TRAIT-DEPENDENT DIVERSIFICATION AND BIOGEOGRAPHY OF CANARY ISLANDS ANGIOSPERMS

Ryan Brewer



Propositions

Belonging to the PhD dissertation

TRAIT-DEPENDENT DIVERSIFICATION AND BIOGEOGRAPHY OF CANARY ISLANDS ANGIOSPERMS

1. Unevenness in species richness between lineages in the Canary Islands combines multiple aspects of the time-for-speciation, diversification rates, and ecological limits hypotheses (This Thesis).
2. Contrasting global studies, there appears to be no relationship between clade age and diversity in the Canary Islands (This Thesis, Chapter Two).
3. Insular woodiness acts as a key innovation, shaping Canary Islands flora (This Thesis, Chapters Two, Three).
4. Topographic complexity is universally important in delineating plant occurrence in the Canary Islands (This Thesis, Chapter Three).
5. The high prevalence of insular woody species in the Canaries results from a change in climate regime and amply available herbaceous colonisers (This Thesis, Synthesis).
6. The 'Zarathustran hypothesis' of Quintero et al. (2025) aptly applies to the Canaries.
7. Monocots do not matter when considering insular woodiness.
8. Some plants work harder than others to make sure they are not understood.
9. "When we cannot formulate a problem with precision, it is often not because the problem is profound: it's because the problem is false", Carlo Rovelli, The Order of Time (2017).

Ryan Brewer

Groningen, 13 October 2026

TRAIT-DEPENDENT DIVERSIFICATION AND BIOGEOGRAPHY OF CANARY ISLANDS ANGIOSPERMS

Ryan Brewer

ISBN: 978-94-6534-599-4
NUR: 949

Cover design: Ron Zijlmans
Design of the plant images: Erik-Jan Bosch, Naturalis Illustrator
Lay-out: RON Graphic Power | www.ron.nu

© Copyright 2026: Ryan Brewer

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, by photocopying, recording, or otherwise, without the prior written permission of the author.

university of groningen

Trait-dependent diversification and biogeography of Canary Islands angiosperms

PhD thesis

to obtain the degree of PhD at the
University of Groningen
on the authority of the
Rector Magnificus Prof. J.M.A. Scherpen
and in accordance with
the decision by the College of Deans.

This thesis will be defended in public on

Tuesday 13 October 2026 at 14.30 hours

by

Ryan Frederick Arthur Brewer

born on 30 June 1996

Supervisors

Prof. R.S. Etienne
Prof. F. Lens
Prof. L.M. Lima Valente

Assessment Committee

Prof. C.J.M. Egas
Prof. B. Gravendeel
Prof. C. Beierkuhnlein

TABLE OF CONTENTS

Chapter 1 Introduction 7

Chapter 2 Ecological limits and trait-dependent diversification drive angiosperm diversity unevenness across the Canary Islands 29

Chapter 3 Climatic seasonality and topographic complexity shape Canary Islands plant growth form distributions 73

Chapter 4 A Model for Trait-Dependent Colonisation and Diversification on Islands 101

Chapter 5 Synthesis 145

Summary 181

Nederlandse samenvatting 187

Acknowledgements 193

Curriculum Vitae 199



Chapter 1

Introduction



Angiosperms and Diversity Unevenness

Land plants are the cornerstone of terrestrial life on Earth. They are the backbone of terrestrial ecosystems (Negrutiu et al. 2020), generate the oxygen we breathe, and provide over 50% of direct human energy intake (Proches et al. 2008). There are currently over 350,000 recognised extant species of land plants, comprising 10 major clades (phyla sensu Niklas 2016): Marchantiophyta (liverworts), Bryophyta (mosses), Anthocerotophyta (hornworts), Lycopodiophyta (clubmosses and quillworts), Polypodiophyta (ferns), Ginkgophyta (*Ginkgo*), Cycadophyta (cycads), Gnetophyta (*Gnetum*, *Welwitschia,* and *Ephedra*), Pinophyta (conifers), and Magnoliophyta (flowering plants/angiosperms). Extant species diversity within these major clades is extremely unevenly distributed, varying from one extant Ginkgophyta species to over 330,000 angiosperms (Hernandez-Hernandez and Wiens 2020; Zuntini et al. 2024). Understanding the underlying reasons to explain this unevenness at the heart of plant diversity across spatial and temporal scales is one of the most important endeavours in (macro)ecology and evolution (Antonelli et al. 2023).

There are three main (non-mutually exclusive) hypotheses to explain the uneven distribution of biodiversity across the tree of life (Valente and Vargas 2013; Etienne et al. 2019). First, the clade age hypothesis stipulates that time is the sole explanatory factor: the older the clade, the more species it will contain due to having more time to accumulate species (McPeek and Brown 2007; Henao Diaz et al. 2019; Yu and Wiens 2024). Second, the diversification rate hypothesis emphasises that clades accumulate species at different rates as a result of morphological, anatomical, physiological, environmental, and/or competitive traits (Gotelli et al. 2009; Wiens 2017). Interestingly, these two hypotheses appear correlated, with evolutionary rates scaling with time (i.e., younger clades having higher evolutionary rates; Henao Diaz et al. 2019; Harmon et al. 2021). The final hypothesis, variable ecological limits on diversity over time, emphasises limits to diversity within a given system based on properties of the system itself (e.g., resource availability or carrying capacity of the system); for this reason, it is sometimes considered as a subset of the diversification rate hypothesis (Valente and Vargas 2013; Rabosky and Hurlbert 2015; Wiens 2017).

To further complicate matters, the mechanisms controlling evolutionary rates have been found to vary between clades. Within some clades, diversity-dependent diversification, the effect of interspecific competition on macroevolutionary rates (Rabosky 2013), causes evolutionary rates to slow down in the presence of many closely related (e.g., congeneric) species. This is often cited as evidence of ecological limits to diversity (Etienne et al. 2016). In contrast, diversity-independent diversification, where rates are unregulated by interactions with other species (Rabosky 2013), has recently been found across

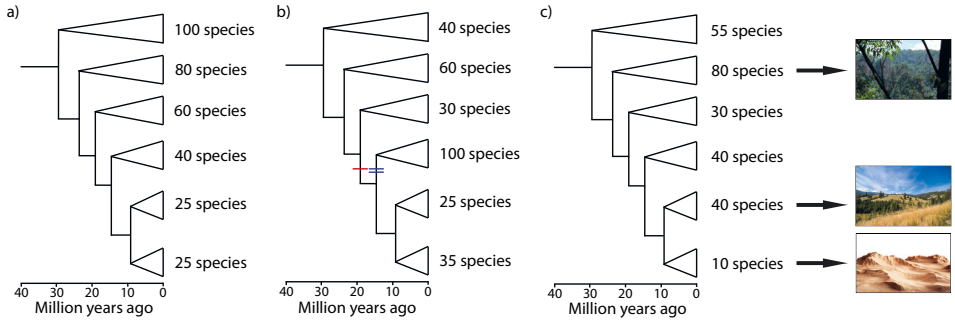


Figure 1. Simplified reconstructions of the three main hypotheses to explain why biodiversity is unevenly distributed across the tree of life. a) Clade age, where diversity is time-dependent. Here, older clades have more species. b) Diversification rate, where the evolution of traits that promote diversification is highlighted in colour. The red bar represents the evolution of a single trait that promotes diversification. The blue bars denote that traits often act synergistically to affect diversification rates, and the evolution of both of these traits together promotes diversification (i.e., a key synnovation; Donoghue and Sanderson 2015). c) Variable ecological limits, where the number of species within a clade is constrained by the area in which the clade lives (the carrying capacity), based on resource availability and/or environmental factors. Three biomes are used as examples: a species-rich tropical rainforest, a steppe system with moderate diversity, and a depauperate desert system. Biome images are sourced from Freepik.

the tree of life (Pie et al. 2023). Given the sheer scale of the variation in diversity across the tree of life, it is perhaps unsurprising that existing studies cannot reach a consensus on the drivers of biodiversity and often arrive at disparate conclusions (see Yu and Wiens 2024 for a summary).

Angiosperms are the youngest clade of land plants, having evolved between 140 and 270 million years ago (mya; Sauquet et al. 2022; Smith and Beaulieu 2024), yet they harbour over 90% of extant land plant species (Zuntini et al. 2024). The drivers of their vast richness in species and forms have been of scientific interest for over 150 years (Sauquet and Magallon 2018). Clearly, angiosperms have a high diversification rate (speciation minus extinction rate) to sustain such a high number of species. Yet, one of the key unresolved questions in plant evolutionary science is identifying which traits promote speciation and extinction (Patiño et al. 2017; Sauquet and Magallon 2018; Hooft van Huysduynen et al. 2021).

Trait-Dependent Diversification in Angiosperms: Patterns and Modelling

Angiosperms are an ideal system for studying trait-dependent diversification (diversification driven by species characteristics) due to their well-resolved phylogenetic backbone (Sauquet and Magallon 2018; Zuntini et al. 2024), their high species diversity, and their extensive use in previous studies (Magallon et al. 2015; Silvestro et al. 2015; Tank et al. 2015; Sauquet et al. 2017; Hernán-

dez-Hernández and Wiens 2020). Despite this, it is only in recent years that leaps in our understanding of the drivers of diversification in angiosperms across taxonomic and spatial scales have been made (Tietje et al. 2022; Helmstetter et al. 2023; Poppenwimer et al. 2023; Perez-Escobar et al. 2024).

Globally, the effects of traits on angiosperm diversification appear idiosyncratic across taxonomic groups, with universal sets of traits affecting diversification not yet uncovered, if they exist (Helmstetter et al. 2023). For example, the effects of flower and fruit morphology (Sabath et al. 2016; Fernandez-Mazueco et al. 2017; Gavrutenko et al. 2020; Xue et al. 2020), and life form have inconsistent effects on diversification (Bouchenak-Khelladi et al 2014; Xue et al. 2020; Koutroumpa et al. 2021; Sperotto et al. 2023). Yet, generalisations can be drawn (Vamosi et al. 2018; Schenk 2021): biotic fertilisation (Givnish 2010; Hernández-Hernández and Wiens 2020), polyploidy (Schenk 2021; but see Mayrose et al. 2011), hybridisation (Curto et al. 2018; White et al. 2020), and the heterogeneity of the clade's environment all appear to promote diversification in angiosperms (Valente et al. 2014; Hernández-Hernández and Wiens 2020; Nurk et al. 2020; Schenk 2021). In addition, rates of genome size evolution (Puttick et al. 2015; Bhadra et al. 2023), dispersal mode and ability (Givnish 1998), and reproductive strategy have all been correlated with elevated rates of net diversification (Givnish 2010). Co-evolution with other classes of organisms has also spurred diversification bilaterally (Strutzenberger and Fiedler 2011; Benton et al. 2022; Blaimer et al. 2023; Peris and Condamine 2024). Nevertheless, it is unlikely that a single trait promotes diversification across all clades (Nürk et al. 2020; Hooft van Huysduynen et al. 2021), yet most existing models to detect trait-dependent diversification cannot capture this biological complexity as they only allow for a single trait to be modelled at a time.

In order to study trait-dependent diversification, a series of models has been developed, based on State-dependent Speciation and Extinction (SSE) models. These include Binary-State SSE (BiSSE; Maddison et al. 2007), a multistate extension of BiSSE (MuSSE; Fitzjohn 2012), and Cladogenetic-State change SSE (ClaSSE; Goldberg and Igic 2012). This series of models works by assigning probabilities of each trait state given the distribution of observed species traits at the tips, which then iterates rootwards, merging probabilities at the nodes, until the root is reached. Then, at the root, the probability of the tree, the rates of speciation and extinction (which can be fixed to zero due to difficulties estimating extinction rates from phylogenies), and trait evolution rates are given. The probability is expressed as likelihood, which can then be used to identify the parameter combination that best fits the data. More recent models overcome some shortcomings of these early SSE models (such as high type I error rates; see Maddison and Fitzjohn 2015 and Rabosky and Goldberg 2015)

by incorporating uncertainty in trait information using a concealed trait. By allowing concealed traits to influence diversification rates, the Several Examined and Concealed SSE model (SecSSE; Herrera-Alsina et al. 2019) avoids the high type I error rates without sacrificing statistical power and allows for multiple traits to be modelled (Beaulieu and O'Meara 2016). Yet, SSE models require densely sampled phylogenetic trees (Maddison et al. 2007; Mynard et al. 2023), and mostly only apply to a single clade (Goldberg and Igic 2012). The requirement of sampling (nearly) all species is often impossible for large clades, due to time, budget and sampling constraints (Sauquet and Magallon 2018; Dehayem Nanwou et al. 2025), impeding studies of trait-dependent diversification in interesting clades and ecological communities that have undergone substantial speciation. Island systems, by virtue of their smaller size and simplified floras relative to continental systems (Losos and Ricklefs 2009), offer more tractable starting points for studying trait-dependent diversification.

Island Biogeography

As natural laboratories of biogeography and evolution (Losos et al. 2010), islands have been of scientific interest since the voyages of Darwin and Wallace (Darwin 1859; Wallace 1869). Their isolation, strict boundaries, and relatively simple but unique biotas compared to continental systems simplify studies on the complex processes of evolution (Losos and Ricklefs 2009). Yet, it was not until the landmark publications of MacArthur and Wilson (1963, 1967) that a theoretical framework of how island properties (size and isolation) interact with evolutionary rates to predict insular species richness was first presented. Entitled the *Theory of Island Biogeography* (TIB), a key tenet of this theory states that insular diversity exists in dynamic equilibrium, where community species richness is maintained via speciation, extinction, and island colonisation rates. Since the publication of the TIB, island biogeography theory has been extended to include how geology and island ontogeny affect speciation, extinction, and colonisation rates across oceanic islands (the general dynamic model, hereafter 'GDM'; Whittaker et al. 2008). Whilst the predictions and assumptions of the TIB have been tested and further developed (Simberloff and Wilson 1969; Lomolino 2000; Rosindell and Phillimore 2011; Triantis et al. 2012; Valente et al. 2020), studies that evaluate the GDM are still in their infancy, but those that exist generally support its conclusions (Valente et al. 2020).

Valente et al. (2020) used the dynamic stochastic model DAISIE (Dynamic Assembly of Island biota through Speciation, Immigration, and Extinction) to estimate rates of colonisation, speciation through anagenesis and cladogenesis, and extinction, with maximum-likelihood methods (Valente et al. 2015).

Using a phylogenetic dataset of 596 terrestrial bird species across 41 oceanic archipelagos worldwide, Valente et al. (2020) tested the effect of island features (area and isolation) on rates of colonisation, speciation, and extinction. Their findings confirm two predictions of MacArthur and Wilson (1967): (1) colonisation rates decline with increasing isolation and (2) extinction rates decline with increasing area. They also further developed the theory presented by MacArthur and Wilson (1967), confirming theory proposed in recent decades that cladogenesis increases through an interactive effect of island size and isolation (Rosindell and Phillimore 2011), and that anagenesis increases with island isolation (Whittaker et al. 2008). However, the inference module of DAISIE does not include the effect of traits on diversification rates and, therefore, is currently not suitable to study trait evolution on islands.

