogeography of the Southern Hemisphere (Moseley, 1879; Darlington, 1959; Muñoz et al., 2004; Sanmartín et al., 2007; Burridge et al., 2012; Winkworth et al., 2015; Crowl et al., 2016; Testo et al., 2018; Baker et al., 2020). Phylogenetic analyses provide a useful tool to address these knowledge gaps or to test biogeographic

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

hypotheses. However, clarification of biogeographic patterns in the Southern Hemisphere remains challenging due to the difficulties in achieving comprehensive taxon sampling across species distributed on disjunct tectonic elements, separated by large ocean basins, or in extremely remote and isolated archipelagos (Moon et al., 2017).

The sub-Antarctic islands are located amidst the primary continental plates of the Southern Hemisphere, positioned between the latitudes of 40°S and 54°S (Fig. 1). Their geographic isolation from major continental landmasses has made these islands some of the last pristine ecosystems remaining on the planet, although they now face significant threats stemming from biological invasions (Frenot et al., 2005; Le Roux et al., 2020; Mairal et al., 2022, 2023) and climate change (Smith, 2002; Nel et al., 2023). The location of these islands between the Subtropical Convergence and the Antarctic Polar Front (Fig. 1) places them in the direct trajectory of major wind (West Wind Drift [WWD]) and water (Antarctic Circumpolar Current [ACC]) currents. Exactly when the ACC and WWD established, with their prevailing west-to-east flow, is still debated; some authors suggest a late Oligocene origin (c. 24 million years (My); Pfuhl & McCave, 2005), while others suggest an earlier time frame (e.g., Sarkar et al., 2019), coincident with the opening of the Drake Passage and the Tasman Gateway during the Oligocene (28–34 My; Livermore et al., 2007; Sanmartín et al., 2007; Bahr et al., 2022). The ACC and the parallel atmospheric circulation of the WWD have been highlighted as drivers of unidirectional dispersal in plants and small animals (Breitwieser et al., 1999; Swenson et al., 2001; Muñoz et al., 2004; Sanmartín & Ronquist, 2004). Yet, several studies have suggested that these circumpolar currents do not impose absolute constraints in dispersal direction, and that westward migration is possible via alternative dispersal mechanisms, such as bird dispersal (Weimerskirch et al., 1985; Winkworth et al., 2002; Sanmartín et al., 2007; Deppe, 2012).

Another interesting characteristic of the sub-Antarctic islands is their widely varying geological ages. For example, South Georgia and the Falkland Islands/Islands Malvinas are “ancient” islands. These landmasses originated from fragments that broke off from the Gondwanan supercontinent. Geological reconstructions suggest that the isolation of South Georgia began around 30–35 My ago, as part of the breakup between the Antarctic Peninsula and southern South America, to which it was closely connected at the time, while the Falkland Islands were originally geologically associated with southern Africa (Ramos et al., 2017). In contrast, other sub-Antarctic islands of volcanic origin, such as Prince Edward Islands, were formed more recently (<1 My, Convey, 2007; Quilty, 2007). This long and complex geological history of the sub-Antarctic islands is hypothesized to have played a key role in maintaining biotic connectivity among the southern continental landmasses of Africa, Australia, New Zealand, and South America over time, likely through extreme long-distance dispersal (Moon et al., 2017; Smith et al., 2018).

The various geological ages of the sub-Antarctic islands, coupled with the presence of directional ocean and wind currents, have led to different hypotheses regarding the diversification and colonization histories of terrestrial organisms in these insular systems (Fig. 1B). The current practice in historical biogeography is to consider both tectonic vicariance and dispersal to explain patterns of extant diversity in the

Southern Hemisphere (Waters, 2008; Sanmartín, 2012; Lopes et al., 2024). In the case of the sub-Antarctic islands, some detached continental fragments, such as South Georgia (30–35 My), Falkland Islands (>120 My; Ramos et al., 2017), and the Kerguelen Archipelago (25–39 My), have existed for over tens of millions of years (Fig. 1A). It is thus possible that these “ancient islands” have harbored currently endemic sub-Antarctic lineages over millions of years. Some authors (Convey, 2007; Fraser et al., 2009, 2012; Convey et al., 2020) have hypothesized that these ancient islands could have acted as refugia for lineages that have remained isolated since at least the Plio-Pleistocene or even earlier (Fig. 1B). The “refugia hypothesis” (Table 1) posits that some sub-Antarctic islands, such as Gondwana-derived fragments and ancient archipelagos, acted as refugia for species or lineages that migrated to these regions before or during the Pleistocene glaciations and geographically persisted in ice-free areas on these islands during the Quaternary (Convey, 2007; Fraser et al., 2009, 2012; Convey et al., 2020). These taxa, or their ancestors, likely migrated from the nearby continental landmasses, for example, from southern South America to South Georgia or the Falkland Islands. These migration events did not necessarily involve large-scale long-distance dispersal; they could have occurred through stepping-stone dispersal or, if old enough, through land dispersal, followed by tectonic vicariance (Michaux & Leschen, 2005). The opposing “recent (re)colonization hypothesis” (Table 1) posits that most colonization events in the sub-Antarctic islands occurred in more recent times, after the Pleistocene glaciations (e.g., Mortimer et al., 2011; Fraser et al., 2012). Convey (2007) hypothesized that sub-Antarctic islands of recent origin, such as the volcanic Marion Island and the uplifted seafloor of Macquarie Island (both less than 1 My old), or those ancient islands significantly affected by the most recent glaciations, like Heard Island, likely acted as sinks for colonization events that postdate the Pleistocene glaciations. These hypotheses can be tested by inferring the age of the earliest colonization events that gave rise to the extant populations of each endemic sub-Antarctic species. The recent (re)colonization hypothesis would be supported if this age postdates the Pleistocene glaciations, whereas an inferred age that predates or falls within the Pleistocene glaciations would support the refugia hypothesis.

Regarding the directionality of dispersal events, the “anisotropic hypothesis” (Fig. 1B; Table 1) postulates that the direction of colonization events follows the prevailing west-to-east ocean and wind currents, the ACC and WWD. Supporting evidence for this hypothesis comes from biogeographic reconstructions and phylogenetic data (Breitwieser et al., 1999; Swenson et al., 2001; Muñoz et al., 2004; Sanmartín & Ronquist, 2004; Bergstrom et al., 2006; Sanmartín et al., 2007). Conversely, the “isotropic hypothesis” (Fig. 1B; Table 1) postulates that colonization may also occur against the prevailing currents, that is, east-to-west. Alternative dispersal mechanisms, such as zoochory, easterly anticyclones, or stepping-stone dispersal via island arc ridges or submerged plateaus (e.g., the Kerguelen or the Campbell Plateaus), might explain such isotropic dispersal events (Breitwieser et al., 1999; Swenson et al., 2001; Sanmartín & Ronquist, 2004; Moon et al., 2017). Distinguishing between these two hypotheses requires inferences of the timing and direction of dispersal events, for example, through biogeographic analysis. It is important to note the challenge of identifying and separating shorter countercurrent dispersal events from

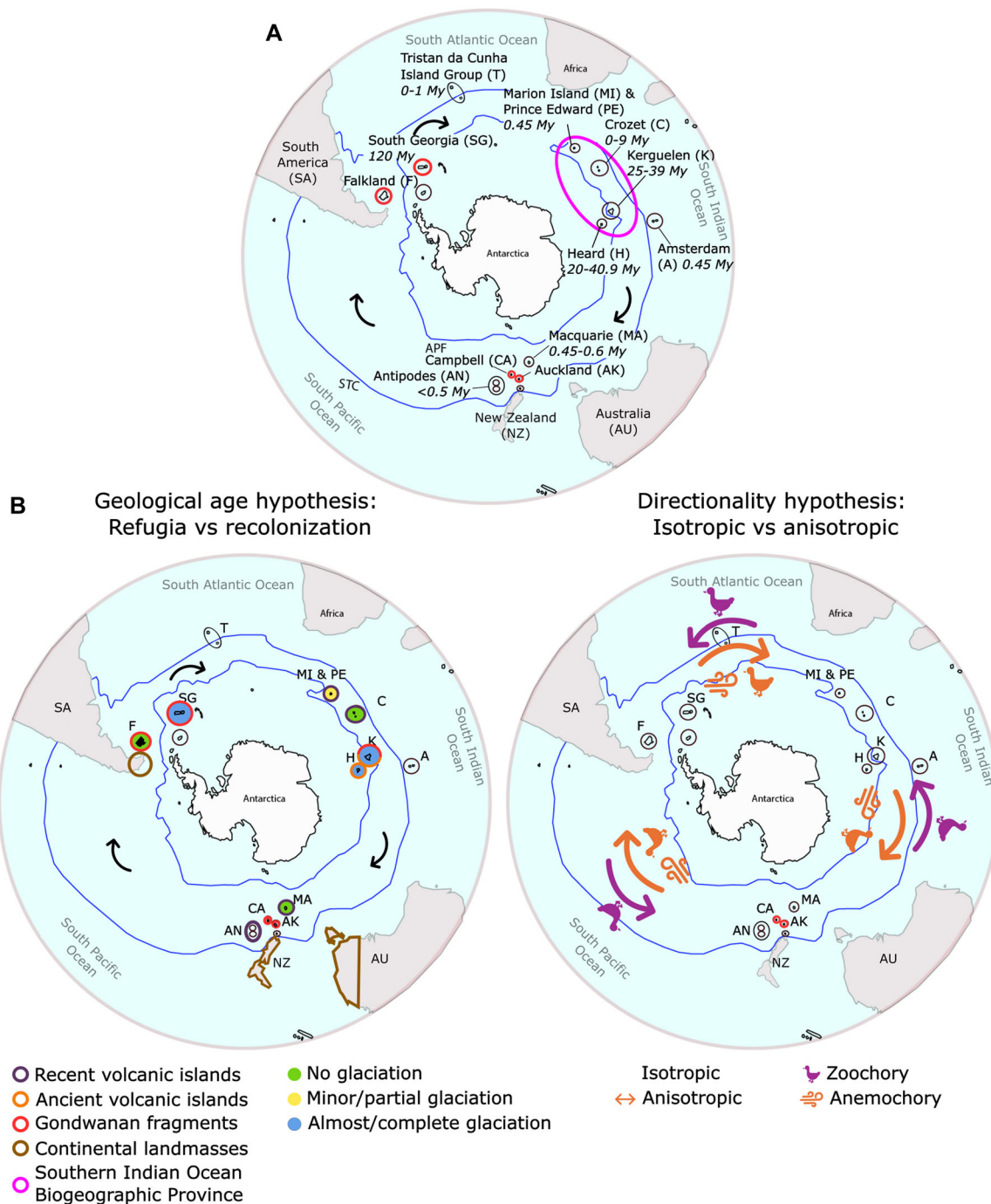


Fig. 1. Maps representing the Southern Hemisphere. **A**, Showing sub-Antarctic delimitation by the Sub-Tropical Convergence (STC) and the Antarctic Polar Front (APF). Gondwanan continental fragments are represented by red circles. The names and ages in million years (Hall, 2004; Convey, 2007; Quilty, 2007) of the principal sub-Antarctic islands and archipelagos are detailed. There are two types of volcanic islands: recent (0.45–9 My) or ancient (20–40 My). Black arrows represent the direction of the wind (West Wind Drift, WWD) and marine (Antarctic Circumpolar Current, ACC) currents that flow west to east around Antarctica. The pink ellipse delineates the Southern Indian Ocean Biogeographic Province (SIOBP). **B**, Outlining the main hypotheses to explain biotic assembly in the sub-Antarctic region, regarding (left) the geological age of the islands/continental fragments in relation to the age of colonization events (long-term refugia versus recent colonization) and (right) the direction of colonization events (isotropic versus anisotropic dispersal events). Orange and purple arrows represent anisotropic and isotropic dispersal directions, respectively. Abbreviations: A, Amsterdam Island; AN, Antipodes Island; AK, Auckland Island; AU, Australia; C, Crozet Islands; CA, Campbell Island; F, Falkland Islands; H, Heard Island; K, Kerguelen Archipelago; MA, Macquarie Island; MI, Marion Island; NZ, New Zealand; PE, Prince Edward Island; SA, South America; SG, South Georgia; T, Tristan da Cunha group.

Table 1 Summary of the four main hypotheses underlying the origin and direction of colonization of the sub-Antarctic

Hypotheses	References	Predictions	Support
Geological age			
Refugia	Convey, 2007; Fraser et al., 2009, 2012; Convey et al., 2020	The sub-Antarctic islands acted as refugia for species that migrated to these regions before or during the Pleistocene glaciations, and geographically persisted in ice-free areas on these islands during the Quaternary.	The age of the first colonization event by the endemic sub-Antarctic species falls within or predates the onset of the Pleistocene glaciations.
Recent (Re) Colonization	Convey, 2007; Mortimer et al., 2011; Fraser et al., 2012; Mairal et al., 2023	The sub-Antarctic islands acted as sinks of recent colonization events that took place after the Pleistocene glaciations.	The age of the first colonization event by the endemic sub-Antarctic species postdates the last Pleistocene glaciation.
Directionality			
Anisotropic	Breitwieser et al., 1999; Swenson et al., 2001; Muñoz et al., 2004; Sanmartín & Ronquist, 2004; Moon et al., 2017	The direction of colonization events follows the west-to-east prevailing ocean and wind currents.	Phylogenetic data and circumstantial evidence support west-to-east colonization events by means of the long-persistent ACC and WWD currents.
Isotropic	Weimerskirch et al., 1985; Winkworth et al., 2002; Sanmartín et al., 2007	The direction of colonization events was predominantly against the west-to-east prevailing ocean and wind currents.	Other dispersal mechanisms, such as zoochory, easterly anticyclones, or stepping-stone dispersal via island ridges or submerged plateaus, explain these dispersal events.

long-distance dispersal events that may circulate nearly or entirely around the globe.