Integrating character traits into diversification models has been identified as the next step in diversification modelling and island biogeography theory (Sauquet and Magallon 2018; Triantis and Matthews 2020; Schrader et al. 2021; Schrader 2025). As the GDM is neutral to traits (Gillespie and Whittaker 2025), a key assumption is that all species have equal opportunities to colonise an island system (meaning that island communities are random subsets of the continental systems from which the island colonists originate). However, the emerging field of 'functional island biogeography' contends the opposite: that functional traits affect island biogeographic predictions (Schrader et al. 2021; Schrader 2025). For example, if traits affected colonisation or extinction rates, we would expect overrepresentation of traits that enhance dispersal ability and that allow for increased persistence (Schrader 2025). The species and trait-abundance in the island community would, therefore, relative to the source of the island community, be different - a globally observed phenomenon known as island disharmony (Ottaviani et al. 2020; Koenig et al. 2021). In order to jointly test functional island biogeography and trait-dependent diversification hypotheses, we need a method that combines mechanisms from the two. In this context, DAISIE and SecSSE are prime candidates for such an integration between island biogeography and trait-dependent diversification, as aspects inherent to the functionality of these models allow their major limitations to be overcome. On the one hand, the sampling requirements of DAISIE are lower than SSE-based trait-dependent diversification models because sampling is focused on the island community plus close continental relatives, rather than the entire phylogeny, and the DAISIE model is specifically designed for islands and can estimate colonisation rates. On the other hand, SecSSE can capture the effects of traits, including a single trait, multiple traits, or those which are not sampled, on evolutionary dynamics. Merging these models would, therefore, capture the impact of multiple-trait-depend-

ent diversification across island assemblages with only island representatives and their continental relatives sampled, providing novel insights into the traits driving diversification and colonisation in an insular context.

Linking Trait-based Macroecology and Diversification Dynamics across Clades

Implementing approaches that accommodate trait evolution and trait-dependent diversification, alone or in conjunction with taxonomic and/or phylogenetic studies, into island biogeography would contribute to addressing roughly two-thirds of the fundamental questions in island biogeography (Patiño et al. 2017), including those ranked as most important (Ottaviani et al. 2020). However, given the current restrictions of diversification models outlined above (only a single focus clade with preferably many (>300) species, near-perfect sampling fraction; Davis et al. 2013; Mynard et al. 2023), coupled with the extremely complex nature of the underlying mechanisms behind trait-based diversification (Anderson et al. 2023), understanding how traits shape diversification remains challenging.

Common plant traits can have opposing effects on diversification, acting to promote diversification in some instances, whilst decreasing diversification in others (Anderson et al. 2023). For example, a smaller maximum plant size could allow for diversification into more niches via niche packing across fine-grained environmental heterogeneity, promoting diversification via allopatry (Boucher et al. 2017); equally, however, a smaller size can reduce competitive and persistence abilities in mature communities (Falster and Westoby 2003), leading to evolutionary stasis (i.e., lower diversification). Current models do not allow for the detection of these differential effects of traits on diversification across clades, or for diversity-(in)dependent diversification, and suggested ways of accounting for these differences (replicated small-scale analyses across independent clades; Anderson et al. 2023), cannot be conducted using existing models. It may be for this reason that so far no trait or set of traits has been found to universally affect diversification (Helmstetter et al. 2023).

In contrast, convergent evolutionary responses in distantly related clades indicate general principles in how plants diversify in space and time (Donoghue et al. 2022; Vasconcelos 2023). For example, annuality in plants dominates in regions with extremely high temperatures and prolonged periods of drought during the hottest quarter (Poppenwimer et al. 2023), with annuality evolving convergently across clades as a heat avoidance mechanism (Boyko et al. 2023). In contrast, perennials dominate at tropical latitudes, strongly driven by the combination of low seasonality, high precipitation, and high temperature (Taylor et al. 2023), suggesting biogeographical drivers of growth forms. Though

generalisable principles in evolution do not necessarily imply changes in diversification rate (but could be attributed to transition rates, for example), they could be evidence that there are broadly generalisable patterns in the drivers of diversification. A particularly prevalent convergent response on islands is herbaceous species repeatedly evolving woodiness following islands' colonisation, an evolutionary phenomenon known as insular woodiness (IW; Lens et al. 2013; Zizka et al. 2022).

Many evolutionary radiations, the dramatic proliferation of taxa within a clade (Simões et al. 2016), on various island systems, tend to feature insular woody species (IWS). These include the Hawaiian silversword alliance (three genera [*Argyroxiphium, Dubautia,* and *Wilkesia*] and 33 species; Baldwin et al. 1998; Carlquist et al. 2003) and woody violets (*Viola*; 10 species; Ballard and Systema 2000), and the Macaronesian *Aeonium* Alliance (three genera [*Aeonium, Aichryson,* and *Monanthes*] and 63 species; Beierkuhnlein et al. 2021; Brilhante et al. 2021) and *Echium* (31 species; Böhle et al. 1996). Though this relationship has led to several authors noting a potential link between the evolution of IW and the promotion of diversification rates (Carine et al. 2010; Lens et al. 2013; Nürk et al. 2019) - insomuch as calling the link 'obvious' for the Canary Islands (Hooft van Huysduynen et al. 2021) - empirical evidence to postulate insular woodiness as a key trait that drives diversification of island clades is equivocal; either without quantitative analysis (Carine et al. 2010; Lens et al. 2013; Hooft van Huysduynen et al. 2021), or restricted to individual clades (Koutroumpa et al. 2021). However, a global analysis found that ~50% of IW clades only contain ≤2 species (Zizka et al. 2022), suggesting that IW does not necessitate elevated diversification rates. Therefore, an empirical analysis on whether IW affects diversification rates across an insular community and at larger taxonomic scales is hitherto lacking.

Several drivers have been proposed to explain the evolution of IW: competition for light, driving selection for height and subsequently stem support, thereby woodiness (Darwin 1859; Givnish 1998); reduced herbivory on islands (selecting for longevity; Carlquist 1974); burial avoidance following volcanic deposits (Beierkuhnlein et al. 2023); and/or drought resistance (Doria et al. 2018; Zizka et al. 2022). Or perhaps the shift from herbaceousness to IW could represent a shift in the large-scale biogeographical drivers of growth forms (Taylor et al. 2023), away from extreme heat towards more climatically stable, aseasonal (i.e., less extreme, frost-free climate), but still arid conditions afforded by islands due to buffering by surrounding oceans (Carlquist 1974), detectable in niche space. Insular woodiness contrasts ancestral woodiness, which has been retained through evolutionary time from the ancestor of angiosperms (Gerrienne et al. 2011; Doyle 2012), and derived woodiness, an evolutionary reversion

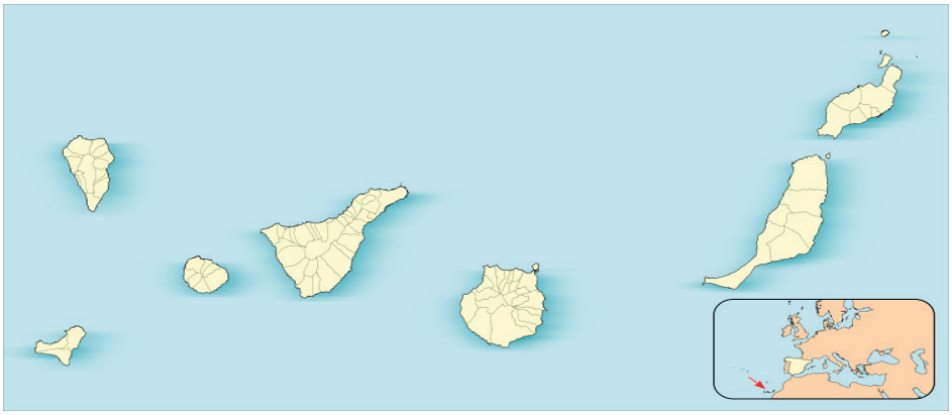


Figure 2. Map of the Canary Islands archipelago and its location ~100km from the coast of North-Western Africa. Images from OpenStreetMap.

of the evolution of woodiness from a herbaceous ancestor in a continental setting (Zizka et al. 2022).

Flora of the Canary Islands as a model system to understand trait-based diversification dynamics in angiosperms

The Canary Islands are an archipelago in the Macaronesian meta-archipelago (alongside the Azores, Madeira, and Cabo Verde archipelagos; Fernandez-Palacios et al. 2024). The Canaries are comprised of seven main islands (El Hierro, Fuerteventura, Gran Canaria, La Gomera, La Palma, Lanzarote, and Tenerife), two smaller islands under 30 km2 (La Graciosa and Alegranza), and several smaller islets. The oldest extant exposed rock emerged ca. 23 million years (myr) ago (Campeny et al. 2024), with many more volcanic islands previously present over the last 70 myrs, which have since disappeared due to erosion and sea-level change (Fernández-Palacios et al. 2011; Arjona et al. 2018; Beierkuhnlein et al. 2021). Located off the northwest coast of Africa, at the easternmost point, the Canary Islands lie approximately 100 km off the coast of Western Sahara. The close geographic proximity to the mainland is a notable difference from other well-studied volcanic archipelagos, such as Hawai'i or the Galápagos Islands, and likely facilitates dispersal and subsequent colonisation of the archipelago (Lens et al. 2013; Caujapé-Castells et al. 2017; Valente et al. 2017; Garcia-Verdugo et al. 2019). This proximity is also hypothesised to allow for multiple colonisations from the same species (Caujapé-Castells et al. 2011, 2017), allowing for elevated genetic diversity relative to other oceanic archipelagoes.

Alongside Hawai'i and the Galápagos Islands, the Canary Islands form one of the best-studied oceanic island systems (Florencio et al. 2021), and the

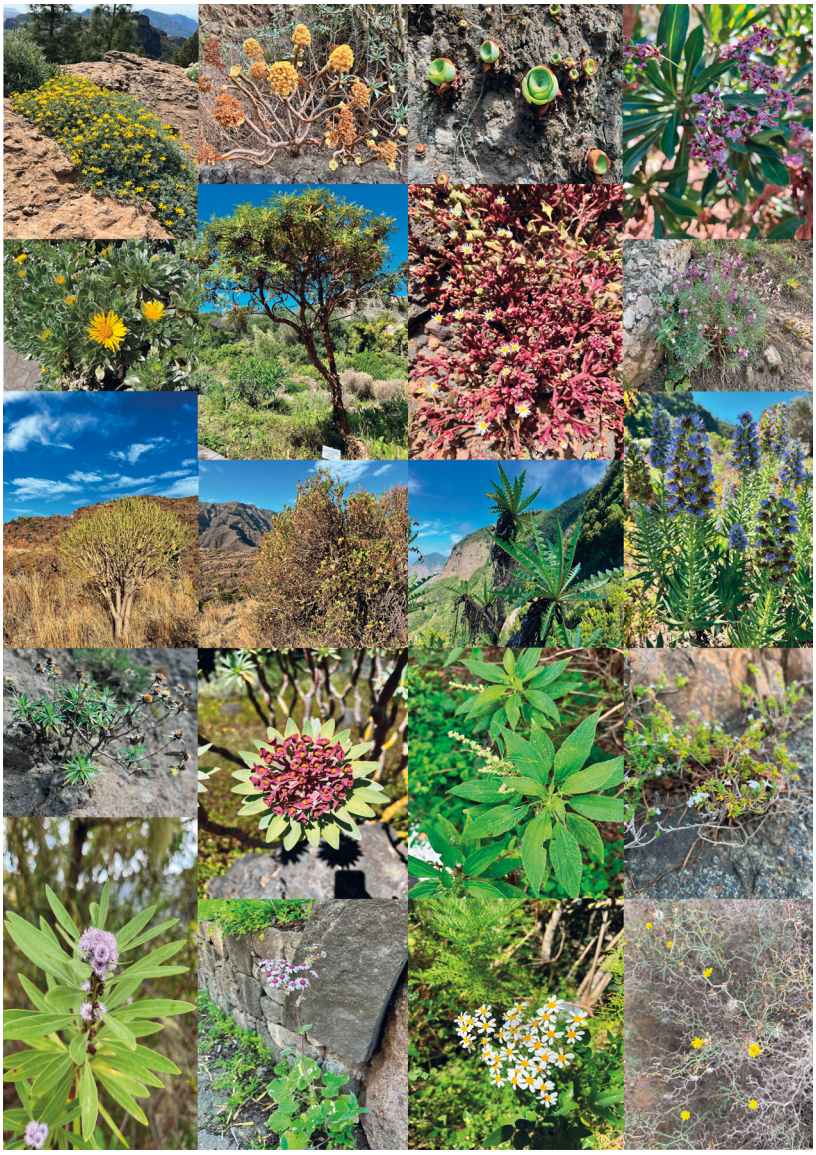


Figure 3. A selection of Canary Islands angiosperms, featuring many of the major evolutionary radiations from Macaronesia. From top left to right: *Lotus spartioides* Webb & Berthel. (Fabaceae); *Aeonium* cf *arboreum* (L.) Webb & Berthel. (Crassulaceae); *Aeonium aureum* (C.Sm. ex Hornem.) T.H.M.Mes [*Greenovia aurea* (C.Sm. ex Hornem.) Webb & Berthel.] (Crassulaceae); *Limonium dendroides* Svent. (Plumbaginaceae); *Asteriscus intermedius* (DC.) Pit. & Proust (Asteraceae); *Bencomia brachystachya* Svent. (Rosaceae); *Mesembryanthemum nodiflorum* L. (Aizoaceae); *Erysimum bicolor* (Hornem.) DC. (Brassicaceae); *Euphorbia berthelotii* Bolle ex Boiss. (Euphorbiaceae); *Rumex lunaria* L. (Polygonaceae); *Sonchus* sp. (Asteraceae); *Echium callithyrsum* Webb ex Bolle (Boraginaceae); *Carlina salicifolia* (L.f.) Cav. (Asteraceae); *Euphorbia atropurpurea* Brouss. ex Willd. (Euphorbiaceae); *Gesnouinia arborea* (L.f.) Gaudich. (Urticaceae); *Micromeria gomerensis* (P.Pérez) Puppo (Lamiaceae); *Globularia salicina* Lam. (Lamiaceae); *Pericallis webbii* (Sch.Bip.) Bolle (Asteraceae); *Launaea arborescens* (Batt.) Murb. (Asteraceae). Pictures by F. Lens.

native angiosperm flora is among the best-studied island floras globally (del Arco-Aguilar and Rodríguez Delgado 2018; Barajas Barbosa et al. 2023). With an estimated ~1,300 native species of angiosperms (Arjona et al. 2018; Beierkuhnlein et al. 2021; Canary Islands Biodiversity Database 2025), the Canarian angiosperm flora has one of the highest levels of species richness of any archipelago (Triantis et al. 2015), contains many evolutionary radiations (Cerca et al. 2023), and holds the highest number of IWS and shifts towards IW globally (Kim et al. 2008; Lens et al. 2013; Hooft van Huysduynen et al. 2021; Zizka et al. 2022). The archipelago is also the origin of the hypothesis linking insular woodiness to drought-related traits (Lens et al. 2013), making it the ideal setting to test if distribution patterns of the extant flora are correlated with environmental variables related to drought (Zizka et al. 2022). Endemicity rates for the archipelago are estimated to be ~40%, making the Canary Islands distinct with many species that evolved only on this small part of the world (Beierkuhnlein et al. 2021).

The diversity and composition of the Canary Island angiosperms are well known (Acebes Ginovés et al. 2009; Beierkuhnlein et al. 2021), as are the habitat and vegetation types in which they occur (del Arco et al. 2010; del Arco and Rodriguez 2018; Fernández-Palacios et al. 2021). The vegetative communities of the islands have been catalogued and mapped in impressive detail, revealing angiosperms to occur in a variety of habitats throughout the archipelago (del Arco and Rodriguez 2018). Such detailed knowledge of the Canary Island angiosperm composition allows for the testing of community assembly hypotheses via direct comparisons with existing data. Additionally, Canary Island angiosperm radiations span multiple habitat types throughout the archipelago, with variation both between and within clades, an important factor to be included in identifying the effect of traits on diversification (Ottaviani et al. 2020). For example, radiations in the genera *Echium* and *Aeonium* gave rise to niche differentiation in climatic niches (Florencio et al. 2021), a particularly pronounced effect between allopatric closely related taxa (Steinbauer et al. 2016). Conversely, niche conservatism has been important in *Helianthemum* diversification (Albaladejo et al. 2021), with inter-island dispersal to the same habitats and allopatric speciation driving the radiation. As Canary Island radiations arise from multiple different colonisation events with varying diversity, they offer the opportunity to use multiple replicates to study the dynamics of insular diversification following contemporary recommendations that trait-dependent diversification studies should use small-scale analyses replicated across independent clades (Anderson et al. 2023). Additionally, by satisfying the criteria required for identifying the drivers of diversification, occurrence in a wide variety of biome types and a broad phylogenetic history (Ottaviani et al. 2020;

Hooft van Huysduynen et al. 2021), the unique Canary Island angiosperm community is a perfect model group to study diversification dynamics.

Thesis Overview

This thesis investigates the processes underlying biodiversity in the Canary Islands' non-monocot angiosperm community by addressing how they are distributed in time (Chapter 2), space (Chapter 3), and across their phylogeny (Chapters 2 and 4). I also provide the required phylogenomic and trait data to test a novel, island-specific trait-dependent diversification model that extends the island biogeography model, DAISIE, by incorporating trait dependence via the SSE framework, using angiosperms of the Canary Islands as a model system (Figure 4). To do so, I generated the most complete phylogenomic reconstruction of the Canary Islands non-monocot angiosperm community (containing ~700 of ~1,160 species) to date, including the closest continental relative of each island clade to time-calibrate 435 colonisation events of the Canaries (Chapter 2). I also included available sequences from species outside the Canaries that have been associated with known fossils across the angiosperm family tree to time-calibrate this phylogeny. I then built a comprehensive trait database for the Canary Islands non-monocot angiosperms community, including 24 intrinsic plant traits and extrinsic environmental variables to investigate if and how these traits have affected diversification within the Canary Islands flora (Chapter 4). I also investigated how distribution patterns of the community are associated with these environmental variables (Chapter 3).

This thesis contains five chapters: the general introduction (Chapter 1, this chapter); a chapter on our phylogenomic reconstruction, biogeography, and the causes of diversity unevenness (time/ecological limits/variable diversification rates) in the Canary Islands non-monocot angiosperm community (Chapter 2); a chapter on how environmental variables shape phylogenetically-informed niche space for the growth forms in the community (Chapter 3); a chapter on trait-dependent diversification, using the phylogenetic and trait data with the novel trait-dependent diversification model to investigate diversification within the community (Chapter 4); and a general discussion including outlines for future research (Chapter 5). Throughout the chapters, I will address key questions in the literature, including whether the distribution of IWS is correlated with drought-related traits (Lens et al. 2013; Zizka et al. 2022), which combinations of traits and environmental variables have modified diversification rates in angiosperms (Sauquet and Magallon 2018), and whether there are universal drivers of diversification among angiosperms (Sauquet and Magallon 2018). Details of each chapter are given below.

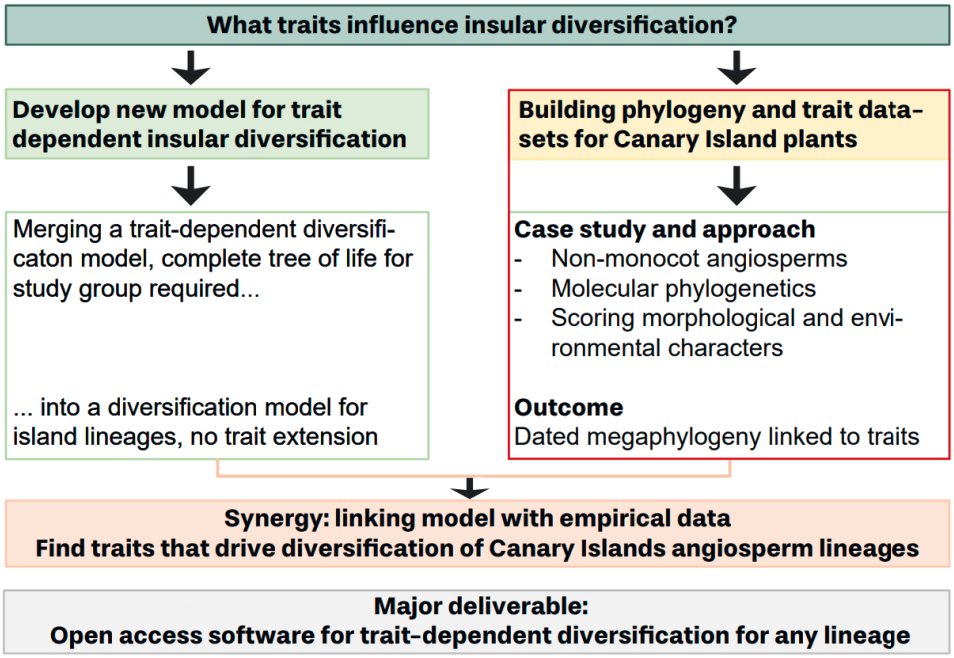


Figure 4. A graphic summary overview of the project linking plant phylogenomics and traits with a novel insular trait-dependent diversification model. The primary contribution of this thesis is highlighted in red; though, given the synergistic nature of this project, I also contributed to the section highlighted in green.

Chapter 2 - This chapter provides the phylogenomic reconstruction of the Canary Islands non-monocot angiosperm community. We identify the number of angiosperm (both monocot and non-monocot) colonisations of the Canary Islands, use this dataset to identify the closest continental ancestor of each colonising lineage, and score phylogenetically-informed growth forms for all the species included in our phylogeny. Our phylogeny contains 669 Canary Islands non-monocot species (~58% of the community), 771 continental sister species, and represents 435 plant lineages (of 676; ~64%). Using this phylogeny, we then address the causes of diversity unevenness across the phylogeny, estimate the rates of colonisation, extinction, and speciation for the community using DAISIE, discuss the biogeographic history of the flora in relation to colonisation times of several hundred lineages, and address whether insular woodiness (amongst other traits) has acted as a key trait driving diversification in the Canary Islands.

Chapter 3 - As growth form was found to significantly shape the Canary Islands non-monocot angiosperm flora in Chapter 2, we looked deeper into the environmental correlates and distribution patterns of various phylogenetically-in-

formed growth forms that take into account lifespan and evolutionary history (annual herbaceous, perennial herbaceous, ancestral woodiness, derived woodiness, and insular woodiness). This chapter reconstructs the phylogenetically-informed environmental niche space for the non-monocot flora of the Canary Islands, across five growth forms, to investigate how niche space is shaped by a suite of environmental variables and which variables are significantly correlated with the distribution of each growth form. We use a novel pipeline to extract environmental data from publicly available occurrence and spatial data, with greater precision offered from existing pipelines and additional cleaning steps. We find that the niche space of the community is shaped by climatic seasonality (in both temperature and precipitation) and topographic complexity, with different variables important for shaping the niche space of the growth forms, despite the unexpectedly high overlap between them. The environmental dataset generated for this chapter also contributes to the trait dataset built to investigate trait-dependent diversification in Chapter 4.

Chapter 4 - This chapter presents the novel trait-dependent diversification model specifically tailored for islands, TRAISIE (TRait-dependent Assembly of Island biota through Speciation, Immigration and Extinction). We first show the efficacy of TRAISIE using simulated datasets of various sizes to establish the conditions by which TRAISIE can accurately estimate evolutionary rates, correctly identify the generating model (constant rates/examined-trait dependence/concealed-trait dependence), and the amount of data required for reliable inference. By combining the phylogenetic and trait datasets presented herein with TRAISIE (see Figure 4), we address how two conspicuous and intertwined traits, growth form and life span, affect diversification in the Canary Islands non-monocot community.

Chapter 5 - In the final chapter of this thesis, a general discussion of the chapters is given to synthesise the findings in the context of the literature. This chapter is split into sections by topic: on trait-dependent diversification, biogeography (split into environmental correlates of plant distributions and colonisations of the Canary Islands by non-monocot angiosperms), diversity unevenness, unexplored projects, and future works. I also expand on the analyses in some of the chapters, to, for example, discuss how the niche space of Canarian radiations is shaped by environmental variables (and how the Canary Island environmental niche space compares to Hawai'i), further explore the relationship between time and diversification rates in shaping Canarian diversity unevenness and compare the relationships to angiosperms globally, and conjecture on why the Canary Islands harbour the most IWS globally.

INTRODUCTION REFERENCES

Acebes Ginovés, J.R.M., Arencibia, M.C.L., Rodríguez Navarro, M.L., del Arco Aguilar, M., Gallo, A.G., Pérez de Paz, P.L., Delgado, O.R., Martín Osorio, V.E., and de la Torre, W.W. Pteridophyta, Spermatophyta. In: Lista de especies silvestres de Canarias. Hongos, plantas y animales terrestres. (2009). Arechavaleta, M., Rodríguez, S., Zurita, N., and García, A. Gobierno de Canarias. p:119-172.

Albaladejo, R.G., Martín-Hernanz, S., Reyes-Betancort, J.A., Santos-Guerra, A., Olangua-Corral, M. and Aparicio, A. (2021). Reconstruction of the spatio-temporal diversification and ecological niche evolution of Helianthemum (Cistaceae) in the Canary Islands using genotyping-by-sequencing data. *Annals of Botany*, *127*(5), pp.597-611.

Anderson, B., Pannell, J., Billiard, S., Burgarella, C., de Boer, H., Dufay, M., & Glemin, S. (2023). Opposing effects of plant traits on diversification. *Iscience*, 26(4).

Antonelli, A., Govaerts, R., Nic Lughadha, E., Onstein, R. E., Smith, R. J., & Zizka, A. (2023). Why plant diversity and distribution matter. *New Phytologist*, 240(4).

Arjona, Y., Nogales, M., Heleno, R. and Vargas, P. (2018). Long-distance dispersal syndromes matter: diaspore–trait effect on shaping plant distribution across the Canary Islands. *Ecography*, *41*(5), pp.805-814.

Baldwin, B. G., & Sanderson, M. J. (1998). Age and rate of diversification of the Hawaiian silversword alliance (Compositae). *Proceedings of the National Academy of Sciences*, 95(16), 9402-9406.

Ballard Jr, H.E. and Sytsma, K.J. (2000). Evolution and biogeography of the woody Hawaiian violets (Viola, Violaceae): Arctic origins, herbaceous ancestry and bird dispersal. *Evolution*, *54*(5), pp.1521-1532.

Barajas Barbosa, M. P., Craven, D., Weigelt, P., Denelle, P., Otto, R., Díaz, S., ... & Kreft, H. (2023). Assembly of functional diversity in an oceanic island flora. *Nature*, 619(7970), 545-550.

Beaulieu, J. M., & O'Meara, B. C. (2016). Detecting hidden diversification shifts in models of trait-dependent speciation and extinction. *Systematic Biology*, 65(4), 583-601.

Beierkuhnlein, C., Walentowitz, A. and Welss, W. (2021). FloCan—A Revised Checklist for the Flora of the Canary Islands. *Diversity*, 13(10), p.480.

Beierkuhnlein, C., Nogales, M., Field, R., Vetaas, O. R., Walentowitz, A., Weiser, F., & Chiarucci, A. (2023). Volcanic ash deposition as a selection mechanism towards woodiness. *npj Biodiversity,* 2(1), 14.

Benton, M. J., Wilf, P., & Sauquet, H. (2022). The Angiosperm Terrestrial Revolution and the origins of modern biodiversity. *New Phytologist, 233*(5), 2017-2035.

Bhadra, S., Leitch, I. J., & Onstein, R. E. (2023). From genome size to trait evolution during angiosperm radiation. *Trends in Genetics*, *39*(10), 728-735.

Blaimer, B. B., Santos, B. F., Cruaud, A., Gates, M. W., Kula, R. R., Mikó, I., & Buffington, M. L. (2023). Key innovations and the diversification of Hymenoptera. *Nature Communications, 14*(1), 1212.

Böhle, U.R., Hilger, H.H. and Martin, W.F. (1996). Island colonization and evolution of the insular woody habit in *Echium* L.(Boraginaceae). *Proceedings of the National Academy of Sciences*, *93*(21), pp.11740-11745.

Bouchenak-Khelladi, Y., Muasya, A. M., & Linder, H. P. (2014). A revised evolutionary history of Poales: origins and diversification. *Botanical Journal of the Linnean Society*, *175*(1), 4-16.

Boucher, F. C., Verboom, G. A., Musker, S., & Ellis, A. G. (2017). Plant size: a key determinant of diversification?. *New Phytologist*, *216*(1), 24-31.

Boyko, J. D., Hagen, E. R., Beaulieu, J. M., & Vasconcelos, T. (2023). The evolutionary responses of life-history strategies to climatic variability in flowering plants. *New Phytologist, 240*(4), 1587-1600.

Brilhante, M., Roxo, G., Catarino, S., Dos Santos, P., Reyes-Betancort, J. A., Cau-

japé-Castells, J. & Romeiras, M. M. (2021). Diversification of Aeonium species across Macaronesian Archipelagos: Correlations between genome-size variation and their conservation status. *Frontiers in Ecology and Evolution, 9*, 607338.

Campeny, M., Menéndez, I., Quevedo, L., Yepes, J., Casillas, R., Ahijado, A., & Mangas, J. (2024). Rare Earth element distribution on the Fuerteventura Basal Complex (Canary Islands, Spain): a geochemical and mineralogical approach. *Solid Earth, 15(*6), 639-656.

Canary Islands Biodiversity Database (2025). Government of the Canary Islands. (https://www.biodiversidadcanarias.es/biota) [10th October 2023].

Carine, M.A., Santos-Guerra, A., Guma, I.R. and Reyes-Betancort, J.A. (2010). *Endemism and evolution of the Macaronesian flora*. In Beyond cladistics (pp. 101-124). University of California Press.

Carlquist, S. (1974). *Island Biology*. Columbia University Press.

Carlquist S., Baldwin, B.G., Carr, G.D. (2003). *Tarweeds and silverswords: evolution of the Madiinae (Asteraceae)*. Missouri Botanical Garden, St. Louis.

Caujapé-Castells, J. (2011). *Jesters, red queens, boomerangs and surfers: a molecular outlook on the diversity of the Canarian endemic flora*. The biology of island floras, pp.284-324.

Caujapé-Castells, J., García-Verdugo, C., Marrero-Rodríguez, Á., Fernández-Palacios, J.M., Crawford, D.J. and Mort, M.E. (2017). Island ontogenies, syngameons, and the origins and evolution of genetic diversity in the Canarian endemic flora. *Perspectives in Plant Ecology, Evolution and Systematics, 27*, pp.9-22.

Cerca, J., Cotoras, D. D., Bieker, V. C., De-Kayne, R., Vargas, P., Fernández-Mazuecos, M., & Martin, M. D. (2023). Evolutionary genomics of oceanic island radiations. *Trends in Ecology & Evolution, 38*(7), 631-642.

Curto, M., Schachtler, C., Puppo, P., & Meimberg, H. (2018). Using a new RAD-sequencing approach to study the evolution of Micromeria in the Canary Islands. *Molecular Phylogenetics and Evolution, 119*, 160-169.

Darwin, C. (1859). *On the origin of species by means of natural selection, or, the preservation of favoured races in the struggle for life*. John Murray, London.

Davis, M. P., Midford, P. E., & Maddison, W. (2013). Exploring power and parameter estimation of the BiSSE method for analyzing species diversification. *BMC Evolutionary Biology,* I(1), 38.

Dehayem, O. N., Brewer, R. F., Valente, L., Lens, F., & Etienne, R. S. (2025). Impact of sampling strategy on inference of community assembly processes in phylogenetic island biogeography. *Methods in Ecology and Evolution*, *16*, 1507-1520.

del Arco Aguilar, M.J., González-González, R., Garzón-Machado, V. and Pizarro-Hernández, B. (2010). Actual and potential natural vegetation on the Canary Islands and its conservation status. *Biodiversity and Conservation*, *19*(11), pp.3089-3140.

del Arco Aguilar, M. J., & Rodríguez Delgado, O. (2018). *Vegetation of the Canary Islands* (pp. 83-319). Springer International Publishing.

Donoghue, M. J., & Sanderson, M. J. (2015). Confluence, synnovation, and depauperons in plant diversification. *New Phytologist*, *207*(2), 260-274.