In order to test the hypotheses outlined above, multi-clade studies of a wide range of organisms are needed, including organisms with different temporal origins and dispersal strategies. This approach will allow for the distinction of general biogeographic patterns from organism-specific ones (Sanmartín et al., 2008; Whittaker et al., 2017). The intrinsic features of the sub-Antarctic islands, such as their small sizes and geographic isolation, pose significant challenges for conducting such studies aimed at disentangling their biogeographic histories, and by extension, austral biogeography. In this study, we selected seven vascular plant species (Fig. 2) representing various evolutionary lineages (fern and angiosperm species from different families) and showing varying distributional ranges and dispersal mechanisms (Table S1). Species with wide distributions across the southern latitudes, including those found on both sub-Antarctic islands and nearby continental landmasses (hereafter referred to as “widespread species”), may offer valuable insights into the historical connections among Southern Hemisphere landmasses and the processes that have influenced them (Sanmartín et al., 2007). Conversely, species with restricted geographic distributions—that is, those endemic to one or a few islands or archipelagos (hereafter referred to as “restricted species”), can help us to explore which hypothesized dispersal barriers have been key in shaping

geographic patterns within the region (Sanmartín et al., 2008; Moon et al., 2017). Taxon sampling in our study encompassed four widespread species (Figs. 2A–2D; Table S1): the angiosperm (Eudicotyledon) species *Acaena magellanica* (Lam.) Vahl (Rosaceae) and *Azorella selago* Hook.f. (Apiaceae), and the fern species *Austroblechnum pennamaryna* (Poir.) Gasper & V.A.O. Dittrich (Blechnaceae) and *Notogrammitis crassior* (Kirk) Parris (Polypodiaceae), as well as three restricted species (Fig. 2E; Table S1): the eudicots *Colobanthus kerguelensis* Hook.f. (Caryophyllaceae) and *Pringlea antiscorbutica* R. Br. ex Hook.f. (Brassicaceae), and the fern *Polystichum marionense* Alston & Schelpe (Dryopteridaceae). The restricted species are distributed exclusively in the four archipelagos of the “Southern Indian Ocean Biogeographic Province” (SIOBP, pink ellipse in Fig. 1A): the Prince Edward Archipelago (PEA, Prince Edward and Marion Islands), Crozet Islands, the Kerguelen Archipelago, and Heard Island (which also includes Macdonald Island). The SIOBP stands out as one of the most remote insular regions globally, with limited natural colonization opportunities from continental sources (Fig. 1A). It is located midway between the southern continental landmasses of Australia–New Zealand and South America and is thus subject to the influence of the ACC and WWD currents. For instance, only a few vascular plant species have managed to colonize and establish on the PEA, which support just 21 native vascular plant species (Chau et al., 2020). In addition

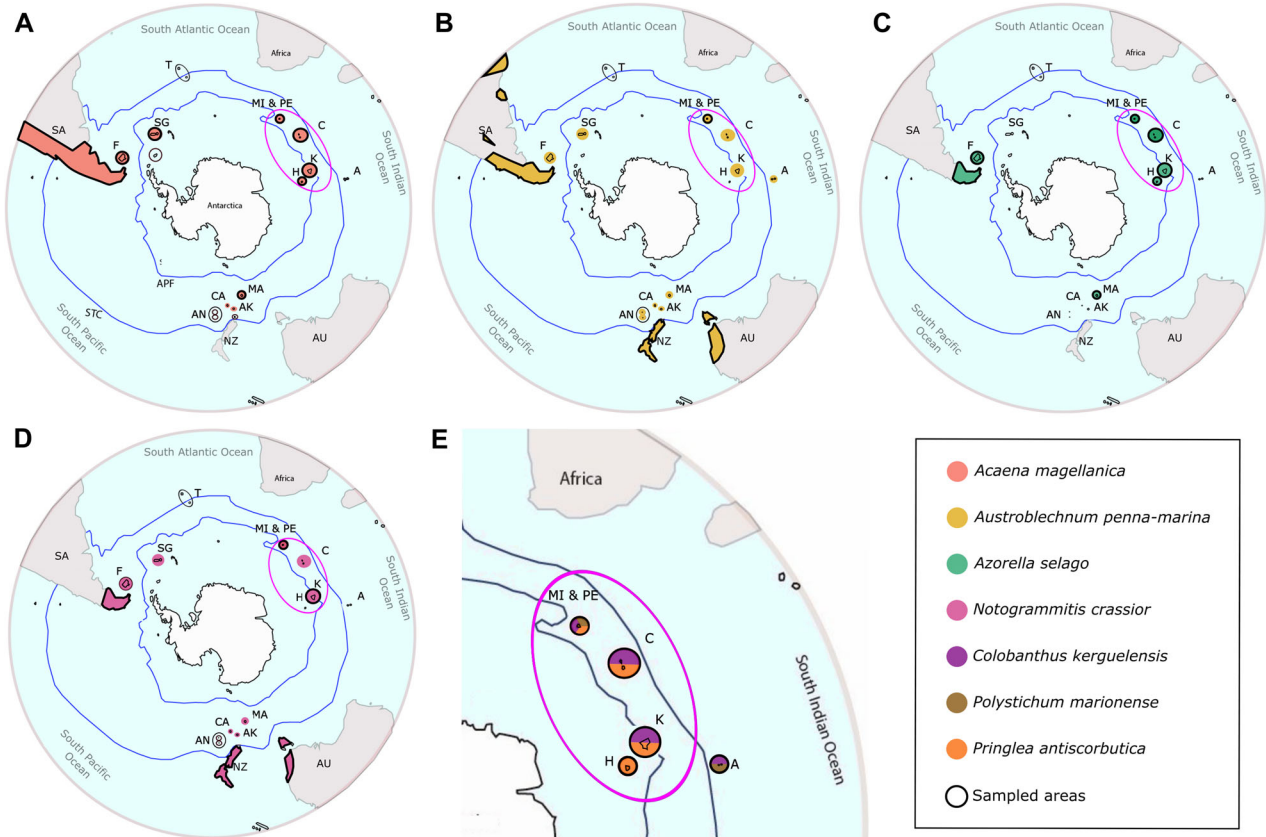


Fig. 2. Maps representing the distribution of the sub-Antarctic species that were included in this study. Widespread species: **A**, *Acaena magellanica* (Lam.) Vahl. **B**, *Austroblechnum penna-marina* (Poir.) Gasper & V.A.O. Dittrich. **C**, *Azorella selago* Hook.f. **D**, *Notogrammitis crassior* (Kirk) Parris. Restricted species to the Southern Indian Ocean Biogeographic Province (SIOBP): **E**, *Colobanthus kerguelensis* Hook.f.; *Pringlea antiscorbutica* R. Br. ex Hook.f.; and *Polystichum marionense* Alston & Schelpe. Outlined areas in black represent sampled locations. Abbreviations: A, Amsterdam Island; AN, Antipodes Island; AK, Auckland Island; AU, Australia; CA, Campbell Island; C, Crozet Islands; F, Falkland Islands; H, Heard Island; K, Kerguelen Archipelago; MA, Macquarie Island; MI, Marion Island; NZ, New Zealand; PE, Prince Edward Island; SA, South America; SG, South Georgia; T, Tristan da Cunha Group.

to variation in their distributions, the plant species that we selected also represent a diverse array of dispersal strategies, including anemochory, hydrochory, and zoochory (Table S1). Another criterion for species selection was the availability of molecular data; the seven selected species and/or their closest relatives are well represented in online repositories (e.g., GenBank; Appendix S1). Some of these species have been subject to recent phylogeographic studies (Chau et al., 2020).

Using our taxonomically and geographically diverse species data set, our aims were to (i) reconstruct the biogeographic history of the seven sub-Antarctic plant lineages; (ii) infer general patterns of vascular plant distribution across the sub-Antarctic region; and (iii) explore possible explanatory processes, including ancient vicariance, geographic persistence, and dispersal across barriers and, for the latter, the timing and direction of dispersal events. Additionally, we explored whether differences in biogeographic patterns exist between widespread and geographically restricted species. Our overarching objective was to examine phylogenetic support (both temporal and spatial) for alternative hypotheses that have been postulated

to explain floristic assembly in the sub-Antarctic region (Table 1; Fig. 1B).

2 Material and Methods

2.1 Taxon sampling and DNA sequencing

Access to the sub-Antarctic archipelagos is restricted and only possible with the support of national Antarctic programs, international consortia, and governments with sovereign authority. Due to these constraints, we obtained DNA sequencing data for our selected plant species from various sources. First, plant tissue samples were collected during field expeditions conducted by some of the authors of this study and their collaborators between 2008 and 2018. This included material obtained from expeditions to South Georgia, Kerguelen, Crozet, and Macquarie, and annual expeditions to the PEA (e.g., Chau et al., 2020; Mairal et al., 2023).

This study focuses on seven target species (Table S1): *Acaena magellanica*, *Austroblechnum penna-marina*, *Azorella selago*, *Colobanthus kerguelensis*, *Notogrammitis crassior*,

Polystichum marionense, and *Pringlea antiscorbutica*. Nomenclature follows POWO (Plants of the World Online: <https://powo.science.kew.org/>) for angiosperms and PPG-I (2016) for ferns. Other species were initially considered, including the monocotyledonous species *Carex dikei* (Nelmes) K.L. Wilson, *Juncus scheuchzerioides* Gaudich., *Poa cookii* (Hook.f.) Hook.f., and *Polypogon magellanicus* (Lam.) Finot. However, the primary reasons for not including these species in our analyses are that they did not have an extended taxon sampling (few available individuals) and/or because of the low clade-level resolution observed in available phylogenetic analyses, which often resulted in polytomies with sister species (e.g., *Carex dikei* and other related species from North America, García-Moro et al., 2021).

For the seven target species for which we collected material, DNA extractions were conducted from leaf tissue of silica-dried samples using the NORGEN Biotech kit (Canada) following the manufacturer's protocol. The TissueLyser LT instrument was used to finely mill tissue samples. The DNA concentration and purity of all extractions were assessed using a Nanodrop (Thermo Scientific; Table S2). We generated DNA sequences for one nuclear region, the Internal Transcribed Spacers ITS1 and ITS2, and four plastid regions, maturase K (*matK*), ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (*rbcl*), ribosomal protein S4 gene (*rps4*) and tRNA-Ser intragenic space (*trnS-rps4*), and tRNA-Leu and tRNA-Phe intragenic space (*trnL-trnF*) gene. These gene regions were chosen based on their use in published studies on our target species [Zhang et al. (2017) and Jauregui-Lazo & Potter (2021) for *A. magellanica*; Gasper et al. (2017) for *A. penna-marina*; Andersson et al. (2006) and Nicolas & Plunkett (2012) for *A. selago*; Fior et al. (2006) and Biersma et al. (2020) for *C. kerguelensis*; Ranker et al. (2004) for *N. crassior*; and Bartish et al. (2012) for *P. antiscorbutica*] and their routine use in plant systematics; for example, *rps4-trnS* and *trnL-trnF* are frequently used in studies on ferns (Trewick et al., 2002; Shaw et al., 2005). All PCR reactions were performed using commercial semi-automated kits from nzytech®, amplified using a thermal cycler Sensequest (Labcycler, Portugal) and purified with the QIAquick PCR Purification kit (Macherey-Nagel, Germany). Conditions for the PCR protocol for each gene region are provided in Tables S3, S4. Sequencing was carried out by Eurofins Genomics. We verified the quality of the electropherograms of each sequence using Geneious 11.0.4 software (www.geneious.com/). Individual reads (forward and reverse sequences) were aligned using the software MEGA X (Kumar et al., 2018) to generate a consensus sequence per sample. In total, we generated 141 consensus sequences: 52 for *A. magellanica* (26 for ITS1-ITS2 and 26 for *matK*), 57 for *A. selago* (28 for *matK* and 29 for *rbcl*), seven for *N. crassior* (five for *trnL-trnF* and two for *trnS-rps4*), 12 for *P. marionense* (five for *trnL-trnF* and seven for *trnS-rps4*), and 13 for *P. antiscorbutica* (six for ITS1-ITS2 and seven for *trnL-trnF*).

We supplemented our genetic data through a comprehensive search for available DNA sequencing data in the GenBank online repository (<https://www.ncbi.nlm.nih.gov/genbank/>), resulting in 48 additional sequences: eight for *A. penna-marina* (five for *rbcl* and three for *trnS-rps4*), eight for *C. kerguelensis* (four for ITS1-ITS2 and four for *atpB-rbcl*), 21 for *N. crassior* (eight for *rbcl*, four for *trnL-trnF*, three for

trnS-rps4, and six for *atpB*), and 11 for *P. antiscorbutica* (four for ITS1-ITS2 and seven for *trnL-trnF*). The final database was representative of the distributional ranges of the seven species: for the widespread taxa *A. magellanica*, *A. penna-marina*, *A. selago*, and *N. crassior* (Figs. 2A–2D), our samples covered both the sub-Antarctic islands and the adjacent continental landmasses; for the restricted species *C. kerguelensis*, *P. marionense*, and *P. antiscorbutica* (Fig. 2E), our sampling spanned their distribution across the SIOBP islands, with some exceptions (Fig. 2E).

For outgroup taxa, we selected species that have been inferred to be closely related to our seven target species in recent molecular phylogenetic studies. We conducted an extensive literature search to find the most comprehensive phylogenetic hypotheses that included our study species. Appendix S1 provides more details on outgroup species, collectors' names, GenBank accession numbers (when applicable), and geo-referenced coordinates of the samples (when available). In the case of *P. marionense*, we used a different approach. Since this species had never been included in a published phylogenetic study, we conducted a BLAST search to identify sequences of *Polystichum* species that were most closely related to those that we generated for *P. marionense*. We downloaded 12 DNA sequences from GenBank generated in a recent study of phylogenetic relationships and divergence times in the genus *Polystichum* (Le Péchon et al., 2016) and aligned them with our sequences of *P. marionense*. This allowed us to generate a preliminary phylogenetic hypothesis to ascertain the phylogenetic placement of this species within *Polystichum* and also to provide calibration points for phylogenetic dating (see below).

In total, our data set contained 48 GenBank sequences for the target species and 82 sequences for the outgroup taxa (Appendix S1); these published sequences were combined with the 141 sequences generated in this study.

2.2 Phylogenetic inference and divergence time estimation

DNA sequences were aligned separately for each gene region in MEGA X (www.ebi.ac.uk/tools/msa/muscle) using the multilocus alignment method MUSCLE. Phylogenetic relationships were inferred using Bayesian Inference (BI) implemented in MrBayes v. 3.2.7 (Ronquist et al., 2012). We used jModelTest v2.10 (Darriba et al., 2012) to select the best-fitting substitution model for each gene region for each target species, using the Akaike Information Criterion (AIC). For some gene regions, the selected models were not available for implementation in MrBayes (e.g., TVM, TPM1uf, TPM3uf, TrNef, TrN, TPM, TIM1). In these cases, we used the next most complex model available (i.e., GTR). We first reconstructed a phylogenetic tree for each individual gene region and target species to check for topological congruence and to identify statistically supported relationships by examining the Bayesian consensus trees. Clade posterior probability values were considered to provide “high” nodal support when ≥ 0.95 (Alfaro et al., 2003). No incongruent clades with high support were found among the individual gene trees. Therefore, we concatenated all markers into a combined nuclear/plastid data set for each target species and their associated outgroup taxa separately, which was used for further analyses (Appendix S1; Table S5).

For some species, we could only successfully sequence plastid gene regions: *A. penna-marina* (*rbcl* and *rps4-trnS*), *A. selago* (*matK* and *rbcl*), *N. crassior* (*atpB*, *matK*, *rbcl* and *trnL-trnF*), and *P. marionense* (*trnL-trnF* and *trnS-rps4*). For *A. magellanica*, *C. kerguelensis*, and *P. antiscorbutica*, for which we were able to obtain both nuclear and plastid gene regions, we partitioned the data set by these two gene regions prior to analyses. We ran MrBayes with four coupled Markov Chains, one heated and three cold, using Metropolis Hastings MCMC simulations, with one million generations per chain; burn-in was set to 25% of the posterior probability sample. Parameter convergence was assessed using the Potential Scale Reduction Factor (PSRF approximates 1.0) and the effective sampling size for each parameter (ESS values > 200). The post-burn-in trees were summarized into a 50% majority rule consensus tree with clade posterior probabilities to approximate the posterior distribution of the phylogeny, using the *sumt* command in MrBayes. The resulting tree was visualized using FigTree v.1.4.4. (<http://tree.bio.ed.ac.uk/software/figtree/>).

Divergence time estimates between lineages were inferred using the Bayesian molecular clock methods implemented in BEAST v.1.10.4 (Suchard et al., 2018). The tree and clock prior were selected using Bayes Factor comparisons of the model marginal likelihood; to compute the marginal likelihood values, we used stepping-stone and path-sampling power posteriors (Baele et al., 2013) implemented in BEAST (Table S6). We compared a model with incomplete taxon sampling (Gernhard, 2008) versus a coalescent model (Kingman, 1982), and a strict clock model versus an uncorrelated log-normal relaxed clock model (UCLD; Drummond et al., 2006). Given that no dated fossils are available for most of our target species, we used secondary age calibrations obtained from the published phylogenies reviewed above (Table S7). The only exception was *A. magellanica*, for which we used two fossils (Jauregui-Lazo & Potter, 2021): a pollen fossil assigned to the genus (5.0 My; Stokes, 1988), which was used to calibrate the MRCA (most recent common ancestor) of all individuals (i.e., the *Acaena* crown-age), and a fossil attributed to *Rosa germerensis* (47.8 My; Edelman, 1975; Zhang et al., 2017), which was used to calibrate the *Rosa-Potentilla* clade in our analysis. For other parameters, we implemented distribution priors that are commonly used in plant systematic studies. Specifically, we applied a uniform prior from 10^{-6} to 10^{-1} per site My^{-1} for the parameter *ucl.d.mean* (Wolfe et al., 1987), and assigned the GTR + G model (Posada, 2009) as the molecular evolutionary model for the concatenated matrix. We ran an MCMC analysis with 10^8 generations, sampling every 1000th generation. We used Tracer v.1.7.2 (Rambaut et al., 2018) to monitor adequate mixing and convergence to the stationary distribution (i.e., ESS values > 200 for each parameter). A Maximum Clade Credibility (MCC) tree was constructed using TreeAnnotator software (within the BEAST software), after discarding 25% of trees as burn-in. The MCC tree was visualized in FigTree v.1.4.4 (<http://tree.bio.ed.ac.uk/software/figtree/>).

2.3 Inferring general dispersal patterns using Bayesian hierarchical models

We used the Bayesian Island Biogeographic (BIB) model described by Sanmartín et al. (2008) to infer general patterns

of dispersal among the sub-Antarctic islands and their adjacent continental landmasses. This model is a hierarchical Bayesian model with conditional dependencies between the biogeographic and molecular parameters. It implements range evolution as a continuous-time Markov Chain (CTMC) process in which transitions between discrete states, that is, migration between single areas, are modeled through a parameter representing time-dependent migration rates. These transition (migration rate) parameters are estimated by Bayesian MCMC simulations from input data comprising the original matrix of DNA sequences and the associated distribution ranges for each taxon in the matrix (Sanmartín et al., 2008). For this, we used the DNA matrices used for phylogenetic dating in BEAST as the input data for the BIB model, after excluding all outgroup taxa. In the case of *A. magellanica*, we excluded the sample “SG4 CAR21” from South Georgia, as this individual appeared as the sister group of *A. cylindristachya* and outside the core *Acaena* clade in the phylogenetic analysis (see Fig. 3A below). This exclusion was based on the possibility that this individual may represent a hybrid, as some studies have suggested that hybridization may occur between *Acaena* taxa in South Georgia (Walton, 1979; Jauregui-Lazo & Potter, 2021).

Unlike most biogeographic models (Ree & Smith, 2008; Matzke, 2022), BIB models do not require a fixed tree topology and can infer phylogenetic relationships and biogeographic parameters simultaneously. Using the hierarchical Bayesian approach, molecular parameters can be modeled individually for each target species, while biogeographic parameters (such as migration rates) are inferred across species, allowing inferences of generalities about dispersal processes (Sanmartín, 2021). Specifically, each lineage/species in the analysis is assigned its own substitution model prior, Birth–Death tree prior, and molecular-clock model prior (Sanmartín et al., 2008). The root age is also modeled individually, that is, some species might be ancient relictual taxa, while others may represent recent endemics. Species can also vary in their dispersal dynamics, that is, zoochorous plant species may have higher dispersal abilities than anemochorous species, or widespread species may have higher overall dispersal rates than endemic species (Arjona et al., 2018). To address this biological variation, the BIB model incorporates a tree height scaler, the “species-specific overall migration rate” (*m*), which is defined as the expected number of dispersal events per number of substitution events per unit time (Sanmartín et al., 2010).

In contrast to these species-specific biological parameters, the BIB model is shared across all taxa by implementing hyperpriors for the dispersal rate parameters in the CTMC process (Sanmartín et al., 2008; Sanmartín, 2022). These are, first, the “carrying capacities” (π) of each island, which are the stationary or equilibrium frequencies of the biogeographic states in the CTMC model. The carrying capacities in the BIB process can be interpreted biologically as the expected number of lineages/species that would be present in each island if the dispersal process modeled by the CTMC is allowed to run for a very long time without external disturbances (Sanmartín, 2022). In other words, the probability of arrival of a plant propagule on each island depends only on the species-specific overall migration rates connecting the islands, assuming that these are constant

over time (Sanmartín, 2022). The second type of parameter in the CTMC model is the “relative dispersal rate” (r), which is the rate of dispersal between every pair of islands, normalized by the number of lineages at the start of the process (i.e., the number of lineages sampled per island). Our dispersal rates were normalized as the number of events per unit of time per lineage and are, therefore, influenced by the number of populations sampled in our study (Sanmartín et al., 2008). If our taxon sampling does not accurately represent the extant diversity, for example, if the Eastern austral archipelagos (EAA) region (SIOBP + Macquarie) is overrepresented in our data set, our estimates will be biased. In the BIB model, carrying capacities represent the proportion of the total number of lineages or sequenced individuals expected to be found on each “island” under equilibrium conditions. Carrying capacities are exclusively informed by the modeled dispersal process and its stationary state frequencies (Sanmartín et al., 2008). Consequently, we do not anticipate finding a correlation between the number of lineages and island area, as suggested by classical Island Biogeography Theory (MacArthur & Wilson, 1967). It is noteworthy that, although the EAA is the best-sampled “BIB-island” in our data set (Appendix S1), this region, particularly the SIOBP, also harbors the greatest diversity of sub-Antarctic endemics. The “EAA” in our model consists of up to six archipelagos/islands (PEA, Macquarie, Crozet, Kerguelen, Heard, and Macquarie), and we aimed to represent each of the sub-Antarctic islands with at least one gene region. The composite “EAA island” covers a large area, with significant variation in geographic distances both among its component islands and between these islands and the continents. For instance, Macquarie Island is much closer to New Zealand—and even more so to the New Zealand sub-Antarctic islands—than to Marion Island (Fig. 1A).

We also explored possible correlations between the BIB parameters and abiotic factors traditionally associated with dispersal in island biogeography (MacArthur & Wilson, 1967). We tested the correlation between the relative dispersal rates and the minimum inter-island geographic distance using the non-parametric Spearman's rank statistical test. For this test, we used the minimum ellipsoidal geographic distance between the coastline of each landmass measured in QGIS 3.20.3, as follows: South America–Gondwanan Fragments = 500 km, South America–EAA = 6667.62 km, South America–Australia + New Zealand = 7713.18 km, Gondwanan Fragments–EAA = 5009.46 km, Gondwanan Fragments–New Zealand + Australia = 9037.8 km, and EAA–New Zealand + Australia = 1079.91 km. We also tested for significant correlations between the carrying capacity and the area of the four “BIB-islands.” The latter was calculated by the target areas shaping with polygons using the “ellipsoidal measure areas” tool in QGIS 3.20.3, based on the distribution of the target species in the continental landmasses (South American = 2 082 445.136 km², Australia + New Zealand sub-Antarctic region = 1 028 991 km²) and considering Quilty's (2007) area measures (EAA = 7830 km² and Gondwanan Fragments = 22 345 km²). Additionally, we performed a Kruskal–Wallis non-parametric analysis to test for differences in the species-specific overall migration rate between the widespread and restricted species.

We implemented the BIB model in the Bayesian software RevBayes (Höhna et al., 2016), using prespecified functions for molecular characters. Carrying capacities were modeled as Dirichlet priors with a broad concentration parameter = 1.0 (simplex) to incorporate uncertainty and lack of prior information. As in Sanmartín et al. (2008), we modeled the inter-island relative dispersal rates as symmetric, which implies that dispersal rates were equal in both directions. These pairwise dispersal rates were also modeled as Dirichlet priors ($\alpha = 1.0$). We implemented independent GTR substitution models for each species, with simplex Dirichlet priors for the four base stationary frequencies and the six relative dispersal rates. The evolutionary clock rate for each group was drawn from an exponential prior distribution with expectation = 0.0003, based on the BEAST estimates (10^{-4} – 10^{-3}). The age of the root was drawn from an exponential prior distribution with $\lambda = 0.6$ (95% credible set: 0.001–13.06 My). This broad prior was chosen to incorporate the variation in the crown age of the target species, as estimated by our BEAST analysis. Additionally, each target species was permitted to evolve under its own net diversification rates (speciation minus extinction) and turnover (the ratio of extinction to speciation). Net diversification rates were modeled using a gamma distribution (shape = 0.2, rate = 1.0), spanning a broad range of rates considered to be realistic for plants (95% credible set: 10^{-6} – 10^{-4} My); the turnover or background extinction rate was modeled with a uniform prior between 0 and 1. The “overall migration” rate per target species (m_i) was modeled using individual and identically distributed gamma distributions with parameters ($\alpha = 0.5$, $\beta = 1.5$), which generate a 50% credible set for m between 0.036 and 0.43. We ran an MCMC analysis with 10^4 generations, sampling from the priors distributions to estimate the mean and 95% confidence interval around the mean representing the posterior probability distribution for each parameter in the model (Sanmartín et al., 2008). We analyzed the results of the common BIB model in R v.4.1.2 (R Core Team, 2022) using graphics such as “box-whisker” and “density” plots. These were generated using the R packages *base*, *graphics*, and *coda* (R Core Team, 2022), as well as the software Tracer.

Initial exploratory analyses revealed that we did not have enough power to run every single island/archipelago or continental landmass as an individual state in the CTMC model, that is, we had 10 islands/landmasses resulting in the same number of carrying capacity parameters (π_i) and 45 pairwise dispersal rates (r_i). We therefore reduced the complexity of our data set by assigning biogeographic states into four categories based on their geological history: (0) South America: corresponding to the sub-Antarctic region of this continental landmass; (1) Gondwanan fragments: corresponding to the Falkland Islands and South Georgia, which are the remains of continental crust that broke away from different parts of the Gondwanan supercontinent; (2) EAA, comprising the sub-Antarctic archipelagos included in the SIOBP region and Macquarie Island; and (3) New Zealand and Australia, encompassing New Zealand's sub-Antarctic region and a continental area in Australia that, while not sub-Antarctic, supports certain sub-Antarctic lineages (Fig. 2).

To recover fine-scale dispersal patterns, especially in relation to timing and direction, we ran individual BIB models on some of the target species for which we had more comprehensive taxon sampling: the widespread species *A. magellanica* and *A. selago*, and the SIOBP-restricted species *P. antiscorbutica*. We first pruned the outgroup taxa from the time trees generated by BEAST and recoded the biogeographic states in the individual BIB models to capture finer-grain biogeographic patterns for these three species; we used four states for *A. selago* (South America, SIOBP, Falkland Islands, and Macquarie Island) and *P. antiscorbutica* (Prince Edward, Marion, Crozet, Kerguelen and Heard islands) and five states for *A. magellanica* (South America, SIOBP, Falkland Islands, South Georgia, and Macquarie Island). Similar to the multi-clade BIB analysis, we modeled carrying capacities and symmetric relative dispersal rates through Dirichlet priors with $\alpha = 1.0$. For the root age prior, we used a more informative gamma prior, spanning the mean value and 95% HPD interval estimated in the individual BEAST analyses. All other priors were identical to those used in the multi-clade BIB analysis. To ensure that these biogeographic models had a non-zero likelihood and that the MCMC started close to the estimated time tree, we used the MCC tree estimated by BEAST as the initial tree in the BIB analyses for each target species. We ran an MCMC analysis with 100 000 generations, sampling every 10th generation. For each MCMC sample, we recorded the joint posterior probability of ancestral states at each node and of changes in geographic state along each branch, estimated using a “rejection-free” stochastic mapping procedure (Freyman & Höhna, 2018). We summarized ancestral states and branch states from the joint posterior distribution as marginal probabilities in the Maximum A Posteriori (MAP) tree, discarding 25% of samples as burn-in. The R packages *RevGadgets* (Tribble et al., 2022) and *ggplot* (Wickham, 2011) were used to plot these results. The scripts to run the multiclade and individual BIB analyses in *RevBayes* are available in the GitHub page (https://github.com/isabelsanmartin/Scripts_RevBayes_SubAntarctic).

2.4 Haplotype network analyses of *A. magellanica*

Given that we investigated intraspecific genealogical relationships with potential gene flow among populations, we explored the possibility of reticulate relationships by reconstructing haplotype networks based on the chloroplast sequencing data. We only performed this analysis for *A. magellanica* since this species had the most extensive intrapopulation sampling in our data set. We removed individual “SG4 CAR21” from this analysis due to its possible hybrid origin (see Material and Methods above). We used the *pegas* R package (R Core Team, 2022; Paradis, 2010) to reconstruct the haplotype network. We first computed a distance matrix from the aligned sequences, using the *dist.dna* function and the RMST model. This model addresses ambiguities caused by the order of sequences in the data set, by randomizing the input order to produce alternative linkages in the haplotype network (Paradis, 2018). The number of randomizations required to reach convergence is affected by the number of sequences and their length and mutation rates, without considering indels (Paradis, 2018). Therefore, we simulated different numbers of random-

izations (5, 10, 20, 50, 100) and sequences (50, 100, 500, 1000), using the original number of haplotypes ($N = 9$) and sequence length ($L = 1326$); the mutation rate for the different states (μ) was constrained using the same priors as those in the *Acaena* BIB model. For the final analysis, we selected five randomizations because this resulted in the highest mean number of additional links across different values of the number of sequences (Table S8). We replicated this procedure up to four times to ensure convergence in the minimum spanning trees (MSTs), which were next consolidated into a summary network with inferred links across replications. To better understand the network structure, a haplotype tree was estimated using the *hclust* function in the *stats* R package (Murtagh & Legendre, 2014). This function allows for the estimation of a hierarchical clustering of haplotypes, represented as a dendrogram, and is based on a matrix of mutational steps. Specifically, it represents the number of mutational steps between haplotypes, calculated as Euclidean distances. Finally, we used the *pegas* R package to compute various genetic diversity metrics for each archipelago, including the number of haplotypes (H_n), nucleotide diversity (π), and haplotype diversity (H_d).

3 Results

3.1 Phylogenetic relationships and molecular dating

The species time trees inferred for the target species using BEAST have been pruned for clarity by excluding outgroups (Fig. 3). Complete phylogenetic trees, inferred with MrBayes for each target species based on concatenated sequence data sets (Figs. S1–S7), along with the complete dated trees (Figs. S8–S14) are presented.

For *Acaena*, we recovered a highly supported ($PP = 1$) group that included *Cliffortia intermedia* Harv., *Acaena cylindristachya* Ruiz & Pav., and *A. magellanica* (Fig. S1). The diversification of these three species was dated to the middle Miocene period (13.37 My; $HPD = 11.46–15.41$; $PP = 1$; Fig. S8). *A. magellanica* was reconstructed as polyphyletic, in agreement with its complex evolutionary history involving both polyploidy and hybridization processes (Walton & Greene, 1971; Walton, 1979, 1982; Jauregui-Lazo & Potter, 2021). Within the *A. magellanica* clade, the primary split reveals two lineages: one clade with a single *A. magellanica* individual of possible hybrid origin from South Georgia (“SG4 CAR21”) that groups with *A. cylindristachya* and a second clade containing all other *A. magellanica* individuals, dated in the late Miocene (stem age: 5.71 My; $HPD = 3.81–8.92$; $PP = 1$; Fig. 3A). The MRCA of this core *A. magellanica* clade was dated in the late Miocene (crown-age: 4.35 My; $HPD = 2.27–7.12$; $PP = 0.99$; Fig. 3A). This node corresponds to a split between a clade including two individuals from Macquarie Island and one from South America, and the remaining *A. magellanica* samples. The divergence between South American and Macquarie Island individuals was dated to the early Pleistocene (1.62 My; $HPD = 0.17–4.03$; $PP = 1$; Fig. 3A). A second clade grouped the individuals of the Falkland Islands and South America, with an intraspecific divergence dated to the middle Pleistocene (0.56 My; $HPD = 0.015–2.26$; $PP = 0.861$; Fig. 3A); the sister group to this clade contained subclades of varying geographic origin,