Donoghue, M. J., Eaton, D. A., Maya-Lastra, C. A., Landis, M. J., Sweeney, P. W., Olson, M. E. & Edwards, E. J. (2022). Replicated radiation of a plant clade along a cloud forest archipelago. *Nature Ecology & Evolution, 6*(9), 1318-1329.

Dória, L. C., Podadera, D. S., del Arco, M., Chauvin, T., Smets, E., Delzon, S., & Lens, F. (2018). Insular woody daisies (Argyranthemum, Asteraceae) are more resistant to drought-induced hydraulic failure than their herbaceous relatives. *Functional Ecology, 32*(6), 1467-1478.

Doyle, J. A. (2012). Molecular and fossil evidence on the origin of angiosperms. Annual Review of Earth and Planetary Sciences, 40, 301–326.

Etienne, R. S., Pigot, A. L., & Phillimore, A. B. (2016). How reliably can we infer diversity-dependent diversification from phylogenies?. *Methods in Ecology and Evolution, 7*(9), 1092-1099.

Etienne, R. S., Cabral, J. S., Hagen, O., Hartig, F., Hurlbert, A. H., Pellissier, L. & Storch, D. (2019). A minimal model for the latitudinal diversity gradient suggests a dominant role for ecological limits. *The American Naturalist, 194*(5), E122-E133.

Falster, D. S., & Westoby, M. (2003). Plant height and evolutionary games. *Trends in Ecology & Evolution, 18*(7), 337-343.

Fernández-Mazuecos, M., & Glover, B. J. (2017). The evo-devo of plant speciation. *Nature Ecology & Evolution, 1*(4), 0110.

Fernández-Palacios, J.M., De Nascimento, L., Otto, R., Delgado, J.D., García-del-Rey, E., Arévalo, J.R. and Whittaker, R.J. (2011). A reconstruction of Palaeo-Macaronesia, with particular reference to the long-term biogeography of the Atlantic island laurel forests. *Journal of Biogeography, 38*(2), pp.226-246.

Fernández-Palacios, J.M., Otto, R., Borregaard, M.K., Kreft, H., Price, J.P., Steinbauer, M.J., Weigelt, P. and Whittaker, R.J. (2021). Evolutionary winners are ecological losers among oceanic island plants. *Journal of Biogeography, 48*(9), pp.2186-2198.

Fernández-Palacios, J. M., Otto, R., Capelo, J., Caujapé-Castells, J., De Nascimento, L., Duarte, M. C. & Whittaker, R. J. (2024). In defence of the entity of Macaronesia as a biogeographical region. *Biological Reviews, 99*(6), 2060-2081.

FitzJohn, R.G. (2012). Diversitree: comparative phylogenetic analyses of diversification in R. *Methods in Ecology and Evolution*, *3*(6), pp.1084-1092.

Florencio, M., Patiño, J., Nogué, S., Traveset, A., Borges, P.A., Schaefer, H., Amorim, I.R., Arnedo, M., Ávila, S.P., Cardoso, P. and de Nascimento, L. (2021). Macaronesia as a fruitful arena for ecology, evolution, and conservation biology. *Frontiers in Ecology and Evolution*, p.752.

García-Verdugo, C., Caujapé-Castells, J. and Sanmartín, I. (2019). Colonization time on island settings: lessons from the Hawaiian and Canary Island floras. *Botanical Journal of the Linnean Society*, *191*(2), pp.155-163.

Gavrutenko, M., Reginato, M., Kriebel, R., Nicolas, A. N., & Michelangeli, F. A. (2020). Evolution of floral morphology and symmetry in the Miconieae (Melastomataceae): multiple generalization trends within a specialized family. *International Journal of Plant Sciences*, 181(7), 732-747.

Gerrienne, P., Gensel, P. G., Strullu-Derrien, C., Lardeux, H., Steemans, P., & Prestianni, C. (2011). A simple type of wood in two Early Devonian plants. Science, 333(6044), 837–837.

Gillespie, R. G., & Whittaker, R. J. (2025). Eco-evolutionary assembly processes within oceanic island systems. *Evolutionary Journal of the Linnean Society*, 4(1), kzaf022.

Givnish, T. (1998). Adaptive radiation of plants on oceanic islands: classic patterns, molecular data, new insights. pp.281-304 in: Grant, P. (ed), *Evolution on Islands*. New York: Oxford University Press.

Givnish, T.J. (2010). Ecology of plant speciation. *Taxon*, *59*(5), pp.1326-1366.

Goldberg, E.E. and Igić, B. (2012). Tempo and mode in plant breeding system evolution. *Evolution: International Journal of Organic Evolution*, *66*(12), pp.3701-3709.

Gotelli, N. J., Anderson, M. J., Arita, H. T., Chao, A., Colwell, R. K., Connolly, S. R., & Willig, M. R. (2009). Patterns and causes of species richness: a general simulation model for macroecology. *Ecology letters, 12*(9), 873-886.

Harmon, L. J., Pennell, M. W., Henao-Diaz, L. F., Rolland, J., Sipley, B. N., & Uyeda, J. C. (2021). Causes and consequences of apparent timescaling across all estimated evolutionary rates. *Annual Review of Ecology, Evolution, and Systematics, 52*(1), 587-609.

Helmstetter, A. J., Zenil-Ferguson, R., Sauquet, H., Otto, S. P., Méndez, M., Vallejo-Marin, M., & Käfer, J. (2023). Trait-dependent diversification in angiosperms: Patterns, models and data. *Ecology Letters, 26*(4), 640-657.

Henao Diaz, L. F., Harmon, L. J., Sugawara, M. T., Miller, E. T., & Pennell, M. W. (2019). Macroevolutionary diversification rates show time dependency. *Proceedings of the National Academy of Sciences, 116*(15), 7403-7408.

Hernández-Hernández, T. and Wiens, J.J. (2020). Why are there so many flowering plants? A multiscale analysis of plant

diversification. *The American Naturalist*, *195*(6), pp.948-963.
Herrera-Alsina, L., van Els, P. and Etienne, R.S. (2019). Detecting the dependence of diversification on multiple traits from phylogenetic trees and trait data. *Systematic Biology*, *68*(2), pp.317-328.
Hooft van Huysduynen, A., Janssens, S., Merckx, V., Vos, R., Valente, L., Zizka, A., Larter, M., Karabayir, B., Maaskant, D., Witmer, Y. and Fernández-Palacios, J.M. (2021). Temporal and palaeoclimatic context of the evolution of insular woodiness in the Canary Islands. *Ecology and Evolution*, *11*(17), pp.12220-12231.
Kim, S.C., McGowen, M.R., Lubinsky, P., Barber, J.C., Mort, M.E. and Santos-Guerra, A. (2008). Timing and tempo of early and successive adaptive radiations in Macaronesia. *PLoS One*, *3*(5), p.e2139.
König, C., Weigelt, P., Taylor, A., Stein, A., Dawson, W., Essl, F., Pergl, J., Py'ek, P., van Kleunen, M., Winter, M. and Chatelain, C. (2021). Source pools and disharmony of the world's island floras. *Ecography*, *44*(1), pp.44-55.
Koutroumpa, K., Warren, B.H., Theodoridis, S., Coiro, M., Romeiras, M.M., Jiménez, A. and Conti, E. (2021). Geo-climatic changes and apomixis as major drivers of diversification in the Mediterranean sea lavenders (Limonium Mill.). *Frontiers in Plant Science*, *11*, p.2127.
Lens, F., Davin, N., Smets, E. and del Arco, M. (2013). Insular woodiness on the Canary Islands: a remarkable case of convergent evolution. *International Journal of Plant Sciences*, *174*(7), pp.992-1013.
Lomolino, M.V. (2000). A call for a new paradigm of island biogeography. *Global Ecology and Biogeography*, *9*(1), pp.1-6.
Losos, J.B. and Ricklefs, R.E. (2009). Adaptation and diversification on islands. *Nature*, *457*(7231), pp.830-836.
Losos, J.B., Ricklefs, R.E., and MacArthur, R.H., eds. (2010). Island Biogeography in the 1960s. In: *The theory of island biogeography revisited*. Princeton, NJ: Princeton University Press.
MacArthur, R. H., & Wilson, E. O. (1963). An equilibrium theory of insular zoogeography. *Evolution*, 373-387.
MacArthur, R.H., and Wilson, E.O. (1967). *The Theory of Island Biogeography*. Princeton University Press.
Maddison, W.P. and FitzJohn, R.G. (2015). The unsolved challenge to phylogenetic correlation tests for categorical characters. *Systematic Biology*, *64*(1), pp.127-136.
Maddison, W.P., Midford, P.E. and Otto, S.P. (2007). Estimating a binary character's effect on speciation and extinction. *Systematic Biology*, *56*(5), pp.701-710.
Magallón, S., Gómez-Acevedo, S., Sánchez-Reyes, L.L. and Hernández-Hernández, T. (2015). A metacalibrated time-tree documents the early rise of flowering plant phylogenetic diversity. *New Phytologist*, *207*(2), pp.437-453.
Mayrose, I., Zhan, S. H., Rothfels, C. J., Magnuson-Ford, K., Barker, M. S., Rieseberg, L. H., & Otto, S. P. (2011). Recently formed polyploid plants diversify at lower rates. *Science*, *333*(6047), 1257-1257.
McPeek, M. A., & Brown, J. M. (2007). Clade age and not diversification rate explains species richness among animal taxa. *The American Naturalist, 169*(4), E97-E106.
Mynard, P., Algar, A. C., Lancaster, L. T., Bocedi, G., Fahri, F., Gubry-Rangin, C. & Herrera-Alsina, L. (2023). Impact of phylogenetic tree completeness and mis-specification of sampling fractions on trait dependent diversification models. *Systematic Biology, 72*(1), 106-119.
Negrutiu, I., Frohlich, M. W., & Hamant, O. (2020). Flowering plants in the Anthropocene: A political agenda. *Trends in plant science, 25*(4), 349-368.
Niklas, K. J. (2016). Plant evolution: an introduction to the history of life. University of Chicago Press, Chicago.
Nürk, N.M., Atchison, G.W. and Hughes, C.E. (2019). Island woodiness underpins accelerated disparification in plant radiations. *New Phytologist*, 224(1), pp.518-531.
Nürk, N.M., Linder, H.P., Onstein, R.E., Larcombe, M.J., Hughes, C.E., Piñeiro Fernández, L., Schlüter, P.M., Valente, L., Beierkuhnlein, C., Cutts, V. and Donoghue, M.J. (2020). Diversification in evolutionary arenas–Assessment

and synthesis. *Ecology and Evolution, 10*(12), pp.6163-6182.

Ottaviani, G., Keppel, G., Götzenberger, L., Harrison, S., Opedal, Ø.H., Conti, L., Liancourt, P., Klimešová, J., Silveira, F.A., Jiménez-Alfaro, B. and Negoita, L. (2020). Linking plant functional ecology to island biogeography. *Trends in Plant Science*, *25*(4), pp.329-339.

Patiño, J., Whittaker, R.J., Borges, P.A., Fernández-Palacios, J.M., Ah-Peng, C., Araújo, M.B., Ávila, S.P., Cardoso, P., Cornuault, J., de Boer, E.J. and de Nascimento, L. (2017). A roadmap for island biology: 50 fundamental questions after 50 years of The Theory of Island Biogeography. *Journal of Biogeography*, *44*(5), pp.963-983.

Pérez-Escobar, O. A., Bogarín, D., Przelomska, N. A., Ackerman, J. D., Balbuena, J. A., Bellot, S., & Antonelli, A. (2024). The origin and speciation of orchids. *New Phytologist, 242*(2), 700-716.

Peris, D., & Condamine, F. L. (2024). The angiosperm radiation played a dual role in the diversification of insects and insect pollinators. *Nature Communications, 15*(1), 552.

Pie, M. R., Divieso, R., & Caron, F. S. (2023). Clade density and the evolution of diversity-dependent diversification. *Nature Communications, 14*(1), 4576.

Poppenwimer, T., Mayrose, I., & DeMalach, N. (2023). Revising the global biogeography of annual and perennial plants. *Nature*, 624(7990), 109-114.

Proches, Ş., Wilson, J. R., Vamosi, J. C., & Richardson, D. M. (2008). Plant diversity in the human diet: weak phylogenetic signal indicates breadth. *Bioscience*, 58(2), 151-159.

Puttick, M.N., Clark, J. and Donoghue, P.C. (2015). Size is not everything: rates of genome size evolution, not C-value, correlate with speciation in angiosperms. *Proceedings of the Royal Society B: Biological Sciences, 282*(1820), p.20152289.

Rabosky, D. L. (2013). Diversity-dependence, ecological speciation, and the role of competition in macroevolution. *Annual Review of Ecology, Evolution, and Systematics, 44(*1), 481-502.

Rabosky, D.L. and Goldberg, E.E (2015). Model inadequacy and mistaken inferences of trait-dependent speciation. *Systematic Biology*, *64*(2), pp.340-355.

Rabosky, D. L., & Hurlbert, A. H. (2015). Species richness at continental scales is dominated by ecological limits. *The American Naturalist, 185*(5), 572-583.

Rosindell, J., & Phillimore, A. B. (2011). A unified model of island biogeography sheds light on the zone of radiation. *Ecology Letters, 14*(6), 552-560.

Sabath, N., Goldberg, E. E., Glick, L., Einhorn, M., Ashman, T. L., Ming, R. & Mayrose, I. (2016). Dioecy does not consistently accelerate or slow lineage diversification across multiple genera of angiosperms. *New Phytologist, 209*(3), 1290-1300.

Sauquet, H., Von Balthazar, M., Magallón, S., Doyle, J.A., Endress, P.K., Bailes, E.J., de Morais, E.B., Bull-Hereñu, K., Carrive, L., Chartier, M. and Chomicki, G. (2017). The ancestral flower of angiosperms and its early diversification. *Nature Communications*, *8*(1), pp.1-10.

Sauquet, H. and Magallón, S. (2018). Key questions and challenges in angiosperm macroevolution. *New Phytologist*, *219*(4), pp.1170-1187.

Sauquet, H., Ramírez-Barahona, S., & Magallón, S. (2022). What is the age of flowering plants?. *Journal of Experimental Botany, 73*(12), 3840-3853.

Schenk, J. J. (2021). The next generation of adaptive radiation studies in plants. *International Journal of Plant Sciences, 182*(4), 245-262.

Schrader, J. (2025). Trait-based island biogeography as a tool for studying future ecological communities. *New Phytologist*.

Schrader, J., Wright, I. J., Kreft, H., & Westoby, M. (2021). A roadmap to plant functional island biogeography. *Biological Reviews, 96*(6), 2851-2870.

Silvestro, D., Cascales-Miñana, B., Bacon, C. D., & Antonelli, A. (2015). Revisiting the origin and diversification of vascular plants through a comprehensive Bayesian analysis of the fossil record. *New Phytologist, 207*(2), 425-436.

Simberloff, D.S. and Wilson, E.O. (1969). Experimental zoogeography of islands: the colonization of empty islands. *Ecology, 50*(2), pp.278-296.

Simões, M., Breitkreuz, L., Alvarado, M., Baca, S., Cooper, J. C., Heins, L. & Lieberman, B. S. (2016). The evolving theory of evolutionary radiations. *Trends in ecology & evolution, 31*(1), 27-34.

Smith, S. A., & Beaulieu, J. M. (2024). Ad fontes: divergence-time estimation and the age of angiosperms. *New Phytologist, 244*(3), 760-766.

Sperotto, P., Roque, N., Acevedo-Rodríguez, P., & Vasconcelos, T. (2023). Climbing mechanisms and the diversification of neotropical climbing plants across time and space. *New Phytologist, 240*(4), 1561-1573.

Steinbauer, M.J., Field, R., Fernández-Palacios, J.M., Irl, S.D., Otto, R., Schaefer, H. and Beierkuhnlein, C. (2016). Biogeographic ranges do not support niche theory in radiating Canary Island plant clades. *Global Ecology and Biogeography*, *25*(7), pp.792-804.