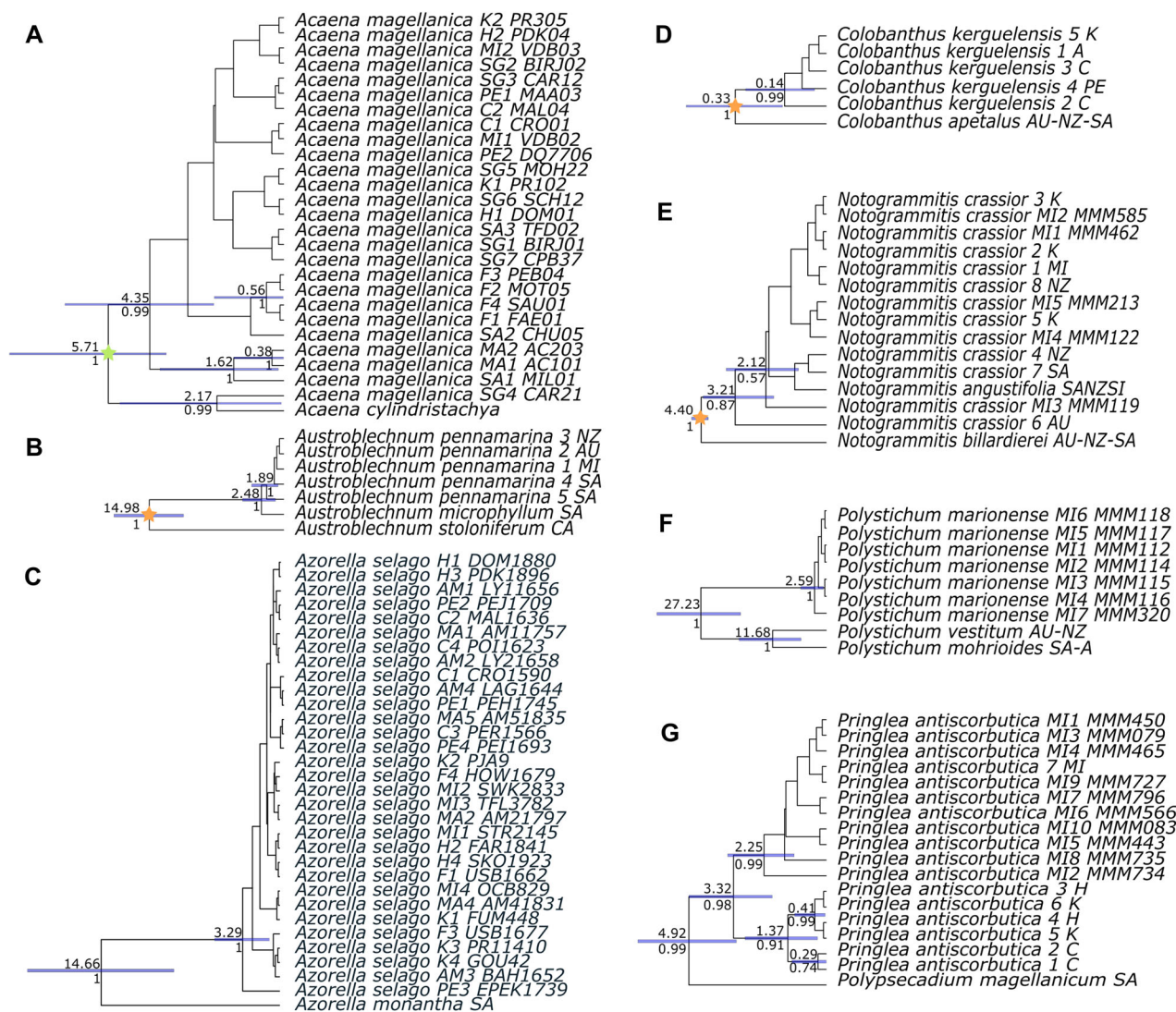


Fig. 3. Maximum clade credibility trees for each target group, depicting phylogenetic relationships and divergence times estimated using BEAST. The trees have been pruned to include only the target sub-Antarctic species and their closest relatives; complete time-calibrated trees for each group are provided in Figs. S8–S14. **A**, *A. magellanica* (Lam.) Vahl. **B**, *A. penna-marina* (Poir.) Gasper & V.A.O. Dittrich. **C**, *A. selago* Hook.f. **D**, *C. kerguelensis* Hook.f. **E**, *N. crassior* (Kirk) Parris. **F**, *P. marionense* Alston & Schelpe. **G**, *P. antiscorbutica* R. Br. ex Hook.f. Green stars represent the fossil calibration points and orange stars represent the secondary calibration points used in phylogenetic dating as listed in Table S7. See Appendix S1 for population codes. Abbreviations: A, Amsterdam Island; AN, Antipodes Island; AK, Auckland Island; AU, Australia; CA, Campbell Island; C, Crozet Islands; F, Falkland Islands; H, Heard Island; K, Kerguelen Archipelago; MA, Macquarie Island; MI, Marion Island; NZ, New Zealand; PE, Prince Edward Island; SA, South America; SG, South Georgia; T, Tristan da Cunha Group.

including accessions from South America, South Georgia, and the remaining sub-Antarctic islands. However, this clade (PP = 0.65) and the internal subclades did not receive significant support (PP < 0.10).

For *Austroblechnum*, we recovered a monophyletic clade (PP = 1; Fig. S2) comprising *A. microphyllum* (Goldm.) Gasper & V.A.O. Dittrich, *A. penna-marina*, and *A. stoloniferum* (E.Fourn.) Gasper & V.A.O. Dittrich. The divergence of *A. penna-marina* from South American *A. microphyllum* was dated to the Plio–Pleistocene (2.48 My; HPD = 0.89–4.57; PP = 1; Fig. 3B). Intraspecific divergence within *A. penna-marina* started in the early Pleistocene, corresponding to the

split of a South American individual (1.89 My; HPD = 0.63–3.57; PP = 1; Fig. 3B).

For *Azorella*, we recovered a well-supported monophyletic group containing all *A. selago* samples (PP = 1; Fig. S3). The divergence of *A. selago* from South American *A. monantha* was dated to the middle Miocene (14.66 My; HPD = 8.81–20.59; PP = 1; Fig. 3C). The first intraspecific divergence in the *A. selago* clade involved a Prince Edward Island accession in the upper Pliocene (3.29 My; HPD = 1.14–5.55; PP = 1; Fig. 3C).

For *Colobanthus*, we recovered a monophyletic clade comprising all samples of *C. apetalus*, *C. kerguelensis*, and *C. quitensis* (PP = 1; Fig. S4). We inferred the divergence

between *C. apetalus*, a species endemic to the southern regions of Australia, New Zealand, and South America, and *C. kerguelensis*, in the middle Pleistocene (0.33 My; HPD = 0.16–0.52; PP = 1; Fig. 3D). The first intraspecific divergence in *C. kerguelensis* occurred in the late Pleistocene, with the divergence of an accession from Crozet Islands (0.14 My; HPD = 0.03–0.29; PP = 0.99; Fig. 3D).

For *Notogrammitis*, all accessions of *N. angustifolia* (New Zealand, South America and the peri-Antarctic islands Tristan da Cunha, New Amsterdam, and Auckland), *N. billardiarei*, *N. crassior* (Australia, New Zealand, South America), and *N. heterophylla* (Australia, New Zealand) were clustered into a well-supported clade (PP = 0.93; Fig. S5). The divergence of *N. angustifolia* and *N. crassior* from *N. billardiarei* was dated to the early Pliocene (4.40 My; HPD = 4.11–4.69; PP = 1; Fig. 3E). The accession of *N. angustifolia* appears embedded within a *N. crassior* clade, suggesting that the latter species may be paraphyletic (Fig. 3E). The first intraspecific divergence within the *N. crassior*–*N. angustifolia* clade was dated to the upper Pliocene (3.21 My; HPD = 1.85–4.34; PP = 0.91; Fig. 3E).

For *Polystichum*, we retrieved a clade comprising all *P. marionense* accessions (PP = 1; Fig. S6). The divergence of *P. marionense* from its sister clade (South Pacific *P. vestitum* (G. Forst.) C.Presl–American *P. mohrioides* (Bory) C.Presl) was dated to the Oligocene (27.22 My; HPD = 18.62–36.75; PP = 1; Figs. 3F, S6). The first intraspecific divergence within *P. marionense* was dated to the upper Pliocene (2.59 My; PP = 1), although with broad confidence intervals for the mean divergence estimate (HPD = 0.67–5.59).

For *Pringlea*, we recovered a well-supported clade containing all *P. antiscorbutica* accessions (PP = 1; Fig. S7), sister to the South American species *Polypsecadium magellanicum* (Juss. ex Pers.) Al-Shehbaz. The divergence between *P. antiscorbutica* and *P. magellanicum* was dated to the early Pliocene (4.92 My; HPD = 3.22–6.75; PP = 0.99; Fig. 3G). The first intraspecific divergence within *P. antiscorbutica* was dated to the upper Pliocene (3.32 My; HPD = 1.94–4.92; PP = 0.98; Fig. 3G). This event corresponds to the divergence between a clade grouping all samples from Marion Island, diverging in the early Pleistocene (2.25 My; HPD = 1.15–3.53; PP = 0.99; Fig. 3G). A second group was retrieved comprising the remaining SIOBP samples: accessions from Crozet Islands appear as sister to a well-supported clade grouping samples from Heard Island and the Kerguelen Archipelago, with divergence in the early Pleistocene (1.37 My; HPD = 0.31–2.90; PP = 0.91; Fig. 3G). Furthermore, two intraspecific divergence events occurred during the middle Pleistocene comprising the Crozet Islands (0.29 My; PP = 0.74; HPD = 0–1.23; Fig. 3G) and the Kerguelen Archipelago and Heard Island (0.41 My; PP = 0.99; HPD = 0.04–1.13; Fig. 3G) samples.

Table 2 provides a summary of the estimated lineage divergence times, including the mean and 95% HPD credibility intervals for the crown and stem nodes of each target species.

3.2 Biogeographic inference

The hierarchical BIB analysis of the complete data set estimated the largest carrying capacity (Fig. S15) for the EAA (SIOBP + Macquarie Island; π_{EAA} = 0.51), followed by South America (SA, π_{SA} = 0.23) and the Gondwanan frag-

Table 2 Mean values and 95% High Posterior Density (HPD) confidence intervals for the ages of the stem and crown nodes in each of the seven plant species analyzed in this study

Species	Stem node (My)	Crown node (My)
Widespread species		
<i>A. magellanica</i>	5.71 [3.81, 8.92]	4.35 [2.27, 7.12]
<i>A. penna-marina</i>	2.89 [0.89, 4.57]	1.89 [0.631, 3.75]
<i>A. selago</i>	14.66 [8.81, 20.59]	3.29 [1.41, 5.55]
<i>N. crassior</i>	4.40 [4.11, 4.69]	3.21 [1.85, 4.34]
Restricted species		
<i>C. kerguelensis</i>	0.33 [0.16, 0.51]	0.14 [0.04, 0.29]
<i>P. marionense</i>	27.23 [18.62, 36.75]	2.59 [0.67, 5.59]
<i>P. antiscorbutica</i>	4.92 [3.22, 6.75]	3.32 [1.94, 4.92]

Abbreviations: My, million years; Restricted, species endemic to the Southern Indian Ocean Biogeographic Province (SIOBP); Widespread, species with wide distributions across the sub-Antarctic latitudes.

ments (GF, Falkland Islands + South Georgia; π_{GF} = 0.16), while the lowest value (π_{AU-NZ} = 0.09) corresponded to the continental landmasses of Australia (AU)–New Zealand (NZ) (Fig. 4A). The highest relative dispersal rate (r_i , Fig. S16) was inferred between South America and Australia–New Zealand ($r_{SA-AU/NZ}$ = 0.234), followed by dispersal rates between the EAA with Australia–New Zealand ($r_{EAA-AU/NZ}$ = 0.19), between GF and South America and GF and the EAA (r_{GF-SA} = 0.17, r_{GF-EAA} = 0.17), and between the EAA and South America (r_{EAA-SA} = 0.14). The lowest relative dispersal rate was found between the GF and Australia–New Zealand ($r_{GF-AU/NZ}$ = 0.10). Species-specific overall migration rates were higher for the widespread sub-Antarctic species (*A. magellanica*, *A. penna-marina*, *A. selago*, *N. crassior*) than for the SIOBP-restricted species (*C. kerguelensis*, *P. marionense* and *P. antiscorbutica*) (Fig. 4B). The fern *A. penna-marina* and the eudicot *P. antiscorbutica* showed the highest and lowest inferred migration rates, respectively (Fig. 4C).

Regarding the species-specific migration rates, we obtained a significant probability supporting higher overall migration rates for widespread species than for species geographically restricted to the SIOBP region (Kruskal–Wallis non-parametric test, chi-squared = 4.5; p = 0.03; Fig. 4C). Additionally, we did not find significant correlations between the BIB parameters and abiotic island features: the relative dispersal rates and the minimum inter-island geographic distance (Spearman's rank test = -0.3188741 , p = 0.5379, n = 6) and the carrying capacities and the actual size of the four “islands” included in the hierarchical BIB analysis (Spearman's R test = -0.8 , p = 0.333, n = 4). Finally, we examined the correlation between the species-specific overall migration rate (response variable) using an ANOVA that included two factors: widespread versus restricted distribution and fern versus angiosperm species. The result showed that there was no significant difference in the mean species-specific overall migration rates between ferns and angiosperms (p = 0.49). Although our dataset does not appear to be affected by these correlations, we acknowledge that testing for differences between ferns and angiosperms

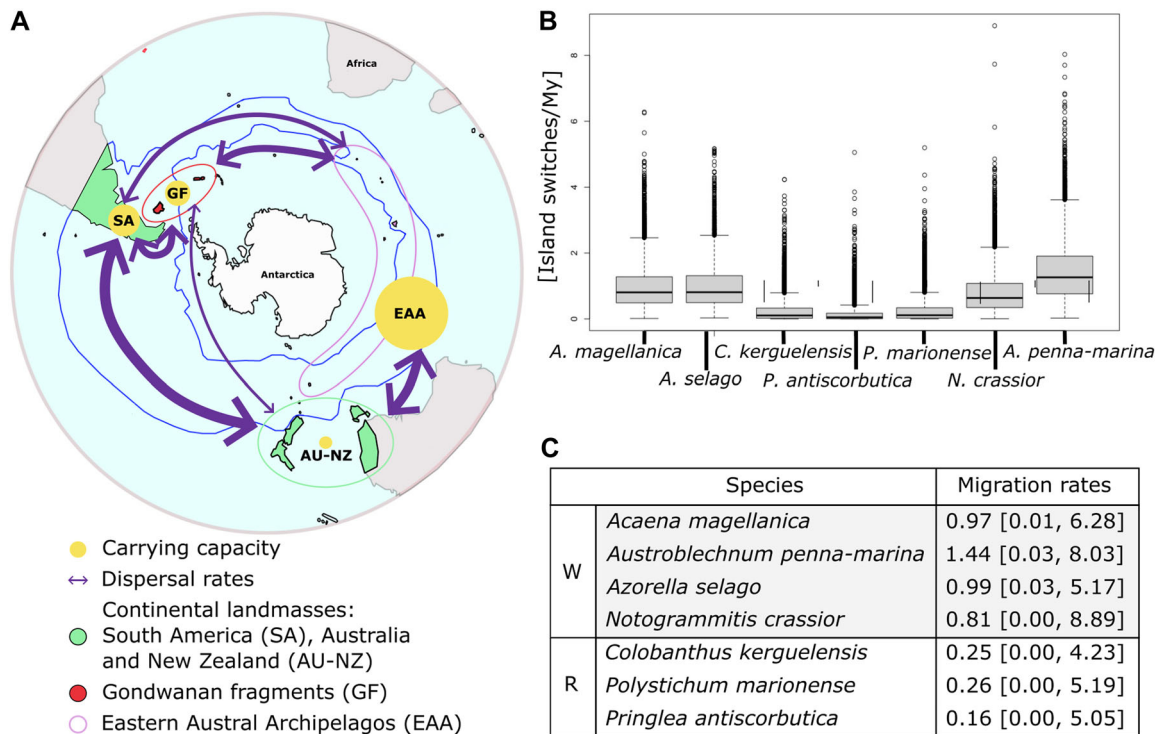


Fig. 4. Reconstruction of migrations for the seven sub-Antarctic plant species using the hierarchical BIB (Bayesian Island Biogeography) analysis. The diagram illustrates the relative dispersal rates (r) between island as blue arrows, with arrows' width proportional to the estimated rates. Yellow circles represent "island" carrying capacities (π). **A**, Map of the sub-Antarctic region. **B**, **C**, Species-specific overall migration rates (m) of the seven sub-Antarctic species: *A. magellanica* (Lam.) Vahl, *A. selago* Hook.f., *C. kerguelensis* Hook.f., *P. antiscorbutica* R. Br. ex Hook.f., *P. marionense* Alston & Schelpe, *N. crassior* (Kirk) Parris, and *A. penna-marina* (Poir.) Gasper & V.A.O. Dittrich. Abbreviations: R, restricted; W, widespread.

would be ideal in future analyses with improved sampling. However, we found that distribution type (widespread or restricted) significantly influenced migration rates ($p < 0.05$), supporting our original findings.

In the individual BIB reconstruction of *A. magellanica*, stochastic character mapping was uncertain for the longest branches, with marginal probability values increasing toward the tips of the tree as expected (Fig. S17). The MRCA (node A, Fig. 5A) was recovered as originating in the SIOBP region, although with a low marginal posterior probability ($PP_{SIOBP} = 0.33$; $PP_{SG} = 0.25$; $PP_{SA} = 0.19$). Node B (Macquarie Island clade, 1.62 My; $PP < 0.5$) recovered with Macquarie Island as the longest probable ancestral range, although also associated with some uncertainty in order of likelihood ($PP_{MA} = 0.32$; $PP_{SA} = 0.28$; $PP_{SIOBP} = 0.16$). A first dispersal event was inferred from SIOBP to South America on the onset of the Pleistocene (c. 2.5 My; $PP = 0.23$; Figs. 5A, S17) and followed slightly later by a second dispersal event from South America to Macquarie Island (c. 2.25 My; $PP = 0.3$; Figs. 5A, S17). Additionally, a back-dispersal event was inferred from Macquarie Island to South America in the middle Pleistocene (c. 1.5 My; see "SA1 MIL01" in Fig. 5A). In node C (Falkland Islands clade), the Falkland Islands are recovered as the most probable ancestral range ($PP_{FA} = 0.5$; $PP_{SA} = 0.24$; $PP_{SIOBP} = 0.11$). This clade originated from a long-distance dispersal event from SIOBP to the Falkland Islands in the early Pleistocene

(c. 2.0 My; $PP = 0.23$; Fig. 5A), followed by a subsequent dispersal from the Falkland Islands to South America during the middle Pleistocene (c. 0.75 My; $PP = 0.6$; Fig. 5A). For the remaining backbone nodes in the tree (node D), SIOBP was recovered as the ancestral area, with larger marginal probability values (Fig. 5A). Several dispersal events from SIOBP to South Georgia were inferred with higher probabilities occurring before the Last Glacial Maximum (LGM): one in the early Pleistocene (c. 1.15 My; $PP > 0.7$; Figs. 5A, S17) and four in the middle-late Pleistocene (c. 0.5 My; with PP values ranging from 0.5 to 1, the PP increases in relation to the depth of the branch; Figs. 5A, S17). We also detected a dispersal event from South Georgia to South America in the middle Pleistocene (Fig. 5A).

The most probable ancestral area for *A. selago* (Fig. 5B) was the SIOBP region ($PP = 0.60$; Fig. S18); this region was also the ancestral range inferred for the remaining nodes in the species' phylogeny, all with marginal probabilities > 0.60 (Fig. S18). Multiple dispersal events were identified connecting SIOBP with other sub-Antarctic regions: the continental landmass of South America, the Gondwanan fragment of the Falkland Islands, and the uplifted seafloor of Macquarie Island. These dispersal events were all inferred to fall within the late Pleistocene period, although pre-LGM (c. 0.25 My; Fig. 5B).

The individual BIB reconstruction of *P. antiscorbutica* (Fig. 5C) recovered this SIOBP-restricted species as having

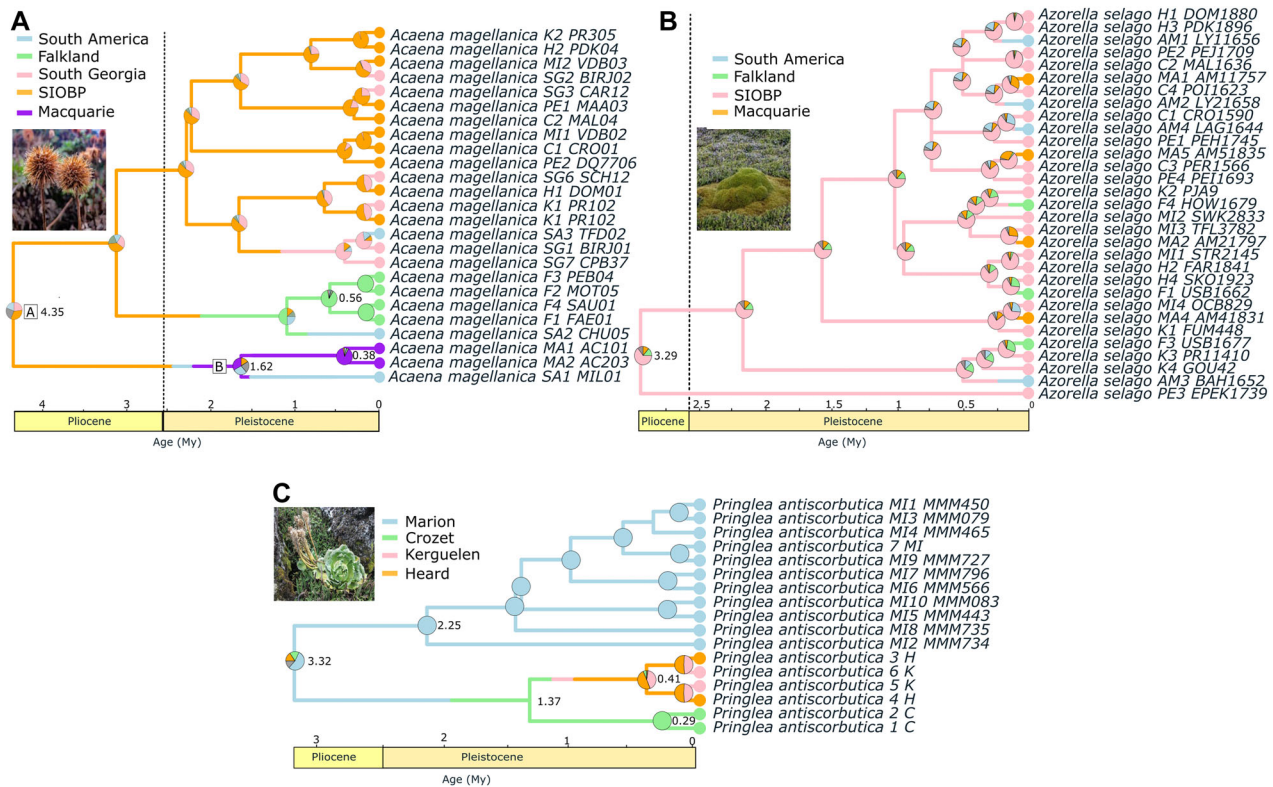


Fig. 5. Reconstruction of the ancestral area based on the individual BIB (Bayesian Island Biographic) analyses for: **A**, *A. magellanica* (Lam.) Vahl. **B**, *A. selago* Hook.f. **C**, *P. antiscorbutica* R. Br. ex Hook.f. Marginal posterior probabilities for alternative ancestral ranges are represented by pie charts at each node of the maximum-a-posteriori (MAP) tree (see legend for colors); marginal posterior probabilities for dispersal events along branches, as estimated by stochastic character mapping, are represented by colors following the legend. Divergence time estimates in million years (My) are shown for well-supported nodes (PP > 0.90). Abbreviations: A, Amsterdam Island; AN, Antipodes Island; AK, Auckland Island; AU, Australia; C, Crozet Islands; CA, Campbell Island; F, Falkland Islands; H, Heard Island; K, Kerguelen Archipelago; MA, Macquarie Island; MI, Marion Island; NZ, New Zealand; PE, Prince Edward Island; SA, South America; SG, South Georgia.

originated on Marion Island (Fig. 5C; PP = 0.58; Fig. S19). Exchanges between Marion Island and Crozet Islands (PP = 0.35) were inferred in the early Pleistocene, and between Crozet Islands and Heard Island and the Kerguelen Archipelago in the late Pleistocene (Fig. 5C).

3.3 Haplotype diversity and network structure of *A. magellanica*

We identified seven haplotypes within *A. magellanica* (Fig. 6A; Table S9). The most prevalent haplotype (H2; 64.0%) had a frequency of 40.0% in the SIOBP region and 24.0% in South Georgia. The remaining haplotypes showed more restricted distributions: H1 and H7 occurred exclusively in South America, H3 and H4 on the Falkland Islands, and H5 and H6 on Macquarie Island. The haplotype network identified haplotype H2 as the ancestral root haplotype (Fig. 6B), being most closely related to the South American haplotype H7, which was separated by one mutational step from another South American haplotype, H1. The two haplotypes found only on Macquarie Island (H5 and H6) were separated by two mutational steps from H1 and were placed in an intermediate position between South American (H1, H7) and Falkland Islands (H3, H4) haplotypes. The Falkland Islands haplotypes

were separated by the greatest number of mutational steps, five steps, from the Macquarie Island haplotypes. Hierarchical clustering grouped the seven haplotypes into three different groups (Fig. S20). One group included the haplotypes from South America, South Georgia, and SIOBP (H1, H2, H7); a second group comprised haplotypes from Macquarie Island (H5, H6); and a third group was composed of haplotypes from the Falkland Islands (H3, H4).

4 Discussion

4.1 Spatio-temporal evolution of the sub-Antarctic vascular flora

Our study emphasizes the pivotal role played by the continental landmasses that reach higher southern and sub-Antarctic latitudes (40°S–54°S)—South America, Australia, and New Zealand—in the colonization of the sub-Antarctic archipelagos. We inferred the sister species of several sub-Antarctic endemics—*A. magellanica*, *A. penna-marina*, *A. selago*, and *P. antiscorbutica*—to be the South American species *Acaena cylindristachya*, *Austroblechnum microphyllum*, *Azorella monantha*, and *Polypsecadium magellanicum*, respectively (Figs. 3A–3C, 3G). We found that the

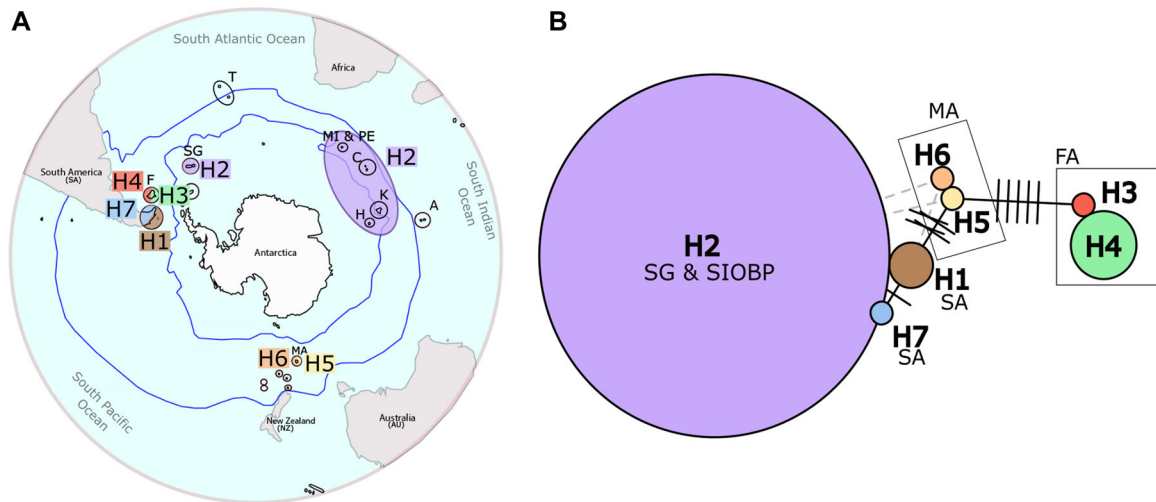


Fig. 6. Haplotype diversity and network structure of *Acaena magellanica*. **A**, Map. **B**, Haplotype distribution of *A. magellanica*. Colors in the map show the geographical location of haplotypes. Population codes are given in Appendix S1. Circle size is proportional to haplotype frequency among populations. Abbreviations: A, Amsterdam Island; AN, Antipodes Island; AK, Auckland Island; AU, Australia; C, Crozet Islands; CA, Campbell Island; F, Falkland Islands; H, Heard Island; K, Kerguelen Archipelago; MA, Macquarie Island; MI, Marion Island; NZ, New Zealand; PE, Prince Edward Island; SA, South America; SG, South Georgia.

sister species of the sub-Antarctic endemics *C. kerguelensis* and *N. crassior* are the New Zealand and Australian endemics *C. apetalus* and *N. billardi*, respectively (Figs. 3D, 3E; see also the individual gene phylogenies in Figs. S4, S5). The sister clade to *P. marionense* comprised *P. vestitum*, endemic to Australia–New Zealand, and the South American species *P. mohrioides* (Figs. 3F, S6). These findings align with other studies that have reported sister-group relationships between sub-Antarctic island endemic plants and species inhabiting the sub-Antarctic regions of the continental landmasses, for example, the angiosperm genus *Ranunculus* (Ranunculaceae; Lehnbach et al., 2017) or the brown algae genus *Durvillaea* (Fucaceae; Fraser et al., 2009). Our inferred phylogenetic relationships also agree with those recovered in previous molecular studies on the target species *A. penna-marina* (Gasper et al., 2017), *C. kerguelensis* (Biersma et al., 2020), *N. crassior* (Sundue et al., 2014; Perrie & Parris, 2012), and *P. antiscorbutica* (Bartish et al., 2012). Notably, none of the sub-Antarctic islands shows a biogeographic connection in our target species with mainland Africa, which is the geographically closest continental landmass (Fig. 1A). This contrasts with other southern islands located a few degrees further north, such as the Tristan da Cunha group in the Atlantic Ocean or Amsterdam Island in the Indian Ocean, where a close connection with Africa has been reported for plant genera such as *Phyllica* (Richardson et al., 2003).

Most of our study species diverged from their sister taxa in the last 23 My, that is, with the species' "stem age" falling within the Neogene or the Quaternary periods (Table 2). The only exception was *P. marionense*, with a stem age dated to the Oligocene period (Table 2). *Acaena magellanica* and *A. selago* showed upper-middle Miocene stem divergences (Figs. 3A, 3C; Table 2), while *A. penna-marina*, *C. kerguelensis*, *N. crassior*, and *P. antiscorbutica* showed younger stem ages,

spanning the Pliocene and Pleistocene periods (Figs. 3B, 3D, 3E, 3G; Table 2). These findings support long-distance dispersal over tectonic vicariance as the most plausible mechanism for the initial colonization of the sub-Antarctic region by these plant species. Moreover, the timing of these colonization events may be even more recent if we consider that the divergence between the insular species and their sister taxa might overestimate the age of arrival. For instance, colonization might have involved other islands that are geographically "in between", followed by species extinction on these intermediate islands, or it could have occurred via stepping-stone dispersal through intermediate landmasses, such as now-submerged islands. This scenario, suggested for the colonization of the Hawaiian and Canarian archipelagos (Fernández-Palacios et al., 2011; García-Verdugo et al., 2019), predicts that the crown age—that is, the onset of intraspecific divergence within an island endemic—provides a more accurate estimate of the timing of island colonization than the stem age. In our study, the "crown ages" of the sub-Antarctic species (Fig. 3; Table 2), ranging from 4.35 My (*A. magellanica*) to 0.14 My (*C. kerguelensis*), substantially postdate the final fragmentation of the Gondwanan supercontinent. This lends further support to the hypothesis that colonization of the sub-Antarctic region occurred primarily through long-distance dispersal events across marine barriers.

The inferred "crown ages" of most target species not only conflict with tectonic vicariance but also, according to our divergence time estimates (including their confidence intervals), predate the Last Glacial Maximum (Table 2). In the case of *A. magellanica*, *A. selago*, *N. crassior*, *P. antiscorbutica*, and *P. marionense*, their "crown ages" also predate the onset of Pleistocene glaciations (2.5 My ago; Table 2). For *A. penna-marina*, the crown age falls within this period (Table 2). These estimates support the "refugia"

hypothesis and contrast with the “recolonization” hypothesis (Table 1). The only exception is *C. kerguelensis*, whose crown age dated to after the LGM, aligning with the “recolonization” hypothesis (Fig. 3D). The “refugia” hypothesis has previously been supported for various taxa (reviewed in Fraser et al., 2012), as suggested by paleobotanical data on vascular plants (van der Putten et al., 2010), and by molecular evidence such as the case of *Pleurophyllum* (Asteraceae). This species is presumed to have survived the Pleistocene glaciation in the sub-Antarctic regions, specifically on Macquarie Island and New Zealand's sub-Antarctic islands (Wagstaff et al., 2011). Further support for this hypothesis comes from phylogeographic studies of other organisms showing deep crown ages predating the LGM, such as the freshwater crustacean *Boeckella poppei* (Maturana et al., 2021) and oribatid mites (Mortimer et al., 2011; van Vuuren et al., 2018). Additional evidence, including intra-island phylogenetic differences among populations of *A. selago*, also suggests colonization events predating the Quaternary period (Mortimer et al., 2008).