Strutzenberger, P., & Fiedler, K. (2011). Temporal patterns of diversification in Andean Eois, a species-rich clade of moths (Lepidoptera, Geometridae). *Journal of Evolutionary Biology, 24*(4), 919-925.

Tank, D. C., Eastman, J. M., Pennell, M. W., Soltis, P. S., Soltis, D. E., Hinchliff, C. E. & Harmon, L. J. (2015). Nested radiations and the pulse of angiosperm diversification: increased diversification rates often follow whole genome duplications. *New Phytologist*, *207*(2), 454-467.

Taylor, A., Weigelt, P., Denelle, P., Cai, L., & Kreft, H. (2023). The contribution of plant life and growth forms to global gradients of vascular plant diversity. *New Phytologist, 240*(4), 1548-1560.

Tietje, M., Antonelli, A., Baker, W. J., Govaerts, R., Smith, S. A., & Eiserhardt, W. L. (2022). Global variation in diversification rate and species richness are unlinked in plants. *Proceedings of the National Academy of Sciences, 119*(27), e2120662119.

Triantis, K.A. and Matthews, T.J. (2020). Biodiversity theory backed by island bird data. *Nature*, doi:10.1038/d41586-020-00426-5.

Triantis, K.A., Guilhaumon, F. and Whittaker, R.J. (2012). The island species–area relationship: biology and statistics. *Journal of Biogeography*, *39*(2), pp.215-231.

Triantis, K. A., Economo, E. P., Guilhaumon, F., & Ricklefs, R. E. (2015). Diversity regulation at macro-scales: species richness on oceanic archipelagos. *Global Ecology and Biogeography, 24*(5), 594-605.

Valente, L. M., & Vargas, P. (2013). Contrasting evolutionary hypotheses between two mediterranean-climate floristic hotspots: the Cape of southern Africa and the Mediterranean Basin. *Journal of Biogeography, 40*(11), 2032-2046.

Valente, L.M., Etienne, R.S. and Phillimore, A.B. (2014). The effects of island ontogeny on species diversity and phylogeny. *Proceedings of the Royal Society B: Biological Sciences*, *281*(1784), p.20133227.

Valente, L.M., Phillimore, A.B. and Etienne, R.S. (2015). Equilibrium and non-equilibrium dynamics simultaneously operate in the Galápagos islands. *Ecology Letters*, *18*(8), pp.844-852.

Valente, L., Illera, J.C., Havenstein, K., Pallien, T., Etienne, R.S. and Tiedemann, R. (2017). Equilibrium bird species diversity in Atlantic islands. *Current Biology*, *27*(11), pp.1660-1666.

Valente, L., Phillimore, A.B., Melo, M., Warren, B.H., Clegg, S.M., Havenstein, K., Tiedemann, R., Illera, J.C., Thébaud, C., Aschenbach, T. and Etienne, R.S. (2020). A simple dynamic model explains the diversity of island birds worldwide. *Nature*, *579*(7797), pp.92-96.

Vamosi, J. C., Magallón, S., Mayrose, I., Otto, S. P., & Sauquet, H. (2018). Macroevolutionary patterns of flowering plant speciation and extinction. *Annual Review of Plant Biology, 69*(1), 685-706.

Vasconcelos, T. (2023). A trait-based approach to determining principles of plant biogeography. *American Journal of Botany*, *110*(2), e16127.

Wallace, A.R. (1869). *The Malay Archipelago: The land of the orang-utan, and the bird of paradise. A narrative of travel, with studies of man and nature*. Macmillan, London.

White, O. W., Reyes-Betancort, J. A., Chapman, M. A., & Carine, M. A. (2020). Geographical isolation, habitat shifts and hybridisation in the diversification of the Macaronesian endemic genus *Argyranthemum* (Asteraceae). *New Phytologist, 228*(6), 1953-1971.

Whittaker, R.J., Triantis, K.A. and Ladle, R.J. (2008). A general dynamic theory of oceanic island biogeography. *Journal of Biogeography, 35*(6), pp.977-994.

Wiens, J.J. (2017). What explains patterns of biodiversity across the Tree of Life? New research is revealing the causes of the dramatic variation in species numbers across branches of the Tree of Life. *BioEssays*, *39*(3), p.1600128.

Xue, B., Guo, X., Landis, J. B., Sun, M., Tang, C. C., Soltis, P. S. & Saunders, R. M. K. (2020). Accelerated diversification correlated with functional traits shapes extant diversity of the early divergent angiosperm family Annonaceae. *Molecular Phylogenetics and Evolution, 142*, 106659.

Yu, D., & Wiens, J. J. (2024). The causes of species richness patterns among clades. *Proceedings of the Royal Society B, 291*(2015), 20232436.

Zizka, A., Onstein, R. E., Rozzi, R., Weigelt, P., Kreft, H., Steinbauer, M. J., & Lens, F. (2022). The evolution of insular woodiness. *Proceedings of the National Academy of Sciences, 119*(37), e2208629119.

Zuntini, A. R., Carruthers, T., Maurin, O., Bailey, P. C., Leempoel, K., Brewer, G. E., & Knapp, S. (2024). Phylogenomics and the rise of the angiosperms. *Nature*, *629*(8013), 843-850.



Chapter 2

Ecological limits and trait-dependent diversification drive angiosperm diversity unevenness across the Canary Islands

Ryan F.A. Brewer[1,2*], Ornela Dehayem Nanwou[1,2], Laura van Hoek[1], Marcelino José del Arco Aguilar[3], Yurena Arjona[4], William J. Baker[5,6], Noor M.S. van den Berg[1], Juli Caujapé-Castells[7], José María Fernández-Palacios[8], Cristina González-Montelongo[9,10], Ruth Jaén-Molina[7], Lucas S. Jansen[1], Louis S. Jay-García[4,9], Águedo Marrero[11], Sara Martín-Hernanz[5,12], Olivier Maurin[5,13], Raquel Negrão[5], Jairo Patiño[4], Stephan Scholz[14], Pablo Vargas[15], Alexandre R. Zuntini[5], Rampal S. Etienne[2,], Luis Valente[1,2,**], Frederic Lens[1,16,**]**

[1] Naturalis Biodiversity Center, Darwinweg 2, 2333 CR Leiden, The Netherlands
[2] Groningen Institute for Evolutionary Life Sciences, University of Groningen, Box 11103, 9700 CC Groningen, The Netherlands
[3] Plant Conservation and Biogeography Research Group. Departamento de Botánica, Ecología y Fisiología Vegetal. Universidad de La Laguna. Ap. Postal 456, 38200. La Laguna, Tenerife, Islas Canarias, Spain.
[4] Island Ecology and Evolution Research Group, Instituto de Productos Naturales y Agrobiología (IPNA-CSIC), La Laguna, Tenerife, Spain
[5] Royal Botanical Gardens, Kew, Richmond, London, TW9 3AE, United Kingdom.
[6] Department of Biology, Aarhus University, Aarhus, Denmark
[7] Department of Molecular Biodiversity & DNA Bank of the Canarian Flora, Jardín Botánico Canario 'Viera y Clavijo'–Unidad Asociada CSIC, Cabildo de Gran Canaria, Las Palmas, Gran Canaria, Spain
[8] Island Ecology and Biogeography Group, Instituto Universitario de Enfermedades Tropicales y Salud Pública de Canarias (IUETSPC), Universidad de La Laguna (ULL), 38200 La Laguna, Canary Islands, Spain
[9] Department of Botany, Ecology and Plant Physiology, Faculty of Pharmacy, University of La Laguna, Apdo, Postal: 456, 38200 San Cristobal de La Laguna, Tenerife, Canary Islands, Spain
[10] TFC Herbarium. Servicio General de Apoyo a la Investigación (SEGAI). Universidad de La Laguna. 38200 San Cristobal de La Laguna, Tenerife, Canary Islands, Spain
[11] Department of Plant Systematics and Herbarium LPA, Jardín Botánico Canario 'Viera y Clavijo'–Unidad Asociada CSIC, Cabildo de Gran Canaria, Las Palmas, Gran Canaria, Spain
[12] Dep. Biodiversity, Ecology and Evolution (Universidad Complutense de Madrid), C/ José Antonio Novais, 12, Ciudad Universitaria, 28040 - Madrid, Spain
[13] Meise Botanic Garden, Nieuwelaan 38, 1860 Meise, Belgium
[14] Jardín Botánico Oasis Wildlife Fuerteventura, Ctra. General de Jandía, 35627 La Lajita, Fuerteventura, Spain
[15] Department of Biodiversity and Conservation, Real Jardın Botánico (RJB-CSIC), Plaza de Murillo 2, 28014 Madrid, Spain
[16] Institute of Biology Leiden, Leiden University, Sylviusweg 72, Leiden 2333 BE, The Netherlands

*Corresponding author: Ryan F. A. Brewer (R.F.Brewer@outlook.com)
**Shared Senior Authors

Under review at *PNAS*

ABSTRACT

Island biotas show highly uneven species richness among lineages, influenced by clade age, diversification rates, and/or ecological limits. Disentangling these drivers remains challenging due to the lack of comprehensive datasets across multiple lineages in the same system. The Canary Islands flora includes hundreds of lineages with contrasting species diversity and harbours the highest number of species that evolved woodiness in-situ ("insular woodiness"). Here, we present a time-calibrated phylogenomic reconstruction for Canary Island non-monocot angiosperms, based on 669 Canarian species and 771 closely related mainland taxa, allowing us to simultaneously study 435 lineages (~65% of non-monocots). Applying dynamic stochastic modelling, we find the flora is shaped by high extinction and colonisation rates, maintaining macroevolutionary equilibrium, with woody, insular woody, and perennial lineages exhibiting higher diversification rates than the remaining flora. Our results suggest uneven diversity within a dynamic insular system can be explained by the interaction of trait evolution and ecological constraints.

INTRODUCTION

Explaining the macroevolutionary processes that have given rise to the unevenness in diversity across the tree of life is one of the major goals of (macro) ecology, evolutionary biology, and biogeography (von Humboldt and Bonpland 1807; Rangel et al. 2018; Vasconcelos 2023). Three main (non-mutually exclusive) explanations of this unevenness have been proposed (Etienne et al. 2019): 1) variation in clade ages, with time differences being the sole explanatory factor (i.e., older lineages have more species than younger lineages; McPeek and Brown 2007; Henao Diaz et al. 2019; Yu and Wiens 2024); 2) variation in diversification rates, emphasising lineages accumulate species at differing rates through evolutionary time due to intrinsic characteristics of the clade (Gotelli et al. 2009; Wiens 2017); and 3) variable ecological limits on diversity over time (limits to species richness over time by resource limitations or the carrying capacity of an ecosystem/clade; Valente and Vargas 2013; Rabosky and Hurlbert 2015; Wiens 2017). Underpinning these three hypotheses are the dynamic biogeographic processes of colonisation, speciation, and extinction, along with their interplay through time and space (Ricklefs 2004; Weigelt et al. 2015). Integrating time-calibrated phylogenetic data with morphological and geographical information has become a powerful approach to better understand the mechanisms underlying these biodiversity patterns (Valente et al. 2015, 2020; Upham et al. 2019; Hernández-Hernández and Wiens 2020; Tietje et al. 2022).

Island systems have long been at the forefront of evolutionary research (Wallace 1855; Darwin 1859; MacArthur and Wilson 1963; Warren et al. 2015; Gillespie 2016; Patiño et al. 2017; Schrader et al. 2024a). The relatively simple biotas of oceanic islands, their well-delimited boundaries compared to continental systems, and variations in the ages and diversities of colonising lineages simplify our interpretation of complex evolutionary processes (Losos and Ricklefs 2009; Warren et al. 2015; Gillespie 2016). Island biogeography theory predicts that the processes of immigration, speciation, and extinction, as well as the age, area, and isolation of the island system, interact to explain how biodiversity accumulates in insular systems and tends towards an equilibrium (MacArthur and Wilson 1963, 1967; Whittaker et al. 2008). While the island diversity equilibrium hypothesis has been contested (reviewed by Warren et al. 2015 and Borregaard et al. 2017), a global test of MacArthur and Wilson's predictions generally supports their conclusions (Valente et al. 2020). Nonetheless, analyses inferring the processes governing biodiversity patterns across the multiple lineages that make up an island assemblage are still scarce (Rabosky and Glor 2010; Valente et al. 2015; Gillespie et al. 2018; Michielson et al. 2023).

Flowering plants (angiosperms) form one of the most diverse groups on islands, with insular systems harbouring over 60,000 endemic angiosperm species (Schrader et al. 2024a; Roeble et al. 2024). Studies on a global scale are beginning to disentangle how insularity affects the biogeographical processes driving angiosperm diversity (Carvajal-Endara et al. 2017). For example, certain traits or trait combinations are often markedly different in insular species than in their mainland relatives, a pattern known as 'island syndrome' (Burns 2019; Veron et al. 2021; Zizka et al. 2022). One of the most conspicuous island syndromes in angiosperms is insular woodiness (hereafter 'IW'; Carlquist 1974; Lens et al. 2013; Hooft van Huysduynen et al. 2021; Zizka et al. 2022), i.e. the evolutionary transition from herbaceous continental colonisers towards woody species following island colonisation. This phenomenon is a classic feature of insular angiosperms, estimated to have independently evolved at least 175 times globally across the eudicot clade, without a single example in monocots that lack a wood-forming vascular cambium (Zizka et al. 2022). Interestingly, several studies have noted a potential positive link between IW and diversification rates (speciation minus extinction; Carine et al. 2010; Lens et al. 2013; Hooft van Huysduynen et al. 2021), with many iconic island radiations (lineages with ≥ 5 island species resulting from in-situ diversification) featuring IW. Carine et al. (2010) posit that the shift towards IW increases diversification rate in Macaronesian lineages, which was corroborated by other studies on the Canary Islands (Lens et al. 2013; Hooft van Huysduynen et al. 2021). In contrast, Zizka et al. (2022) found that 51% of IW lineages globally contain only one to two species, challenging this IW-diversification hypothesis. Because empirical tests of the effect of IW on diversification to date have been restricted to single clades (Koutroumpa et al. 2021), we present an island community study investigating how IW drives speciation and extinction across multiple lineages. In addition, because woodiness is a trait linked to perenniality, we also assess the role of life span in explaining diversification patterns across Canary Island angiosperms, classifying species as annuals (single-growing season herbaceous species) or perennials (longer-lived herbaceous and woody species).

To investigate the processes governing insular biodiversity patterns, we selected the iconic angiosperm flora of the Canary Islands. The biota of this oceanic archipelago is one of the best studied globally and has a long history of evolutionary research (Carine et al. 2010; Caujapé-Castells et al. 2011; del Arco Aguilar and Rodríguez Delgado 2018; Price et al. 2018). Moreover, the Canarian angiosperm flora has one of the highest levels of species richness of any oceanic island archipelago (>1,300 species, including 225 monocots that mainly belong to lineages composed of a single species, i.e. singletons), harbours marked une-

venness with numerous singleton lineages and many radiations, and has been identified as one of the two major IW hotspots globally (Triantis et al. 2015; Hooft van Huysduynen et al. 2021; Zizka et al. 2022; Cerca et al. 2023; Roeble et al. 2024). Despite the potential of the Canarian flora as a model to study diversity unevenness across lineages, a comprehensive standardised phylogenetic dataset that allows for simultaneously analysing multiple Canary Island angiosperm lineages to test for causes of this unevenness under the same approach is lacking. In fact, such a dataset does not currently exist for any insular flora in the world.

Here, we analyse the processes that shaped Canary Island angiosperm biodiversity patterns in three ways. First, to quantify the number of independent colonisation events, the number of radiations and species per lineage, and how these vary amongst families across all native angiosperm genera of the archipelago, we update the estimated number of independent colonisation events using the phylogenetics literature and new phylogenetic data generated in this study.