We note that some of our divergence time estimates are apparently in conflict with geological information, such as instances of the inferred origin of an island endemic predating the emergence of the island it inhabits. For example, the first divergence events within *A. selago*, *P. marionense*, and *P. antiscorbutica* correspond to the split of a population endemic to Prince Edward Archipelago (PEA). However, the geological age of these two islands (c. 0.45 My; Quilty, 2007) substantially postdates the crown ages of *A. selago* (3.29 My), *P. marionense* (2.59 My), and *P. antiscorbutica* (3.32 My) inferred by our analyses (Figs. 3B, 3F, 3G). Incomplete taxon sampling of the target species and among the outgroup taxa might have decreased the accuracy of our divergence time estimates, through its effects on branch lengths and calibration priors (Ho et al., 2015). However, we notice that these results hold even after accounting for the broad credibility intervals that surround our mean divergence time estimates: the geological age of PEA postdates the lowest bound of the 95% HPD credibility interval in *A. selago* (1.41 My), *P. antiscorbutica* (1.94 My), and *P. marionense* (0.67 My; Table 2). A more plausible explanation is that the populations of these three species endemic to the PEA originated on another geologically older island or landmass. After colonizing the PEA, the original source population(s) went extinct. This scenario is supported by the fact that most of the sub-Antarctic islands on both sides of the PEA experienced near-complete glaciations (South Georgia, Kerguelen Archipelago, and Heard Island; Fig. 1A), while the islands of the PEA and Crozet either avoided glaciation or were affected to a lesser extent (Convey, 2007; Quilty, 2007). This scenario has been suggested for other archipelagos (García-Verdugo et al., 2019). For example, in the case of the Canary Island endemic *Euphorbia balsamifera*, the youngest volcanic island in the archipelago, El Hierro Island, was identified as the source of propagules that dispersed to geologically older islands (Rincón-Barrado et al., 2024). Another possibility is that these species are present on other islands that were not sampled in our study (Fig. 2). Increasing population sampling efforts within these species, and among their closest relatives, are needed to test this hypothesis.

In summary, initial colonization events of the sub-Antarctic region in our data set—defined by either the stem or the crown ages of the study species—generally postdate the final stages of Gondwanan fragmentation. In addition, despite the limitations in our taxon sampling, our findings largely support the refugia hypothesis and favor across-oceanic dispersal as the main process driving floristic assembly in the sub-Antarctic region.

4.2 Connectivity and direction of colonization within the sub-Antarctic vascular flora

The archipelagos and islands located in the sub-Antarctic region are considered the most remote insular systems in the world, separated by thousands of kilometers from each other and any continental landmasses. Despite their remoteness and geographic isolation, our results show that these islands have maintained significant biological connectivity (Fig. 4). This connectivity is particularly evident in the fact that the highest inferred dispersal rate connects the most distant BIB unit areas—South America and Australia–New Zealand (Fig. 4A)—which are separated by thousands of kilometers across the South Pacific Ocean, with no intermediate sub-Antarctic islands that could serve as stepping stones. Furthermore, we inferred a significantly higher species-specific overall migration rate for widespread species compared to geographically restricted species in the area (Figs. 4B, 4C).

The inferred relative dispersal rates, ranging from 0.10 to 0.234 dispersal events per lineage per million years (Fig. 4A), are similar to those reported by Sanmartín et al. (2008) for the Atlantic Canary Islands, despite the latter being geographically much closer to the relevant continental landmass. However, unlike Sanmartín et al. (2008), we found no strong correlation between the relative dispersal rates and the minimum inter-island geographic distance. This is in agreement with Biersma et al. (2017), who inferred that bipolar Polytrichales mosses were connected by very infrequent long-distance dispersal events, as few as four in 10 million years depending on the dating analysis used. Carrying capacities in our study also appeared to be unrelated to the actual area of the four “islands” used in the BIB analysis. They were estimated to be largest for the EAA region (7830 km²) and lowest for Australia–New Zealand (1028991 km²) (Fig. 4A). Sanmartín et al. (2008) reported similar patterns in the Canary Islands, with island carrying capacities unrelated to their current sizes. Carrying capacities in the BIB model are estimated directly from the inferred transition (dispersal) probabilities between the “island” states (Sanmartín et al., 2008). Akin to a molecular evolutionary process, dispersal in BIB is modeled as a time-homogeneous process that reaches a stationary state if left undisturbed, at which point the carrying capacities represent the equilibrium frequencies of island diversity, that is, the number of lineages expected on each island if this only depends on the model's dispersal probabilities (Sanmartín, 2022). Since parameters in the BIB model are not informed by any abiotic factors, such as island size or inter-island geographic distance, a direct correlation between these factors and the inferred carrying capacities and dispersal rates is not necessarily expected, even

though these abiotic factors have probably contributed to the dispersal history of the sub-Antarctic archipelagos.

Regarding the direction of dispersal events, our species-specific individual BIB models provide examples that potentially support both the anisotropic (eastward) and isotropic (westward) hypotheses. A notable case is that of *A. magellanica*, where the majority of the dispersal events were inferred to be in the opposite direction to the sub-Antarctic marine currents. For instance, as shown in Fig. 5A, five dispersal events are inferred from the SIOBP to South Georgia, one event from South Georgia to South America, one event from the SIOBP to the Falkland Islands, and one event from the Falkland Islands to South America. This asymmetric connectivity between regions was also supported by the haplotype analysis (Fig. 6B). Specifically, the most frequent haplotype (H2), shared between the SIOBP and South Georgia, had the fewest intermediate mutation steps separating it from the other South American haplotypes (H1 and H7). *Aceana magellanica* is a zoochorous species, so dispersal against the prevailing WWD and ACC currents is plausible. The species' fruit also matures in late summer, when summer breeding birds such as gulls and skuas leave South Georgia and migrate toward South America. There are seabirds that require more frequent breaks for rest and foraging while crossing the South Pacific, particularly when facing strong headwinds that oppose the WWD (Guilford et al., 2009; Landers et al., 2011). Zoochorous plants, such as *A. magellanica*, might be dispersed attached to the feathers of these seabirds, facilitating colonization of land where birds stop as part of a "stepping-stone" dispersal strategy. Observations have documented seeds of *A. magellanica* attached to the feathers or feet of flying seabirds (e.g., Bergstrom et al., 2006). Mazibuko et al. (2024) experimentally demonstrated that more than half (53%) of this species' seeds tested on Marion Island for their ability to attach to seabird feathers could potentially disperse via zoochory.

Within *A. selago*, both westward and eastward dispersal events were inferred. For instance, four isotropic events (from the SIOBP to Macquarie Island) and seven anisotropic events (four from the SIOBP to South America and three from the SIOBP to the Falkland Islands) are inferred (Fig. 5B). This species is anemochorous, which agrees with the inference of anisotropic dispersal following the direction of the WWD current, for example, from SIOBP to Macquarie Island (Fig. 5B). Isotropic events against the prevailing currents are more difficult to explain. Some authors have suggested explanations for westward dispersal in the southern latitudes, such as the existence of easterly anticyclones, moving from east to west, or short-distance stepping-stone dispersal using island arc ridges or now-submerged plateaus connecting landmasses separated by water barriers (Michaux & Leschen, 2005; Sanmartín et al., 2007; Winkworth et al., 2015). An additional alternative is the possibility of propagules drifting with the prevailing eastward currents all the way around the Southern Hemisphere rather than going against these currents, a scenario that has been suggested for drifting buoys, tree trunks, and seaweed rafts entrained in the ACC (Patterson, 1985; Macaya et al., 2016).

In the geographically restricted *P. antiscorbutica*, all inferred dispersal events occurred in an anisotropic eastward direction, from the westernmost islands of the SIOBP to the eastern archipelagos. This dispersal route seems to have followed a stepping-stone progression, from Prince Edward Archipelago in the west to the Crozet Islands, Kerguelen Archipelago, and Heard Island in the east (Fig. 5C). An eastward geographic progression is further supported by the more severe glaciation in the eastern SIOBP islands compared to the western islands, with the latter acting as refugia and sources of later migration events. Eastward dispersal may have been facilitated by the prevailing marine currents, with hydrochory proposed as a probable dispersal mechanism. The mucilaginous seeds of *P. antiscorbutica* could have enabled rafting, at least along the coasts of the Kerguelen Archipelago, as suggested by Bartish et al. (2012). Supporting this, experiments on seed viability and buoyancy of *P. antiscorbutica* revealed that 77% of viable seeds were buoyant (Mazibuko et al., 2024).

4.3 The critical role and conservation value of sub-Antarctic islands as biodiversity refugia

The sub-Antarctic region, due to its latitudinal position and isolation, has emerged as a unique biodiversity cradle. Its biota is characterized by a small number of species that show high levels of endemism (Convey, 2007; Chau et al., 2020), underscoring the region's conservation significance. Our study supports the hypothesis that the sub-Antarctic islands have acted as long-term refugia for species that colonized these regions from the southern latitudes of South America or Australia–New Zealand prior to periods of glaciation (Hulton et al., 2002; Convey, 2007). Our crown-age estimates indicated that these colonization events have spanned at least the last four million years and possibly much longer if sister-group divergence timings are representative of colonization timings. Among the different sub-Antarctic islands and archipelagos, those that are Gondwanan fragments (the Falkland Islands and South Georgia) are likely to have played a significant role as sources for secondary colonization events (Figs. 5A–5C). Extended isolation in Gondwanan fragments could also have facilitated the maintenance or emergence of polyploid lineages (van de Peer et al., 2017). This is likely the case for the "Falkland Islands clade" of *A. magellanica* (Figs. 3A, 5A), which groups all individuals from the Falkland Islands into a single clade that may consist of tetraploid cytotypes, in contrast to the diploid cytotypes found in populations across South America and the rest of the sub-Antarctic region (Marticorena & Cavieres, 2000).

The SIOBP region has been recognized as a cradle of unique phylogenetic diversity in phylogeographic studies (Chown, 1994; Chown et al., 1998), with numerous intra-specific diversification events identifying it as the ancestral range. This suggests that allopatric speciation, mediated by marine barriers between islands, has played a significant role in shaping the region's flora. Some of the SIOBP islands were not heavily/completely glaciated during the LGM and previous glacial maxima (e.g., PEA, Crozet Islands), unlike other ancient islands such as South Georgia, Kerguelen Archipelago, and Heard Island, which were almost completely or completely glaciated (Convey, 2007; Fraser

et al., 2009). The younger SIOBP Archipelagos might have acted as climatic “refugia within refugia” (Gómez & Lunt, 2007), consistent with the higher carrying capacity detected for this region (Fig. 4A). These archipelagos could also have served as source areas, with recolonization potentially occurring toward other SIOBP islands, ancient Gondwanan fragments, or the continental landmasses. In contrast, Gondwanan fragments such as South Georgia and the sub-Antarctic regions of the continental landmasses were more heavily glaciated (Clapperton, 1990). Additionally, the ocean's tempering effect may have enabled the persistence of plant lineages in the SIOBP archipelagos (Fraser et al., 2012). The role of islands as crucial refuges and sources of diversity for the continental mainlands is supported in numerous phylogeographic studies on vascular plants (Fernández-Mazuecos & Vargas, 2011; García-Verdugo et al., 2021) and also for the sub-Antarctic islands in two studies on crabs (Xu et al., 2009) and oribatid mites (Mortimer et al., 2011).

Although conservation is not the main focus of the study, it is important to note that preserving this unique intra-specific genetic diversity is crucial, as it supports evolutionary potential and enables species to adapt to environmental change (Mairal et al., 2018).

5 Conclusions

Our study of the spatio-temporal evolution of the sub-Antarctic vascular flora supports long-distance dispersal as the primary mode of colonization. We highlight the critical role of sub-Antarctic islands as unique biodiversity refugia over millions of years, with some areas experiencing only mild glaciation, thus allowing the persistence of some species during glacial maxima. Additionally, we identify instances where these islands have acted as evolutionary cradles for further diversification (Schrader et al., 2024), with phylogenetic analyses indicating that several endemic species have evolved in a context of isolation. Despite their remote locations, biological connectivity—both isotropic and anisotropic—has facilitated the persistence and enhancement of genetic diversity of plant lineages within the sub-Antarctic islands. We have detected anisotropic colonization events that align with the prevailing eastward winds and ocean currents in the region, as well as several isotropic events, potentially mediated by zoochory. Future research into the evolutionary and biogeographic history of more endemic species of the sub-Antarctic islands is needed for refining our understanding of the biogeographic processes mediating plant community assembly in this region. Given these islands' isolation at the Earth's southernmost latitudes and the ongoing threats posed by anthropogenic activities and invasive species to their unique floras and faunas, we argue that urgent conservation measures are required for the sub-Antarctic regions.

Acknowledgements

We thank Bruce Dyer, Daniela Monsanto, and John H. Chau for help with fieldwork. Financial and logistical support was provided by the South African National Research Foundation

(NRF) and by the South African National Antarctic Programme (SANAP). MM was supported by the National Research Foundation (grant 89967). Thanks are due to Dr. Michael Sundue for assisting us with nomenclatural questions regarding the genus *Notogrammitis* and Dr. Germain Rouhan for clarifications on the distribution of *P. marionense*. We also thank Dr. Igor Bartish for providing us with crucial information about the biology of *P. antiscorbutica*. We thank the Herbarium of the Musée National d'Histoire Naturelle de Paris (P) for the loan of specimens necessary for this study (P05441843; P05441844; P00157990; P00157986; P00749907; P06140118; and P00749922). AA was supported by a contract from the programme YO INVESTIGO (13-2022-005867), autonomous government of Madrid (CAM), PRTR and NextGenerationEU, and the predoctoral grant (PIPF-2022/ECO-24376) funded by CAM and supervised by IS. IS was supported by projects PID2019-108109GB-I00 funded by MCIN/AEI/10.13039/501100011033; and PID2023-153023NB-I00 by MICIU/AEI/10.13039/501100011033 and FEDER/UE. PC is supported by NERC core funding to the British Antarctic Survey's “Biodiversity, Evolution and Adaptation” Team.

Author Contributions

Ángela Aguado-Lara: investigation, formal analysis, and writing (original draft and revisions); Isabel Sanmartín: conceptualization, formal analysis, and writing (original draft and revisions); Johannes J. Le Roux: investigation and writing (review and editing); Carlos García-Verdugo: investigation and writing (review and editing); Sonia Molino: investigation and writing (review and editing); Peter Convey: writing (review and editing); Bettine Jansen van Vuuren: investigation and writing (review and editing); Mario Mairal: conceptualization, fieldwork, investigation, formal analysis, and writing (original draft and revisions).

Conflicts of Interest

The authors declare no conflict of interest.

References

- Alfaro ME, Zoller S, Lutzoni F. 2003. Bayes or bootstrap? A simulation study comparing the performance of Bayesian Markov chain Monte Carlo sampling and bootstrapping in assessing phylogenetic confidence. *Molecular Biology and Evolution* 20: 255–266.
- Andersson L, Kocsis M, Eriksson R. 2006. Relationships of the genus *Azorella* (Apiaceae) and other hydrocotyloids inferred from sequence variation in three plastid markers. *Taxon* 55: 270–280.
- Arjona Y, Nogales M, Heleno R, Vargas P. 2018. Long-distance dispersal syndromes matter: Diaspore-trait effect on shaping plant distribution across the Canary Islands. *Ecography* 41: 805–814.
- Bahr A, Kaboth-Bahr S, Karas C. 2022. The opening and closure of oceanic seaways during the Cenozoic: Pacemaker of global climate change? *Geological Society London Special Publications* 523: 141–171.
- Baele G, Lemey P, Vansteelandt S. 2013. Make the most of your samples: Bayes factor estimators for high-dimensional models of sequence evolution. *BMC Bioinformatics* 14: 1–18.