Second, to estimate time of island colonisation and speciation for Canary Island lineages, we provide a comprehensive standardised time-calibrated phylogenomics dataset comprising 1,244 newly generated and 501 published DNA sequences for 669 species of non-monocot angiosperms native to the Canary Islands and their close mainland relatives, using hybrid capture target enrichment of the Angiosperms353 bait set (Baker et al. 2021) and 166 fossil calibration points across angiosperms (Ramírez-Barahona et al. 2020). Monocots were excluded from our target group for DNA sequencing because we are primarily interested in the effects of IW on diversification, and monocots are never able to develop a vascular cambium that produces wood (Carlquist 2012; Lens et al. 2013). To generate this phylogenomic dataset, we combined our newly generated sequences with existing published sequences of more distantly related mainland species across angiosperms (both monocots and non-monocots) to account for all the internal nodes that are associated with the fossils used for calibration (Zuntini et al. 2024). Our molecular dataset, therefore, includes Canary Island non-monocots and both monocots and non-monocots from outside the Canary Islands for time calibration purposes. From this phylogenomic dataset, we were able to estimate the time of island colonisation for hundreds of non-monocot Canary Island plant lineages, allowing us to statistically assess the relationship between lineage age and diversity for ~50% of the native angiosperm flora and ~60% of the native Canary non-monocot species.

Third, to estimate colonisation and diversification rates, compare rates of diversification between IW lineages and other lineages, and infer ecological limits to diversity, we fitted a dynamic stochastic model of island biogeogra-

phy (DAISIE, Dynamic Assembly of Island biota through Speciation, Immigration, and Extinction; Valente et al. 2015) to the time-calibrated phylogenomic information. DAISIE allows us to estimate lineage-specific rates of speciation, extinction, and colonisation for the complete Canary Islands angiosperm assemblage, as well as clade-level diversity carrying capacity, i.e., whether rates of cladogenesis and colonisation decline as the number of species increases (Valente et al. 2015, 2020). In addition, we performed sensitivity analyses to study the role of IW and life span on diversification based on Canary Island angiosperms with and without monocots.

By combining these three approaches, we provide an archipelago-wide empirical assessment of evolutionary rates of the Canary Islands angiosperm flora. This allows us to test whether the unevenness in species diversity between lineages can be best explained by the clade age, (trait-dependent) diversification rate, or ecological limits hypothesis, and to investigate how two of the most conspicuous plant traits on islands (i.e., insular woodiness and life span) affect diversification rates on the archipelago.

RESULTS AND DISCUSSION

High number of colonisations shape the Canary Islands' flora

Our updated list of the native Canary Islands angiosperm flora based on published checklists recognises 1,385 species (608 endemics, ~44%), representing 427 genera from 83 families, or 1,320 species across ~400 genera when removing possibly native species (Supplementary Data S1, S2; Beierkuhnlein et al. 2021; BIOTA 2023; Sauerbier et al. 2023). Combining this list with phylogenetic data, we identify 881 'likely' colonisation events of the Canary Islands (minimum: 772, maximum: 968; Supplementary Data S1, see Methods), with the vast majority of colonisations being non-monocot angiosperms (676 likely colonisations, 605 minimum, 752 maximum; Supplementary Data S1). Flowering plant diversity is deeply uneven between lineages, ranging from one to 63 species per Canarian lineage. At least 480 lineages are singletons, 27 have more than five species, and six have 20 or more species (Supplementary Data S1).

Colonisation of the Canary Islands is widespread across the angiosperm tree of life, spanning 33 angiosperm orders (Figure 1, Supplementary Data S1). The families with the highest numbers of colonisations are Asteraceae (249 species, 125 likely colonisations), Poaceae (130 species, 121 likely colonisations), and Fabaceae (130 species, 85 likely colonisations) (Figures 1, 2). Native insular species richness per family is positively correlated with the number of colonisations (Figure 1; Supplementary Data S3; Kendall's rank correlation test,

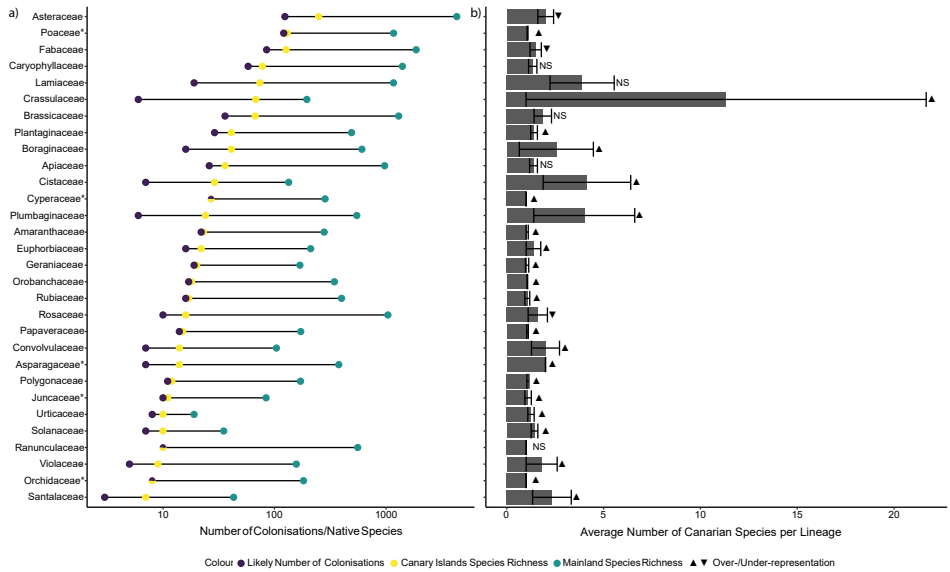


Figure 1. Top 30 most species-rich angiosperm families on the Canary Islands. a) Number of colonisations, total Canary Island species richness and mainland richness per family. Semicircles denote families where the number of colonisations and total Canary Island species richness is equal (e.g., Orchidaceae, Ranunculaceae). b) Average number of native Canarian species (both endemic and non-endemic) per lineage and over-/under-representation relative to the mainland, where error bars indicate standard error. Significant island disharmony in comparison to the mainland is represented by triangles, upwards for over-represented, downwards for under-represented, and NS for not significant disharmony. Families with an asterisk are monocot angiosperms. Values for all families native to the Canaries are given in Supplementary Data S3 and S4.

$\tau = 0.741$, $p < 0.01$), with eight of the ten most species-rich families overlapping with the ten families with the highest number of colonisations (Supplementary Data S3). Native insular species richness per family is also positively correlated with mainland species richness ($\tau = 0.544$, $p < 0.01$), where we use the Mediterranean Basin as our estimated mainland source pool, as this region accounts for 74 of the 83 families native to the Canaries (see Methods). The most diverse mainland families also have the most colonisations of the Canaries ($\tau = 0.505$, $p < 0.01$). Among the ten most species-rich families on the Canaries, only Crassulaceae (whose Canarian diversity is mostly the product of the largest radiation on the archipelago) is neither in the ten most frequently colonising families, nor in the ten most diverse mainland families, reflecting the importance of in-situ diversification in this family. Conversely, families commonly assumed to have a strong dispersal ability have very few native Canarian species per lineage but high native species richness (i.e. resulting from multiple colonisations). For example, Poaceae (130 species; 121 colonisations; 1.07 average native Canarian species per lineage), Cyperaceae (27/27/1), and Amaranthaceae (23/22/1.05) are

represented almost entirely by single-species lineages (singletons) but are in the 15 most species-rich families (Supplementary Data S3).

Our phylogenetically informed estimate of 881 angiosperm colonisations of the Canary Islands (Supplementary Data S1, S2) is over 50% higher than previous estimates for all vascular plants. Using a stricter definition of native insular diversity, only considering species that are 'Surely Native' (BIOTA 2023), Price et al. (2018) estimated 479 colonisations of native vascular plants. When restricting our results to only include 'Surely Native' species from (BIOTA 2023), we estimate 282 angiosperm colonisations (Supplementary Table S1). The drastic difference between the estimated number of likely colonisations and colonisations of 'Surely Native' species reflects the 'Probably Native' species included in our study. Nevertheless, our estimate of in-situ diversification (44%, Supplementary Data S1) is similar to the values obtained by Price et al. (2018) and Schrader et al. (2024a), at 47% and 42%, respectively.

Using binomial tests to assess island disharmony, we identify 22 of the top 30 species-rich angiosperm families as significantly overrepresented on the Canaries relative to the mainland Mediterranean Basin source pool, with only three families (Asteraceae, Fabaceae, and Rosaceae) being underrepresented (Figure 1). Interestingly, these three families are the only ones among the 74 families in total that are under-represented on the Canaries compared to the Mediterranean mainland, a finding that is in line with recent studies on island disharmony that found Asteraceae, Fabaceae, and Rosaceae are consistently underrepresented on islands globally (König et al. 2021; Roeble et al. 2024; Schrader et al. 2024a). Of the remaining 71 families native to the Canaries present in the Mediterranean Basin source pool, 58 are over-represented, and 13 are equally represented (Supplementary Data S4).

A time-calibrated phylogeny for Canary Island non-monocots

To estimate colonisation and branching times across the Canary Islands angiosperm flora, we built a time-calibrated phylogenomic tree of the Canary Islands angiosperms, focusing on the non-monocot flora. We focus on non-monocot angiosperms for DNA sequencing and evolutionary analyses as monocots never produce wood, and to not bias our analyses against the effects of IW on diversification (although we include published monocot sequences associated with fossils from outside the Canaries for calibration of our phylogenetic reconstruction, and we account for the inclusion of Canarian monocots in our diversification analyses via modelling, see below). Our phylogenetic reconstruction is based on 669 native Canarian species (~60% of native Canarian non-monocot angiosperms, ~50% of all native angiosperms) across 683 tips, and 771 closely related mainland taxa of Canarian species that are used as outgroups for each

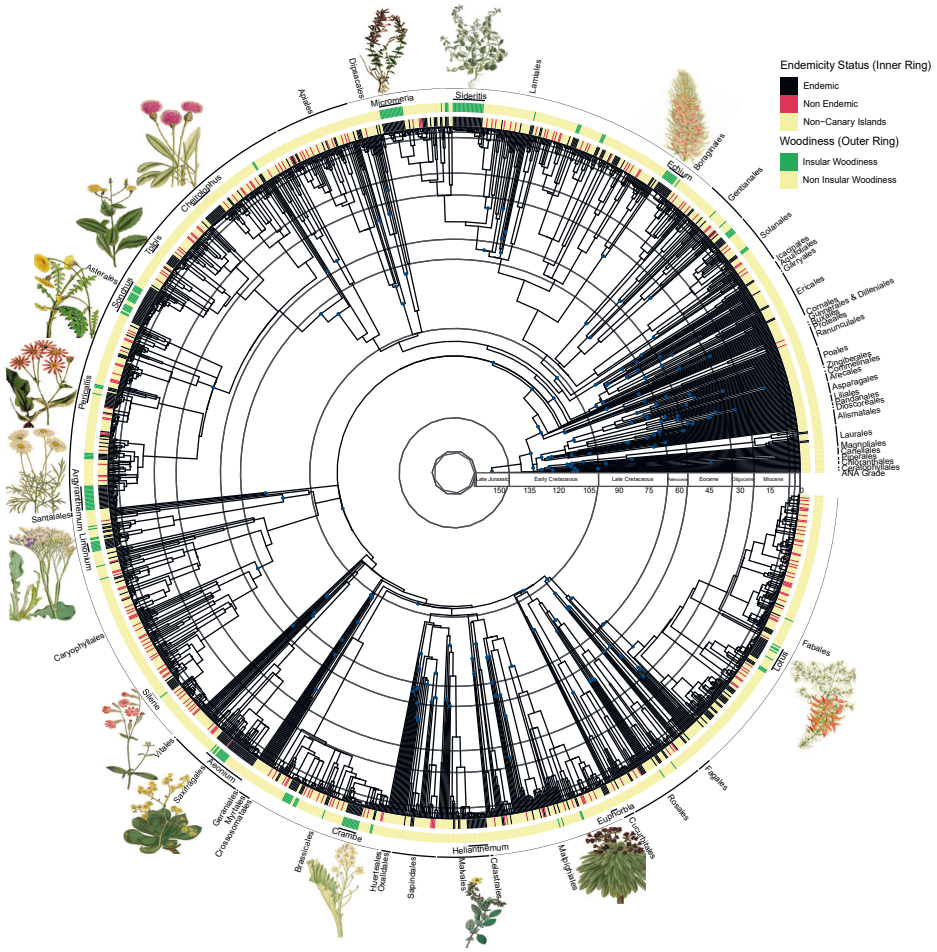


Figure 2. Time-calibrated phylogenetic tree for the Canary Islands non-monocot angiosperms, with the top 15 most species-rich lineages highlighted with illustrations. Taxonomic orders are given outside of the tree. The inner coloured ring denotes Canary Island endemicity status, with mainland relatives and additional taxa required for fossil calibration as 'non-Canary Islands'. The outer ring shows the distribution of insular woodiness across the Canary Islands angiosperm tree. The blue circles at the nodes indicate fossil-calibrated nodes, given in Supplementary Data S5. The grey circles within the tree denote geological periods. For details on the illustrations, see Supplementary Data S6.

Canarian lineage. In total, we sequenced 1,244 new samples for this study, with the remaining tips from previously published studies (Hendriks et al. 2023; Arjona et al. 2025; Martín-Hernanz et al. 2025), adding another 210 tips. We embedded these Canarian lineages into a broader phylogenetic angiosperm framework containing an additional 291 tips (thus adding up to a total of 1,745 tips; (Zuntini et al. 2024)), to enable reconstructing a time-calibrated evolutionary history across all sampled Canarian lineages using the same phylogenetic

framework and fossil calibrations (see 'Sampling Strategy' in Methods; Figure 2, Supplementary Data S2, S5, S7, S8). While our focus is on non-monocots, we do include existing non-Canary Islands monocot sequence data in our tree (59 of the 501 additional tips), but only to assist with time calibration (to take advantage of the many existing monocot fossils outside the Canary Islands). For phylogenetic support values (Felsenstein's bootstrap proportions) for each lineage, see Supplementary Data S9. The 669 native Canarian species and 771 mainland relatives sampled in our tree represent 435 of our 676 estimated non-monocot lineages (~65%) and 186 of the 241 IW Canarian species (77%; Figure 2; Supplementary Data S10).

From our phylogenetic reconstruction, we recover many established relationships across the backbone of angiosperms and within Canarian lineages, alongside peculiarities (Figure 2, Supplementary Data S8). Major lineages previously identified within angiosperms are recovered as monophyletic (magnoliids-Chlorantales, monocots (and the magnoliid-monocot sister relationship recovered in (52)), and eudicots (super-asterids and -rosids)). The main peculiarity that affected the time estimation of a considerable number of Canarian lineages is the position of Caryophyllales (68 sequenced Canarian species, 51 lineages) and Santalales (1 sequenced Canarian species), which are retrieved as more closely related to rosids than asterids (cf. Zuntini et al. 2024 but with Caryophyllales diverging earlier; see Ruhfel et al. 2014, Angiosperm Working Group 2016, or Gitzendanner et al. 2018 for a position closer to asterids rather than rosids). In addition, as with Ruhfel et al. (2014) and Gitzendanner et al. (2018), Vitales are recovered as sister to the rest of the rosids, rather than Saxifragales. As these peculiarities are towards the root of rosids/asterids, the topological differences can affect the colonisation and branching times obtained for Canarian lineages. However, changes in the backbone of our phylogeny do not affect the number of inferred colonisations.

We recover all large, charismatic radiations (defined as lineages with ≥ 5 species resulting from in-situ diversification) within the native non-monocot flora, comprising the most species-rich lineages, as monophyletic (Figure 2, Supplementary Data S8). Unexpectedly, the five endemic *Viola* species are largely recovered as independent lineages, with the only sister relationship retrieved being between *V. palmensis* and *V. odorata*. Interestingly, several native non-endemic Canary Island species are recovered as more closely related to Canarian congeners than to conspecifics from the continent. For example, *Clinopodium menthifolium* and *C. nepeta* specimens from the Canary Islands are recovered as more closely related to each other than to mainland populations of their own species (Supplementary Data S8). This pattern was also found in *Cuscuta, Rubia, Calendula, Filago, Senecio, Hippocrepis,* and *Ononis*. Whilst

this pattern could be caused by multiple factors, such as incomplete lineage sorting or taxonomic gaps resulting from cryptic diversity, it is compatible with the surfing syngameon hypothesis (Caujapé-Castells et al. 2011, 2017), which proposes that intrageneric hybridisation between congeners becomes possible following recurrent island colonisation in phylogenetically closely related taxa previously geographically isolated on the mainland but occurring in sympatry on the island.

Insular woodiness is phylogenetically widespread in the Canary Islands eudicots, occurring in a total of 41 lineages from a diversity of orders and families. However, in terms of species diversity, IW is highly concentrated in a few lineages (Figure 2, outer ring). IW is present in 11 of the 15 most diverse island lineages, which account for 173 of the 241 IW species in the Canaries (72%; Supplementary Data S10, S11). Further, of the 41 lineages containing IW species, 24 have radiated into three or more species, pointing to a possible link between IW and diversification. The *Aeonium* Alliance (*Aeonium, Aichryson, Monanthes*, and former *Greenovia*; Crassulaceae) and *Sideritis* sect. *Marrubiastrum* (Lamiaceae) both have the most IW species in a single lineage, with 27 species each. For comparison, of the remaining 635 lineages without IW species, only 31 have three or more species (Supplementary Data S1, S10). However, there are also 11 IW singleton lineages demonstrating that IW is not always linked to speciation (e.g., *Ixanthus viscosus, Urtica morifolia*; Supplementary Data S11), with these singleton lineages not consistently recovered as the youngest IW lineages (i.e., there are younger IW lineages that have diversified more; see Figure 3 below). These results align with those of Zizka et al. (2022), who found that although IW is widespread throughout non-monocot angiosperms, diversity is clustered within phylogenetically distantly related families.

No evidence for clade-age hypothesis in non-monocot lineages

Based on our angiosperm-wide phylogeny that was time-calibrated with 166 fossils (Supplementary Data S5, see 'Sampling Strategy' in Methods), we estimate both stem (divergence from the closest mainland relative on the tree) and crown ages (earliest branching event within the Canary Islands, applicable only to lineages that have undergone in-situ cladogenetic speciation) of native Canary Island lineages. While there is much discussion about whether stem or crown age is the best proxy for colonisation time (García-Verdugo et al. 2019; Martín-Hernanz et al. 2023), they provide an upper (stem) and lower (crown) bound for the colonisation event of a lineage. For cases where only a stem is available (lineages that have not diversified and are represented by a single tip in our tree), this can be assumed to be the maximum time of colonisation, where there is no lower bound. Based on this approach, we estimated

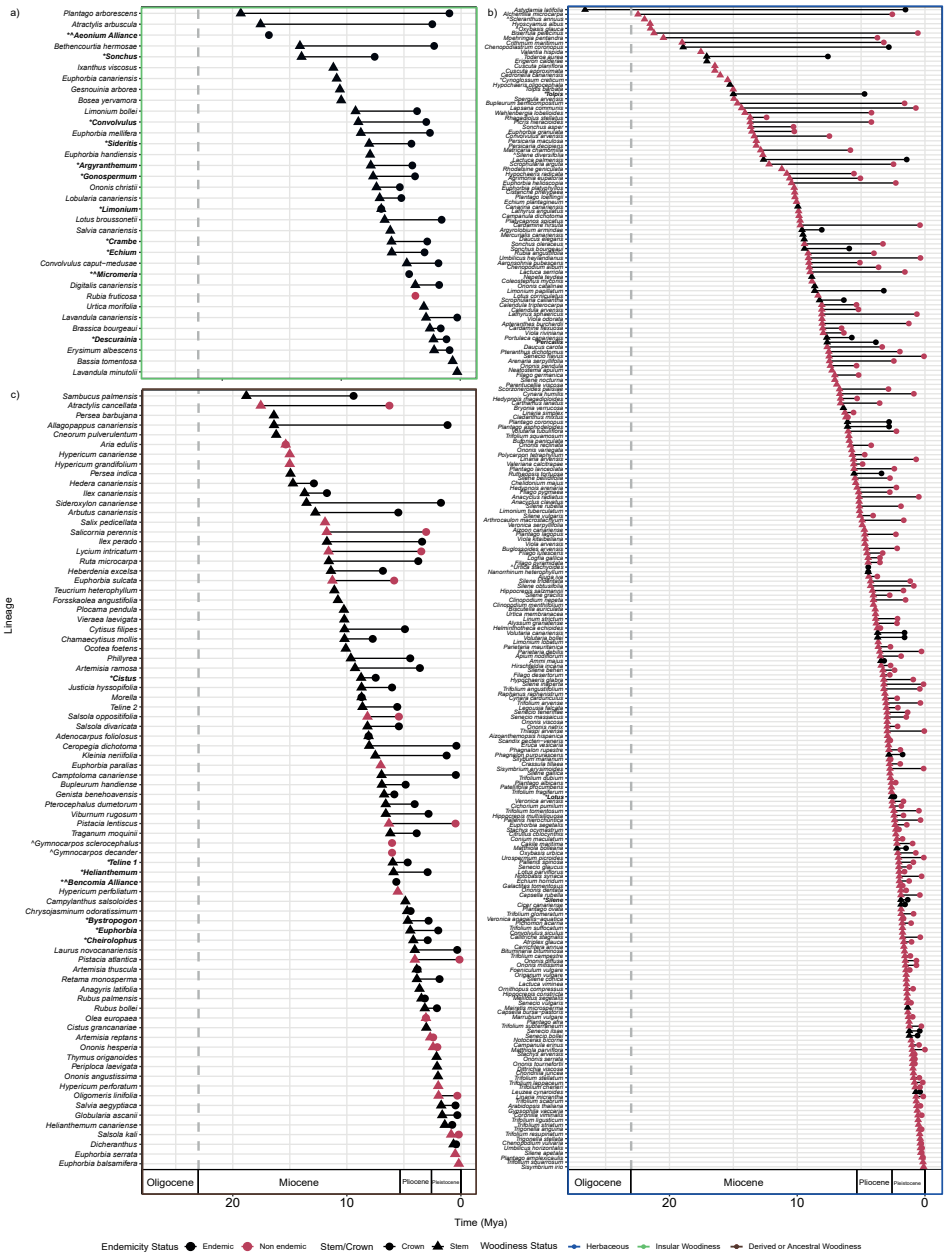


Figure 3. Colonisation intervals of Canary Island non-monocot lineages, classified by growth form, with a) insular woody (green), b) herbaceous (blue), and c) ancestrally and derived woody lineages combined (brown). Emboldened lineages with an asterisk (*) denote radiations (≥ 5 species). Lineages with a caret (^) indicate those for which their mainland relative was not successfully sequenced, so their colonisation time may be an overestimate. If they are singleton lineages, the stem age is given; if they are non-singleton, the crown age is given. The vertical dashed line represents the age of the oldest extant Canary Islands rock (23 My). For lineages recovered as polyphyletic, a conservative estimate of the age of the clade is used (see Supplementary Data S12). *For more legibility, please see the associated preprint. DOI: 10.1101/2025.09.19.677265.*

stem and crown ages for 362 Canary Island lineages, representing 751 species (those from Figures 2, 3 for which we could confidently estimate colonisation time based on sequencing depth); for 289 of these lineages, there is no previously published colonisation estimate (Supplementary Data S12, S13). Across all lineages, the mean stem age is 6.6 ± 0.5 million years (My) and the mean crown age is 4.4 ± 0.4 My (Figure 3, Supplementary Data S12). Confidence intervals for each clade age estimate (crown and stem ages of Canary Island lineages) are given in Supplementary Data S14.

As we are particularly interested in how growth form has shaped the Canarian non-monocot flora, we compared lineage ages between different growth forms. We define growth form as herbaceous (i.e., island lineages that remained herbaceous after the arrival of herbaceous colonisers), ancestrally and derived woody combined (i.e., lineages that colonised the islands as woody species, either because they were always woody, or evolved woodiness on the mainland, hereafter 'ancestrally+derived woody'), or insular woody (i.e., in-situ woodiness evolution on the islands; see (Hooft van Huysduynen et al. 2021) for further explanation of these growth forms). Woody island lineages (both ancestral+derived woody combined, and insular woody) are generally recovered as older than herbaceous lineages (Figure 3, Supplementary Data S12), with the only overlap between insular woody and herbaceous lineages. Mean stem and crown ages of insular woody lineages are 7.7 ± 1.6 and 4.6 ± 1.3 My, respectively (Figure 3a), whereas ancestrally+derived woody lineages combined have mean stem and crown ages of 7.6 ± 1.0 and 5.4 ± 0.9 My (Figure 3c). In comparison, mean stem and crown ages for herbaceous lineages are 5.9 ± 0.6 and 3.9 ± 0.5 My (Figure 3b). Under high extinction rates (see below), these comparatively younger mean ages of herbaceous insular lineages might indicate a potential increased ability of herbaceous species for long-distance dispersal required to colonise the Canaries (see Table 1), as well as the shift towards more open vegetation observed in the Pliocene Western Mediterranean (Jiménez-Moreno et al. 2010, 2013).

Whilst variance with existing estimates is expected, given differences in sampling and dating methodologies used, our colonisation time intervals are consistently older than existing estimates. The mean stem and crown ages for the 73 Canarian angiosperm lineages with published colonisation time estimates are 5.5 ± 1.3 My and 3.0 ± 1.0 My, respectively (García-Verdugo et al. 2019; Martín-Hernanz et al. 2023), while our mean colonisation times for these lineages are 7.8 ± 1.3 and 4.2 ± 1.2 My, respectively (Supplementary Data S13). Nevertheless, our consistently older divergence times overlap at the extreme ends of the confidence intervals and do not radically change inferences on the biogeographic history of the Canary Islands. We find that all divergence times

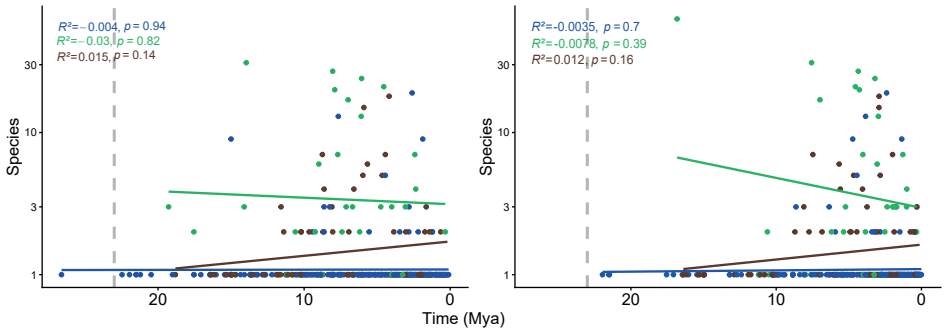


Figure 4. The relationship between lineage species richness and stem (left) and crown (right) age estimates on Canarian non-monocots. The y-axis has a logarithmic scale. The different colours refer to different growth forms: green is insular woodiness (n = 34), blue is herbaceousness (n = 248), and brown is ancestral and derived woodiness combined (n = 80; see Figure 3). The vertical dashed line represents the age of the oldest extant Canary Islands rock (23 My).

estimated here fall within the estimated age of the currently emerged islands of the archipelago (age of Fuerteventura, 23 My (Campeny et al. 2024)), except *Astydamia latifolia* with an estimated stem age of 26.6 ± 9.0 My. Additionally, our estimates of colonisation time intervals overlap with the late Miocene aridification and intensification of drought in the Western Mediterranean, Northern Africa, and Macaronesia, as previously reported (García-Maroto et al. 2009; Jiménez-Moreno et al. 2010; García-Verdugo et al. 2019; Hooft van Huysduynen et al. 2021). These events, along with the onset of the Mediterranean climate and the Plio-Pleistocene glaciations, could have abruptly altered the Canarian climate and likely resulted in extinction events within the Canarian flora (Hooft van Huysduynen et al. 2021), driving taxon turnover. Taken together, these findings contribute to the growing body of evidence that extant Canarian angiosperm diversity likely reflects recent rapid taxon turnover mediated by, among others, high palaeoclimate-driven extinction and high colonisation rates due to nearby continental sources (García-Verdugo et al. 2019).

To test the clade-age hypothesis, we investigated the link between the age of Canary Island lineages and species richness using a linear regression with the age estimates (both stem and crown) from our time-calibrated phylogenetic tree (Figure 4). In contrast to regional and global studies addressing the effect of clade age on species richness (Valente et al. 2013; Henao Diaz et al. 2019; Tietje et al. 2023), we find no significant relationship between lineage age and species richness. These findings are also robust to grouping all growth forms (i.e., no significant relationship between lineage age and species richness; Figure S1), log-transforming the data, removing singleton lineages (both endemic and non-endemic), and accounting for phylogenetic non-independence using phylogenetic least squares regression.

Evidence of ecological limits & differential diversification

To test the diversification rates and ecological limits hypotheses, we applied macroevolutionary dynamic stochastic modelling to our phylogenetic data. The colonisation and branching time data from our time-calibrated phylogeny were input into DAISIE, a phylogenetic model of island biogeography, to estimate biogeographical and diversification rates per lineage (i.e., speciation, extinction, and colonisation) for the flora of the Canary Islands. We fitted a set of 30 DAISIE models to our dataset (see Methods). Briefly, we fitted four models (M1-M4) that assume homogeneous rates of speciation, colonisation and (natural) extinction across the angiosperm flora of the Canaries (background models in Table 1). In addition, to test whether growth form and life span affect diversification rates, we fitted a set of "two-type" models (M5-M30), which allow for a subset of the flora to have evolutionary rates that differ from the background rates for the archipelago (Valente et al. 2015). Specifically, we tested whether lineages that were herbaceous vs woody at the time of colonisation have different diversification rates, whether lineages with predominantly IW species have different rates than lineages without any IW species, and whether lineages with predominantly perennial species have different rates than lineages with mostly annual species (to investigate whether the effects of woodiness and perenniality on diversification may be confounded). We fit DAISIE to both the non-monocot flora, for which we have branching and colonisation time data from our sequencing data, and for the whole angiosperm flora (i.e., monocots and non-monocots). In the latter DAISIE analysis, we included information for monocot lineages by using previously published phylogenetic information to infer the number of lineages and species per lineage and treated branching and colonisation time data as missing information (see Methods for more details).

Table 1 shows the preferred DAISIE models for the background, the herbaceous versus woody analyses, the IW analyses, and the annual versus perennial analyses. Under the preferred background model (M4), diversification in non-monocot angiosperms can be explained by just three parameters: cladogenetic speciation, extinction, and colonisation (Non-Monocot Only section of Table 1). In the M4 model, the system is diversity-independent, that is, we find no evidence that diversity leads to a slowdown in rates of colonisation or speciation, and no detectable upper bound to clade-level diversity on the Canaries. Rather, the model indicates that the system acts as an evolutionary sink - the archipelago needs to receive new colonists from the mainland to maintain diversity, given that the rate of extinction greatly exceeds speciation. The colonisation rate is exceptionally high compared to other studies, at approximately 234 colonisations per million years (My; gamma * mainland pool size;

0.012 * 19,631), further supporting the findings of García-Verdugo et al. (2019) that extant Canarian flora is maintained by high extinction and high colonisation rates, as predicted by the surfing syngameon hypothesis (Caujapé-Castells et al. 2011, 2017). In comparison, Canary Island land birds have been estimated to have 45 colonisations per My (Valente et al. 2017a). Under the background model, the diversity of Canary Islands non-monocot angiosperms is found to be at a dynamic diversity equilibrium (diversity remains relatively constant through time, Figure S2), as was the case in Canary Islands birds (Valente et al. 2017a), in contrast with angiosperm diversity globally (Magallón et al. 2019). Finally, whilst the presence of many lineages with just a single endemic species would suggest that rates of anagenetic speciation (without lineage splitting within the archipelago) are high, the estimated anagenesis rate is surprisingly low. This suggests that the data do not contain a signal on anagenesis and that the presence of singleton endemics can instead be explained by cladogenesis followed by extinction.