- Baker CM, Boyer SL, Giribet G. 2020. A well-resolved transcriptomic phylogeny of the mite harvestman family Pettalidae (Arachnida, Opiliones, Cyphophthalmi) reveals signatures of Gondwanan vicariance. *Journal of Biogeography* 47: 1345–1361.
- Bartish IV, Aïnouche A, Jia D, Bergstrom D, Chown SL, Winkworth RC, Hennion F. 2012. Phylogeny and colonization history of *P. antiscorbutica* (Brassicaceae), an emblematic endemic from the South Indian Ocean Province. *Molecular Phylogenetics and Evolution* 65: 748–756.
- Bergstrom DM, Hodgson DA, Convey P. 2006. The physical setting of the Antarctic. In: Bergstrom DM, Convey P, Huiskes AHL eds. *Trends in Antarctic terrestrial and limnetic ecosystems: Antarctica as a global indicator*. Dordrecht: Springer. 15–33.
- Biersma EM, Jackson J, Hyvönen J, Koskinen S, Linse K, Griffiths H, Convey P. 2017. Global movements in bipolar moss species. *Royal Society Open Science* 4: 170147.
- Biersma EM, Torres-Díaz C, Molina-Montenegro MA, Newsham, Kevin K, Vidal MA, Collado GA, Acuña-Rodríguez IS, Ballesteros GI, Figueroa CC, Goodall-Copestake WP, Leppe MA, Cuba-Díaz M, Valladares MA, Perterra LR, Convey P. 2020. Multiple late-Pleistocene colonization events of the Antarctic pearlwort *Colobanthus quitensis* (Caryophyllaceae) reveal the recent arrival of native Antarctic vascular flora. *Journal of Biogeography* 47: 1663–1673.
- Breitwieser I, Glenn DS, Thorne A, Wagstaff SJ. 1999. Phylogenetic relationships in Australasian Gnaphalieae (Compositae) inferred from ITS sequences. *New Zealand Journal of Botany* 37: 399–412.
- Burridge CP, McDowall RM, Craw D, Wilson MV, Waters JM. 2012. Marine dispersal as a pre-requisite for Gondwanan vicariance among elements of the galaxiid fish fauna. *Journal of Biogeography* 39: 306–321.
- Chau JH, Mtsi NIS, Münbergová Z, Greve M, Le Roux PC, Mairal M, Le Roux JJ, Dorrington RA, Jansen, van Vuuren B. 2020. An update on the indigenous vascular flora of sub-Antarctic Marion Island: Taxonomic changes, sequences for DNA barcode loci, and genome size data. *Polar Biology* 43: 1817–1828.
- Chown SL. 1994. Historical ecology of sub-Antarctic weevils (Coleoptera: Curculionidae): Patterns and processes on isolated islands. *Journal of Natural History* 28: 411–433.
- Chown SL, Gremmen NJM, Gaston KJ. 1998. Ecological biogeography of Southern Ocean islands: Species-area relationships, human impacts, and conservation. *American Naturalist* 152: 562–575.
- Clapperton CM. 1990. Quaternary glaciations in the Southern Hemisphere: An overview. *Quaternary Science Reviews* 9: 299–304.
- Convey P. 2007. Influences on and origins of terrestrial biodiversity of the sub-Antarctic islands. *Papers and Proceedings of the Royal Society of Tasmania* 141: 83–93.
- Convey P, Biersma EM, Casanova-Katny A, Maturana CS. 2020. Refuges of Antarctic diversity. In: Oliva M, Ruiz-Fernández J eds. *Past Antarctica*, Chapter 10. Burlington: Academic Press. 181–200.
- Crowl AA, Miles NW, Visger CJ, Hansen K, Ayers T, Haberle R, Cellinese N. 2016. A global perspective on Campanulaceae: Biogeographic, genomic, and floral evolution. *American Journal of Botany* 103: 233–245.
- Darlington PJ. 1959. Area, climate, and evolution. *Evolution* 13: 488–510.
- Darriba D, Taboada GL, Doallo R, Posada D. 2012. jModelTest 2: More models, new heuristics and parallel computing. *Nature Methods* 9: 772.
- Deppe L. 2012. *Spatial and temporal patterns of at-sea distribution and habitat use of New Zealand albatrosses*. Ph.D. Dissertation. Christchurch: University of Canterbury.
- Drummond AJ, Ho SYW, Phillips MJ, Rambaut A. 2006. Relaxed phylogenetics and dating with confidence. *PLoS Biology* 4: e88.
- Edelman DW. 1975. *The Eocene Germer Basin flora of south-central Idaho*. M.S. Thesis. Moscow, Idaho: University of Idaho.
- Fernández-Mazuecos M, Vargas P. 2011. Genetically depauperate in the continent but rich in oceanic islands: *Cistus monspeliensis* (Cistaceae) in the Canary Islands. *PLoS One* 6: e17172.
- Fernández-Palacios JM, De Nascimento L, Otto R, Delgado JD, García-del-Rey E, Arévalo JR, Whittaker RJ. 2011. A reconstruction of Palaeo-Macaronesia, with particular reference to the long-term biogeography of the Atlantic island laurel forests. *Journal of Biogeography* 38: 226–246.
- Fior S, Karis PO, Casazza G, Minuto L, Sala F. 2006. Molecular phylogeny of the Caryophyllaceae (Caryophyllales) inferred from chloroplast matK and nuclear rDNA ITS sequences. *American Journal of Botany* 93: 399–411.
- Fraser CI, Nikula R, Ruzzante DE, Waters JM. 2012. Poleward bound: Biological impacts of Southern Hemisphere glaciation. *Trends in Ecology & Evolution* 27: 462–471.
- Fraser CI, Nikula R, Spencer HG, Waters JM. 2009. Kelp genes reveal effects of subantarctic sea ice during the Last Glacial Maximum. *Proceedings of the National Academy of Sciences USA* 106: 3249–3253.
- Frenot Y, Chown SL, Whinam J, Selkirk PM, Convey P, Skotnicki M, Bergstrom DM. 2005. Biological invasions in the Antarctic: Extent, impacts and implications. *Biological Reviews* 80: 45–72.
- Freyman WA, Höhna S. 2018. Cladogenetic and anagenetic models of chromosome number evolution: A Bayesian model averaging approach. *Systematic Biology* 67: 195–215.
- García-Verdugo C, Caujapé-Castells J, Sanmartín I. 2019. Colonization time on island settings: Lessons from the Hawaiian and Canary Island floras. *Botanical Journal of the Linnean Society* 191: 155–163.
- García-Verdugo C, Mairal M, Tamaki I, Msanda F. 2021. Phylogeography at the crossroad: Pleistocene range expansion throughout the Mediterranean and back-colonization from the Canary Islands in the legume *Bituminaria bituminosa*. *Journal of Biogeography* 48: 1622–1634.
- Gasper AL, Almeida TE, Dittrich VAO, Smith AR, Salino A. 2017. Molecular phylogeny of the fern family Blechnaceae (Polypodiales) with a revised genus-level treatment. *Cladistics* 33: 429–446.
- Gernhard T. 2008. New analytic results for speciation times in neutral models. *Bulletin of Mathematical Biology* 70: 1082–1097.
- Gómez A, Lunt DH. 2007. Refugia within refugia: Patterns of phylogeographic concordance in the Iberian peninsula. In: Weiss S, Ferrand N eds. *Phylogeography of Southern European refugia: Evolutionary perspectives on the origins and conservation of European biodiversity*. Netherlands: Springer. 155–188.
- Guilford T, Meade J, Willis J, Phillips RA, Boyle D, Roberts S, Collett M, Freeman R, Perrins CM. 2009. Migration and stopover in a small pelagic seabird, the Manx shearwater *Puffinus puffinus*: Insights from machine learning. *Proceedings of the Royal Society B: Biological Sciences* 276: 1215–1223.
- Hall K. 2004. Quaternary glaciation of the sub-Antarctic Islands. In: Ehlers J, Gibbard PL eds. *Quaternary glaciations—Extent and chronology, Part III*. Canada: Elsevier. 339–345.
- Ho SYW, Duchêne S, Duchêne D. 2015. Simulating and detecting autocorrelation of molecular evolutionary rates among lineages. *Molecular Ecology Resources* 15: 688–696.
- Höhna S, Landis MJ, Heath TA, Boussau B, Lartillot N, Moore BR, Huelsenbeck JP, Ronquist F. 2016. RevBayes: Bayesian Phylogenetic Inference Using Graphical Models and an Interactive Model-Specification Language. *Systematic Biology* 65: 726–736.

- Hooker JD, Fitch WH. 1844. *The botany of the Antarctic voyage of HM discovery ships Erebus and Terror in the years 1839–1843: Under the command of Captain Sir James Clark Ross (1)*. London: Reeve.
- Hulton NRJ, Purves RS, McCulloch RD, Sugden DE, Bentley MJ. 2002. The Last Glacial Maximum and deglaciation in southern South America. *Quaternary Science Reviews* 21: 233–241.
- Jauregui-Lazo J, Potter D. 2021. Phylogeny and biogeography of *Acaena* (Rosaceae) and its relatives: Evidence of multiple long-distance dispersal events across the globe. *Systematic Botany* 46: 998–1010.
- Kingman JFC. 1982. The coalescent. *Stochastic Processes and their Applications* 13: 235–248.
- Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution* 35: 1547–1549.
- Landers TJ, Rayner MJ, Phillips RA, Hauber ME. 2011. Dynamics of seasonal movements by a transPacific migrant, the Westland petrel. *The Condor* 113: 71–79.
- Le Péchon T, Zhang L, He H, Zhou XM, Bytebier B, Gao XF, Zhang LB. 2016. A well-sampled phylogenetic analysis of the polystichoid ferns (Dryopteridaceae) suggests a complex biogeographical history involving both boreotropical migrations and recent transoceanic dispersals. *Molecular Phylogenetics and Evolution* 98: 324–336.
- Le Roux JJ, Clusella-Trullas S, Mokotjomela TM, Mairal M, Richardson DM, Skein L, Wilson JR, Weyl OLF, Geerts S. 2020. Biotic interactions as mediators of biological invasions: Insights from South Africa. *Biological Invasions in South Africa* 35: 387–427.
- Lehnebach CA, Winkworth RC, Becker M, Lockhart PJ, Hennion F. 2017. Around the pole: Evolution of sub-Antarctic *Ranunculus*. *Journal of Biogeography* 44: 875–886.
- Livemore R, Hillenbrand CD, Meredith M, Eagles G. 2007. Drake Passage and Cenozoic climate: An open and shut case? *Geochemistry, Geophysics, Geosystems* 8: Q01005.
- Lopes JC, Fonseca LHM, Johnson DM, Luebert F, Murray N, Nge FJ, Rodrigues-Vaz C, Soulé V, Onstein RE, Lohmann LG, Couvreur TL. 2024. Dispersal from Africa to the Neotropics was followed by multiple transitions across Neotropical biomes facilitated by frugivores. *Annals of Botany* 133: 659–676.
- MacArthur R, Wilson EO. 1967. *The theory of island biogeography*. Princeton, NJ: Princeton University Press.
- Macaya EC, López B, Tala F, Tellier F, Thiel M. 2016. Float and raft: Role of buoyant seaweeds in the phylogeography and genetic structure of non-buoyant associated flora. In: Hu Z-M, Fraser C eds. *Seaweed phylogeography: Adaptation and evolution of seaweeds under environmental change*. Dordrecht: Springer. 97–130.
- Mairal M, Caujapé-Castells J, Pellissier L, Jaén-Molina R, Álvarez N, Heuertz M, Sanmartín I. 2018. A tale of two forests: Ongoing aridification drives population decline and genetic diversity loss at continental scale in Afro-Macaronesian evergreen-forest archipelago endemics. *Annals of Botany* 122: 1005–1017.
- Mairal M, Chown SL, Shaw J, Chala D, Chau JH, Hui C, Kalwij JM, Münzbergová Z, Jansen Van Vuuren B, Le Roux JJ. 2022. Human activity strongly influences genetic dynamics of the most widespread invasive plant in the sub-Antarctic. *Molecular Ecology* 31: 1649–1665.
- Mairal M, García-Verdugo C, Le Roux JJ, Chau JH, Van Vuuren BJ, Hui C, Münzbergová Z, Chown SL, Shaw JD. 2023. Multiple introductions, polyploidy and mixed reproductive strategies are linked to genetic diversity and structure in the most widespread invasive plant across Southern Ocean archipelagos. *Molecular Ecology* 32: 756–771.
- Marticorena AE, Cavieres LA. 2000. *A. magellanica* (Lam.) Vahl (Rosaceae). *Gayana Botánica* 57: 107–113.
- Maturana CS, Rosenfeld S, Biersma EM, Segovia NI, González-Wevar CA, Díaz A, Naretto J, Duggan IC, Hogg ID, Poulin E, Convey P, Jackson JA. 2021. Historical biogeography of the Gondwanan freshwater genus *Boeckella* (Crustacea): Timing and modes of speciation in the Southern Hemisphere. *Diversity and Distributions* 27: 2330–2343.
- Matzke NJ. 2022. Statistical comparison of DEC and DEC+J is identical to comparison of two ClaSSE submodels and is therefore valid. *Journal of Biogeography* 49: 1805–1824.
- Mazibuko N, Greve M, le Roux PC. 2024. Dispersal potential does not predict recent range expansions of sub-Antarctic plant species. *Polar Biology* 47: 1–16.
- Michaux B, Leschen RAB. 2005. East meets west: Biogeology of the Campbell Plateau. *Biological Journal of the Linnean Society* 86: 95–115.
- Moon KL, Chown SL, Fraser CI. 2017. Reconsidering connectivity in the sub-Antarctic: Reconsidering connectivity in the sub-Antarctic. *Biological Reviews* 92: 2164–2181.
- Mortimer E, Jansen van Vuuren B, Lee JE, Marshall DJ, Convey P, Chown SL. 2011. Mite dispersal among the Southern Ocean Islands and Antarctica before the last glacial maximum. *Proceedings of the Royal Society B: Biological Sciences* 278: 1247–1255.
- Mortimer E, McGeoch MA, Daniels SR, Van Vuuren BJ. 2008. Growth form and population genetic structure of *A. selago* on sub-Antarctic Marion Island. *Antarctic Science* 20: 381–390.
- Moseley HN. 1879. On the structure of the Stylas-Teridae a family of hydroid stony corals. *Nature* 20: 339–341.
- Muñoz J, Felicísimo ÁM, Cabezas F, Burgaz AR, Martínez I. 2004. Wind as a long-distance dispersal vehicle in the Southern Hemisphere. *Science* 304: 1144–1147.
- Murtagh F, Legendre P. 2014. Ward's hierarchical agglomerative clustering method: Which algorithms implement Ward's criterion? *Journal of Classification* 31: 274–295.
- Nel W, Hedding DW, Rudolph EM. 2023. The sub-Antarctic islands are increasingly warming in the 21st century. *Antarctic Science* 35: 124–126.
- Nicolas AN, Plunkett GM. 2012. Untangling generic limits in *Azorella*, *Laretia* and *Mulinum* (Apiaceae: Azorelloideae): Insights from phylogenetics and biogeography. *Taxon* 61: 826–840.
- Paradis E. 2018. Analysis of haplotype networks: The randomized minimum spanning tree method. *Methods in Ecology and Evolution* 9: 1308–1317.
- Paradis E. 2010. pegas: An R package for population genetics with an integrated-dmodular approach. *Bioinformatics* 26: 419–420.
- Patterson SL. 1985. Surface circulation and kinetic energy distributions in the Southern Hemisphere oceans from FGGE drifting buoys. *Journal of Physical Oceanography* 15: 865–884.
- Perrie L, Parris B. 2012. Chloroplast DNA sequences indicate the gammitid ferns (Polypodiaceae) in New Zealand belong to a single clade, *Notogrammitis* gen. nov. *New Zealand Journal of Botany* 50: 457–472.
- Pfuhl H, McCave IN. 2005. Evidence for late Oligocene establishment of the Antarctic Circumpolar Current. *Earth and Planetary Science Letters* 235: 715–728.
- Posada D. 2009. Selection of models of DNA evolution with jModelTest. In: Posada D ed. *Bioinformatics for DNA sequence analysis*. USA: Humana Press. 537: 93–112.

- PPG I. 2016. A community-derived classification for extant lycophytes and ferns. *Journal of Systematics and Evolution* 54: 563–603.
- Quilty P. 2007. Origin and evolution of the sub-Antarctic islands: The foundation. *Papers and Proceedings of the Royal Society of Tasmania* 141: 35–58.
- R Core Team. 2022. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing.
- Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA. 2018. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology* 67: 901–904.
- Ramos VA, Cingolani C, Junior FC, Naipauer M, Rapalini A. 2017. The Malvinas (Falkland) Islands revisited: The tectonic evolution of southern Gondwana based on U-Pb and Lu-Hf detrital zircon isotopes in the Paleozoic cover. *Journal of South American Earth Sciences* 76: 320–345.
- Ranker TA, Smith AR, Parris BS, Geiger JMO, Haufler CH, Straub SCK, Schneider H. 2004. Phylogeny and evolution of Grammitid Ferns (Grammitidaceae): A case of rampant morphological homoplasy. *Taxon* 53: 415–428.
- Ree RH, Smith SA. 2008. Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. *Systematic Biology* 57: 4–14.
- Richardson JE, Fay MF, Cronk QC, Chase MW. 2003. Species delimitation and the origin of populations in island representatives of *Phyllica* (Rhamnaceae). *Evolution* 57: 816–827.
- Rincón-Barrado M, Villaverde T, Perez MF, Sanmartín I, Riina R. 2024. The sweet tabaiba or there and back again: Phylogeographical history of the Macaronesian *Euphorbia balsamifera*. *Annals of Botany* 135: 883–904.
- Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. 2012. MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology* 61: 539–542.
- Sanmartín I. 2012. Historical biogeography: Evolution in time and space. *Evolution: Education and Outreach* 5: 555–568.
- Sanmartín I. 2022. Analytical approaches in biogeography: Advances and challenges. In: Guilbert E ed. *Biogeography: An integrative approach of the evolution of living*. Hoboken, NJ: Wiley. 27–58.
- Sanmartín I, Anderson CL, Alarcon M, Ronquist F, Aldasoro JJ. 2010. Bayesian island biogeography in a continental setting: The Rand Flora case. *Biology Letters* 6: 703–707.
- Sanmartín I, Ronquist F. 2004. Southern hemisphere biogeography inferred by event-based models: Plant versus animal Patterns. *Systematic Biology* 53: 216–243.
- Sanmartín I, Van Der Mark P, Ronquist F. 2008. Inferring dispersal a Bayesian approach to phylogeny-based island. *Journal of Biogeography* 35: 428–449.
- Sanmartín I, Wanntorp L, Winkworth RC. 2007. West Wind Drift revisited: Testing for directional dispersal in the Southern Hemisphere using event-based tree fitting. *Journal of Biogeography* 34: 398–416.
- Sarkar S, Basak C, Frank M, Berndt C, Huuse M, Badhani S, Bialas J. 2019. Late Eocene onset of the proto-Antarctic circumpolar current. *Scientific Reports* 9: 10125.
- Schrader J, Weigelt J, Cai L, Westoby M, Fernández-Palacios JM, Cabezas FJ, Plunkett GM, Ranker TA, Triantis KA, Trigas P, Kubota Y, Kreft H. 2024. Islands are key for protecting the world's plant endemism. *Nature* 634: 1–7.
- Shaw J, Lickey EB, Beck JT, Farmer SB, Liu W, Miller J, Siripun KC, Winder CT, Schilling EE, Small RL. 2005. The tortoise and the hare II: Relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. *American Journal of Botany* 92: 142–166.
- Smith VR. 2002. Climate change in the sub-Antarctic: An illustration from Marion Island. *Climatic Change* 52: 345–357.
- Smith TM, York PH, Broitman BR, Thiel M, Hays GC, van Sebillie E, Putman NF, Macreadie PI, Sherman CDH. 2018. Rare long-distance dispersal of a marine angiosperm across the Pacific Ocean. *Global Ecology and Biogeography* 27: 487–496.
- Stokes S. 1988. R13/f0171. FRED. The Fossil Record Electronic Database. Available from <https://fred.org.nz/index.jsp> [accessed October 2024].
- Suchard MA, Lemey P, Baele G, Ayres DL, Drummond AJ, Rambaut A. 2018. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evolution* 4: vey016.
- Sundue MA, Parris BS, Ranker TA, Smith AR, Fujimoto EL, Zamora-Crosby D, Morden CW, Chiou WL, Chen CW, Rouhan G, Hirai RY, Prado J. 2014. Global phylogeny and biogeography of grammitid ferns (Polypodiaceae). *Molecular Phylogenetics and Evolution* 81: 195–206.
- Swenson U, Backlund A, McLoughlin S, Hill RS. 2001. Nothofagus biogeography revisited with special emphasis on the enigmatic distribution of subgenus *Brassospora* in New Caledonia. *Cladistics* 17: 28–47.
- Testo W, Field A, Barrington D. 2018. Overcoming among-lineage rate heterogeneity to infer the divergence times and biogeography of the clubmoss family Lycopodiaceae. *Journal of Biogeography* 45: 1929–1941.
- Tribble CM, Freyman WA, Landis MJ, Lim JY, Barido-Sottani J, Kopperud BT, Höhna S, May MR. 2022. RevGadgets: An R package for visualizing Bayesian phylogenetic analyses from RevBayes. *Methods in Ecology and Evolution* 13: 314–323.
- Trewick SA, Morgan-Richards M, Russell SJ, Henderson S, Rumsey FJ, Pintér I, Barrett JA, Gibby M, Vogel JC. 2002. Polyploidy, phylogeography and Pleistocene refugia of the rockfern *Asplenium ceterach*: Evidence from chloroplast DNA. *Molecular Ecology* 11: 2003–2012.
- Van de Peer Y, Mizrahi E, Marchal K. 2017. The evolutionary significance of polyploidy. *Nature Reviews Genetics* 18: 411–424.
- Van der Putten N, Verbruggen C, Ochrya R, Verleyen E, Frenot Y. 2010. Subantarctic flowering plants: Pre-glacial survivors or post-glacial immigrants? *Journal of Biogeography* 37: 582–592.
- Van Vuuren BJ, Lee JE, Convey P, Chown SL. 2018. Conservation implications of spatial genetic structure in two species of oribatid mites from the Antarctic Peninsula and the Scotia Arc. *Antarctic Science* 30: 105–114.
- Wagstaff SJ, Breitwieser I, Ito M. 2011. Evolution and biogeography of *Pleurophyllum* (Astereae, Asteraceae), a small genus of megaherbs endemic to the subantarctic islands. *American Journal of Botany* 98: 62–75.
- Walton DW. 1979. Studies on *Acaena* (Rosaceae): III. Flowering and hybridization on South Georgia. *British Antarctic Survey Bulletin* 48: 1–13.
- Walton DWH. 1982. Floral phenology in the South Georgian vascular flora. *British Antarctic Survey Bulletin* 55: 11–25.
- Walton DWH, Greene SW. 1971. The South Georgian species of *Acaena* and their probable hybrid. *British Antarctic Survey Bulletin* 23: 29–44.
- Waters JM. 2008. Driven by the West Wind Drift? A synthesis of southern temperate marine biogeography, with new directions for dispersalism. *Journal of Biogeography* 35: 417–427.
- Weimerskirch H, Jouventin P, Mougín JL, Stahl JC, Beveren MV. 1985. Banding recoveries and the dispersal of seabirds breeding in

- French Austral and Antarctic territories. *Emu—Austral Ornithology* 85: 22–33.
- Whittaker RJ, Fernández-Palacios JM, Matthews TJ, Borregaard MK, Triantis KA. 2017. Island biogeography: Taking the long view of nature's laboratories. *Science* 357: eaam8326.
- Wickham H. 2006. An introduction to ggplot: An implementation of the grammar of graphics in R. *Statistics* 1–8. <https://ggplot2.tidyverse.org>
- Winkworth RC, Hennion F, Prinzing A, Wagstaff SJ. 2015. Explaining the disjunct distributions of austral plants: The roles of Antarctic and direct dispersal routes. *Journal of Biogeography* 42: 1197–1209.
- Winkworth RC, Wagstaff SJ, Glenny D, Lockhart PJ. 2002. Plant dispersal news from New Zealand. *Trends in Ecology & Evolution* 17: 514–520.
- Wolfe KH, Li WH, Sharp PM. 1987. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proceedings of the National Academy of Sciences USA* 84: 9054–9058.
- Xu J, Pérez-Losada M, Jara CG, Crandall KA. 2009. Pleistocene glaciation leaves deep signature on the freshwater crab *Aegla alacalufi* in Chilean Patagonia. *Molecular Ecology* 18: 904–918.
- Zhang S, Jin J, Chen S, Chase MW, Soltis DE, Li H, Yang J, Li D, Yi T. 2017. Diversification of Rosaceae since the Late Cretaceous based on plastid phylogenomics. *New Phytologist* 214: 1355–1367.

Supplementary Material

The following supplementary material is available online for this article at <http://onlinelibrary.wiley.com/doi/10.1111/jse.13170/supinfo>:

Appendix S1. Sample information: GenBank accession numbers, geographical coordinates, voucher information, and summary statistics for each sample's DNA alignments.

Fig. S1. Maximum Clade Credibility (MCC) tree of *Acaena magellanica* showing the phylogenetic relationships based on the combined sequences using Bayesian Inference.

Fig. S2. Maximum Clade Credibility (MCC) tree of *Austroblechnum penna-marina* showing the phylogenetic relationships based on the combined sequences using Bayesian inference.

Fig. S3. Maximum Clade Credibility (MCC) tree of *Azorella selago* showing the phylogenetic relationships based on the combined sequences using Bayesian Inference.

Fig. S4. Maximum Clade Credibility (MCC) tree of *Colobanthus kerguelensis* showing the phylogenetic relationships based on the combined sequences using Bayesian Inference.

Fig. S5. Maximum Clade Credibility (MCC) tree of *Notoگرامmitis crassior* showing the phylogenetic relationships based on the combined sequences using Bayesian Inference.

Fig. S6. Maximum Clade Credibility (MCC) tree of *Polystichum marionense* showing the phylogenetic relationships based on the combined sequences using Bayesian Inference.

Fig. S7. Maximum Clade Credibility (MCC) tree of *Pringlea antiscorbutica* showing the phylogenetic relationships based on the combined sequences using Bayesian Inference.

Fig. S8. Maximum Clade Credibility (MCC) tree showing the divergence ages of *Acaena magellanica*. The orange stars represent the secondary calibration point and green stars represent the fossil calibration points. The purple bars in

nodes show the confidence interval 95% HPD (High Posterior Density). Crozet Islands, C; Falkland Islands, F; Heard Island, H; Kerguelen Archipelago, K; Macquarie Island, MA; Marion Island, MI; Prince Edward Island, PE; South America, SA; and South Georgia, SG.

Fig. S9. Maximum Clade Credibility (MCC) tree showing the divergence ages of *Austroblechnum penna-marina*. The orange stars represent the calibration point. The purple bars in nodes show the confidence interval 95% HPD (High Posterior Density). Australia, AU; Marion Island, MI; New Zealand, NZ; and South America, SA.

Fig. S10. Maximum Clade Credibility (MCC) tree showing the divergence ages of *Azorella selago*. The orange stars represent the calibration point and the purple bars in nodes. The purple bars in nodes show the confidence interval 95% HPD (High Posterior Density). Crozet Islands, C; Falkland Islands, F; Heard Island, H; Kerguelen Archipelago, K; Macquarie Island, MA; Marion Island, MI; Prince Edward Island, PE; and South America, SA.

Fig. S11. Maximum Clade Credibility (MCC) tree showing the divergence ages of *Colobanthus kerguelensis*. The orange stars represent the calibration point. The purple bars in nodes show the confidence interval 95% HPD (High Posterior Density). Amsterdam Island: A; Crozet Islands, C; Kerguelen Archipelago, K; and Prince Edward Island, PE.

Fig. S12. Maximum Clade Credibility (MCC) tree showing the divergence ages of *Notoگرامmitis crassior*. The orange stars represent the calibration point. The purple bars in nodes show the confidence interval 95% HPD (High Posterior Density). Australia, AU; Kerguelen Archipelago, K; Marion Island, MI; New Zealand, NZ; and South America, SA.

Fig. S13. Maximum Clade Credibility (MCC) tree showing the divergence ages of *Polystichum marionense*. The orange stars represent the calibration point. The purple bars in nodes show the confidence interval 95% HPD (High Posterior Density). MI: Marion Island.

Fig. S14. Maximum Clade Credibility (MCC) tree showing the divergence ages of *Pringlea antiscorbutica*. The orange stars represent the calibration point. The purple bars in nodes show the confidence interval 95% HPD (High Posterior Density). Amsterdam Island: A; Antipodes Island, AN; Auckland Island, AK; Australia, AU; Campbell Island, CA; Crozet Islands, C; Falkland Islands, F; Heard Island, H; Kerguelen Archipelago, K; Macquarie Island, MA; Marion Island, MI; New Zealand, NZ; Prince Edward Island, PE; South America, SA; South Georgia, SG; and Tristan da Cunha Islands Group, T.

Fig. S15. Graphic showing carrying capacity analysis of the sub-Antarctic areas. We identified four areas: South America, Gondwanan fragments (Falkland Islands and South Georgia), the Eastern Austral Archipelagos (SIOBP and Macquarie), and Australia/New Zealand.

Fig. S16. Graphic showing the dispersal rates between the identified sub-Antarctic areas in Fig. 1. SA: South America; GF: Gondwanan Fragments; EAA: Eastern Austral Archipelagos; AU: Australia; and NZ: New Zealand.

Fig. S17. Stochastic mapping showing the posterior probability of the changes between the ancestral states of *Acaena magellanica*. The posterior probability is represented by a gradient between the maximum in red (1) and the minimum in yellow (0).

Fig. S18. Stochastic mapping showing the posterior probability of the changes between the ancestral states of *Azorella selago*. The posterior probability is represented by a gradient between the maximum in red (1) and the minimum in yellow (0).

Fig. S19. Stochastic mapping showing the posterior probability of the changes between the ancestral states of *Pringlea antiscorbutica*. The posterior probability is represented by a gradient between the maximum in red (1) and the minimum in yellow (0).

Fig. S20. Hierarchical cluster of haplotypes for *Acaena magellanica* (Lam.) Vahl showing the mutational steps between the different haplotypes found in the sub-Antarctic region. Abbreviations: FA, Falkland; MA, Macquarie; SA, South America; SIOBP, Southern Indian Ocean Biogeographic Province; SG, South Georgia.

Table S1. Biological information about the seven target plant species. Abbreviations: a, anemochory; c, coastal; cc, cinder cones; fl, fellfield; fn, fernbrake; h, hydrochory; m, mire; p, polar desert; s, springs; z, zoochory.

Table S2. Extraction information of the DNA samples per each species showing the concentration (ng/ μ L), the purity (A280/A260), and the amplified molecular markers.

Table S3. PCR conditions used in this study. The rules described in the protocol of the commercial kit of extraction (nzytech, Portugal) were applied using the following

reactives: 2.5 μ L of forward primer forward (5 μ M), 2.5 μ L of reverse primer (5 μ M), 2 μ L of DNA sample, 5 μ L of Buffer Reaction, 2.5 μ L of MgCl₂ (50 mM), 2.5 μ L of dNTPs (10 mM, 1 mL per each nucleotide), 0.5 μ L of DNA-polymerase (Supreme NZYTaQ II, 5U/ μ L), and 32.5 μ L of distilled water.

Table S4. Sequences of primers used for the seven target taxa.

Table S5. Summary of the basic information on the seven taxa: alignment length (in bp), percentage of identical sites (%), number of sequences, and GC content (%).

Table S6. Summary of the likelihood of the Stepping-Stone comparisons in BEAST software. Bold format represents the selected models. UCLN: Uncorrelated clock; SC: Strict clock; BD: Birth-Death prior; and COAL: Coalescence prior.

Table S7. Summary of the calibration points used for the molecular dating. Fossils are marked with an asterisk. My: millions of years; STDE: standard deviation.

Table S8. Results of the simulations performed for exploring the mean number of additional links found in the RMST algorithm (N: number of sequences; NUD: mean number of unique distances).

Table S9. List of haplotypes and associated genetic statistics in *A. magellanica*. Abbreviations: H(n), number of haplotypes; SIOBP, Southern Indian Ocean Biogeographic Province.