We found that two-type models that allow differential rates based on growth form are preferred over the background models. For the herbaceous versus woody non-monocot analysis, the preferred model was M18 (Non-Monocot Only section of Table 1; Supplementary Data S15). Under this model, woody species at the time of colonisation (including ancestrally woody and derived woody, but not IW species) have a lower extinction rate than species that were herbaceous at the time of colonisation. Woody lineages also have lower cladogenetic speciation rates, in agreement with previous angiosperm-wide diversification studies that found that herbaceous species with shorter generation times have higher rates of molecular evolution and therefore higher speciation rates (Smith and Donoghue 2008; Givnish 2010). Notably, however, anagenetic speciation rates are higher in woody lineages under this model. Furthermore, differences in diversity between herbaceous and woody species are explained by pronounced differences in colonisation rates: the rate of colonisation for herbaceous species is estimated to be 0.016, equivalent to 247.2 colonisations per My (0.016 * 19,631 * 0.77), the rate of colonisation for woody species is only 0.00407, equivalent to 18.4 colonisations per My (gamma * mainland pool size * type 2 proportion; 0.00407 * 19,631 * 0.23). Improved colonisation capabilities of herbaceous compared to woody colonisers on oceanic islands have been posited since the time of Darwin (1859), largely due to expected differences in propagule size (Arjona et al. 2018). Generally, small-seeded species are better dispersers than large-seeded species once plant height is accounted for (Thomson et al. 2011; Schrader et al. 2024b), but a differential ability of herbaceous species to colonise islands, relative to woody species, based solely on growth form, remains hitherto untested.

Table 1. Best fitting DAISIE models for non-monocot angiosperms and for all angiosperms (monocots and non-monocots). Results are shown only for the preferred background models and the three preferred differential rate two-type models. λ^c- cladogenetic speciation; μ - extinction rate; *K* - carrying capacity; ɣ - colonisation rate; λ^a - anagenetic speciation; ∞ - infinite. Background rates represent estimated parameter rates for the entire community (background model), the 'herbaceous' type (herbaceous vs woody), the 'non-IW' type (non-IW vs IW), and the 'annual' type (annual vs perennial). "Two-type" model rates represent estimated rates for species of the 'woody' type (herbaceous vs woody), the 'IW' type (non-IW vs IW), and the 'perennial' type (annual vs perennial). 'Prop' is the proportion of the mainland pool of the "two-type" species (woody, herbaceous species belonging to lineages leading to IW, or perennial species) in the "two-type" model. For explanations of how the proportions were calculated, see Methods and Supplementary Data S11. For the full DAISIE analysis results for non-monocots, see Supplementary Data S15- S17, and S18-20 for all angiosperms.

Non-Monocots Only													
		Background Rates					Two-Type Rates						
Model Type	Model	λ^c	μ	K	ɣ	λ^a	λ^c	μ	K	ɣ	λ^a	Prop	df
Background Model	M4	0.285	0.505	∞	0.012	0	-	-	-	-	-	-	3
Herbaceous vs Woody	M18	0.304	0.566	∞	0.016	0	0.216	0.325	∞	0.004	0.148	0.233	6
Non-IW vs IW	M18	0.172	0.473	∞	0.014	0	0.479	0.321	288.711	0.003	1.694	0.071	8
Annual vs Perennial	M30	0.060	0.781	∞	0.069	5.1E-14	0.376	0.498	∞	0.006	5.1E-14	0.8	8
Whole Angiosperm Flora													
		Background Rates					Two-Type Rates						
Model Type	Model	λ^c	μ	K	ɣ	λ^a	λ^c	μ	K	ɣ	λ^a	Prop	df
Background Model	M4	0.226	0.471	∞	0.013	0	-	-	-	-	-	-	3
Herbaceous vs Woody	M18	0.223	0.523	∞	0.017	0	0.217	0.352	∞	0.004	0.116	0.217	8
Non-IW vs IW	M18	0.135	0.451	∞	0.015	0	0.488	0.341	315.731	0.003	0.164	0.058	8
Annual vs Perennial	M30	0.046	0.706	∞	0.072	3.4E-14	0.316	0.472	∞	0.007	3.4E-14	0.82	7

For the non-monocot analysis of IW versus other growth forms, we find that all parameters are different for lineages with the potential to evolve IW under the preferred model (M18; Non-Monocot Only section of Table 1; Supplementary Data S16). Both cladogenetic and anagenetic speciation are substantially higher for IW lineages, whereas extinction is lower. Therefore, the diversification rate of IW lineages is higher than that of the other growth forms. The colonisation rate of IW lineages is also substantially lower than the colonisation rate for non-IW lineages, at 3.78 colonisation events per My (gamma * mainland pool size * type 2 proportion; 0.00274 * 19,631 * 0.0706; Table 1), meaning the diversity of IW lineages on the Canaries is driven by high speciation on the island, rather than the sink dynamics of the background community. This result supports the diversification rate hypothesis to explain differences in diversity between IW lineages and the background community. Furthermore, IW lineages are identified as having diversity-dependence in rates of colonisation and cladogenesis, in support of the ecological limits to diversity hypothesis. In this scenario, the more IW species there are on the island, the lower the rates of colonisation and cladogenesis for IW lineages. This is not the case for non-IW species, for which no diversity-dependence is detected. The causes for the evolution of IW are incompletely understood; drought resistance, favourable climate and reduced herbivory (selecting for longevity), competition for light (thereby selecting for increasing height), and potentially burial avoidance following volcanism are all suggested global IW drivers that may interact with each other in a different way for each lineage (Carlquist 1974; Zizka et al. 2022; Beierkuhnlein et al. 2023).

For the annual versus perennial non-monocot analysis, we find the only parameters that vary for perennial lineages are cladogenetic speciation, extinction, and colonisation rates (Non-Monocot Only section of Table 1; Supplementary Data S17). The preferred model was M30, with perennial lineages having a substantially higher cladogenetic speciation rate and lower extinction rate (therefore, they have a higher rate of diversification). Colonisation rate of perennial lineages is 96.7 colonisation events per My (0.0062 * 19,631 * 0.8), substantially lower than that of annual lineages, at 269.7 colonisation events per My. The diversity of perennial lineages is therefore also maintained by in-situ diversification, rather than the sink dynamics of annual species. As 19 of the 20 most species-rich lineages in the Canaries primarily consist of perennial species (Supplementary Data S1, S12), the differential diversification rates recovered here are to be expected. Interestingly, previous studies are equivocal, with no clear effect of annuality or perenniality on diversification across taxonomic groups and scales (Helmstetter et al. 2023). As IW and perennial lineages are estimated to have similar effects on diversification by promoting speciation

rates and reducing extinction rates (see 'two-type' rates of 'Non-IW vs IW' and 'Annual vs Perennial' in the Non-Monocot Only section of Table 1), it is possible that the effects of these traits on diversification could be confounded. However, given the dissimilarity in the estimated evolutionary rates between these analyses (see the Non-Monocot Only section of Table 1), this suggests that the effects of woodiness and perenniality on diversification in the Canary Islands are somewhat independent. Under the preferred models for all three non-monocot growth form analyses, the inference of equilibrium in Canarian non-monocot angiosperms was unchanged (Figure S2).

Next, to assess whether or not considering monocots in the DAISIE modelling affected our inference, we reran DAISIE including monocot lineages (see Methods). Overall, our results remain unchanged when monocot lineages are included in the DAISIE analysis (see 'Whole Angiosperm Flora' section of Table 1). IW lineages still have faster speciation rates (across both cladogenesis and anagenesis), lower extinction and colonisation rates than the background community, and the best fitting model choice remains the same (M18; Supplementary Data S18). Similarly, perennial lineages still have faster cladogenetic speciation rates, lower extinction rates, and lower colonisation rates, with anagenetic speciation rates the same as annual lineages. The best fitting model choice is also the same as when monocots are not included (M30; Supplementary Data S19). For the whole angiosperm herbaceous versus woody analysis, model M18 is still preferred, with woody lineages (at the time of colonisation) having higher net speciation rates, lower extinction, and lower colonisation rates, confirming our non-monocot-only results (Supplementary Data S20). The inference of equilibrium in Canary Islands angiosperms was also found when monocots are included, except in the IW versus other growth forms analysis, where diversity is slowly increasing (Figure S3).

Finally, to assess how well the estimated parameters fit our data, we generated goodness of fit plots by simulating 500 datasets using the best fitting parameters for each analysis. We then compared summary statistics (total number of species, total number of colonisations, largest clade size, number of endemic clades, number of endemic species, number of endemic singleton lineages, number of type 1 species (i.e., herbaceous type, non-IW type, and annual type), and number of type 2 species) in the simulated datasets to our empirical dataset (Supplementary Data S21). Figure 5 shows the goodness of fit plots for each non-monocot analysis (for the whole angiosperm analyses, see Supplementary Figure S4). Generally, whilst the total number of species and colonisations is well estimated, the fit is poor for the other summary statistics. We expect these poor fits are caused by the combination of three factors: 1) DAISIE does not account for observed trait heterogeneity within lineages

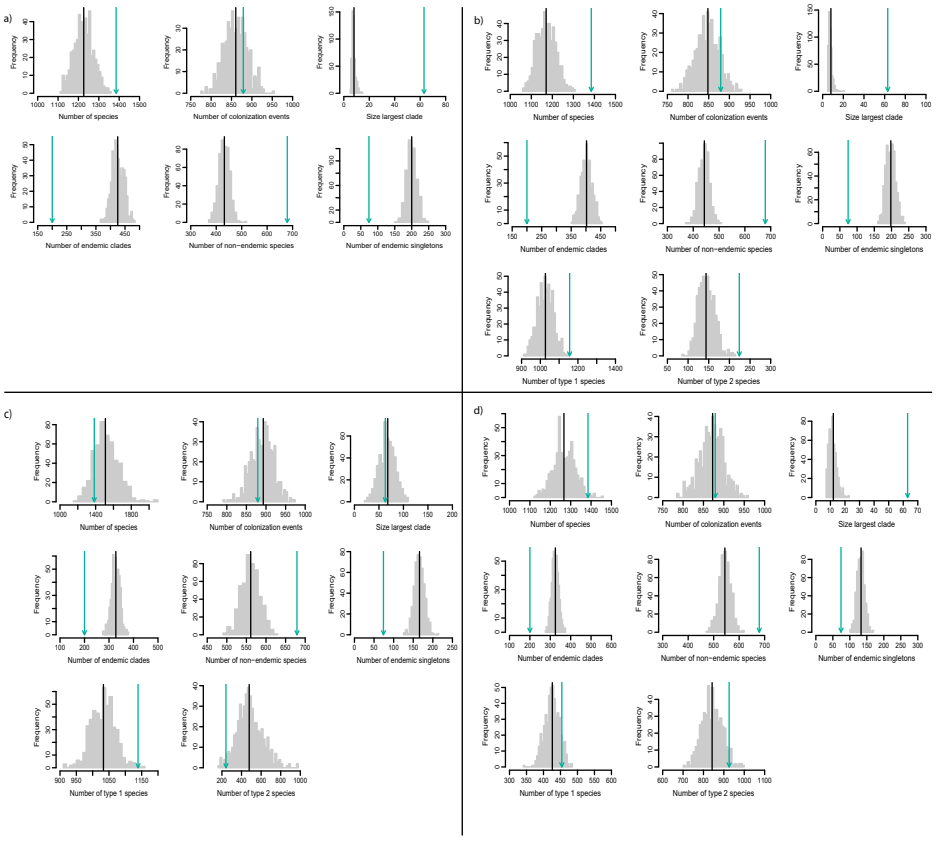


Figure 5. Goodness of fit plots for the non-monocot only DAISIE analyses, where 500 datasets were simulated using the best fitting parameters from a) the Background M4 rates, b) the Herbaceous vs Woody M18 rates, c) the Non-IW vs IW M18 rates, and d) the Annual vs Perennial M30 rates from Table 1. The green arrow shows the empirical value from Supplementary Data S21, and the black bar shows the median value of the simulated datasets.

(i.e., all species within a lineage are assumed to have the same trait state); 2) the dramatic diversity unevenness observed between lineages, where there are very many singleton lineages but also several large radiations with more than 20 species, is difficult to explain with a single set of rates (Supplementary Data S1); and 3), the effect of missing data, which has been shown to increase error in the rates inferred from DAISIE (Dehayem Nanwou et al. 2025). The relatively poor fit for some of the summary statistics does not mean, however, that our DAISIE results are incorrect, as they match with previously published results for the Canaries as explained above and are in line with trait variation among the top 20 most species rich genera; instead it means that our DAISIE results do not tell the full story yet, and we need to develop more complex trait-based models for a fuller understanding of its macroevolutionary assembly.

In conclusion, using our phylogenomic data, we have shown that of the three main global hypotheses to explain global biodiversity patterns, there is strong evidence for the (trait-dependent) diversification rate hypothesis and some evidence for the ecological limits hypothesis. First, we find that the non-monocot flora of the Canary Islands is shaped by high extinction and colonisation rates (Table 1) and that there is no relationship between lineage age and diversity, supporting a scenario with a high turnover of lineages (Caujapé-Castells, et al. 2017; García-Verdugo et al. 2019), and therefore rejecting the clade age hypothesis (Figure 4). Second, IW, woody, and perennial lineages behave distinctly compared to the total archipelago-wide flora of the Canary Islands, with higher speciation and lower extinction rates and therefore higher diversification rates, than the rest of the angiosperm flora, supporting the diversification rate hypothesis to explain unevenness in diversity (Table 1). Third, IW angiosperm lineages have a detectable carrying capacity, meaning that as the number of IW species increases, rates of cladogenesis and colonisation decline (Table 1), in line with the ecological limits hypothesis. Finally, most models show that the Canarian angiosperm flora have achieved a macroevolutionary equilibrium in terms of species richness (Table 1, Figure S2), except in the whole angiosperm non-IW versus IW analysis (Figure S3). With our time-calibrated phylogeny, many new avenues of research into the Canary Islands angiosperm flora become possible, including more focus on trait(-dependent) evolution (Helmstetter et al. 2023), island-wide vs clade-specific diversification (Etienne et al. 2023), and evolutionary return time (Valente et al. 2017b; Michielson et al. 2023), whilst the archipelago checklist identifies the remaining lineages to be sequenced in order to complete the tree of life for this iconic flora (Supplementary Data S1).

METHODOLOGY

Canary Islands as a Model System

The Canary Islands are an archipelago comprised of seven main islands (El Hierro, Fuerteventura, Gran Canaria, La Gomera, La Palma, Lanzarote, and Tenerife), several smaller islets under 30 km2 of the 'Chinijo' archipelago (located to the northeast of Lanzarote, including La Graciosa and Alegranza), and Lobos Island between Fuerteventura and Lanzarote (4 km2). The oldest extant exposed rock emerged ca. 23 million years (My) ago (Campeny et al. 2024), with many more volcanic islands previously present over the last 70 My, which have since become submerged (Fernández-Palacios et al. 2011; Arjona et al. 2018; Beierkuhnlein et al. 2021). Located in the Macaronesian meta-archi-

pelago off the northwest coast of Africa, at the easternmost point, the Canary Islands lie only approximately 100 km off the coast of Western Sahara.

Alongside Hawai'i and the Galápagos Islands, the Canary Islands form one of the best-studied oceanic island systems (Florencio et al. 2021), and the native angiosperm flora is among the best-studied island floras globally (Caujapé-Castells et al. 2011; Price et al. 2018). With an estimated ~1,300 native species of angiosperms (Arjona et al. 2018; Beierkuhnlein et al. 2021; BIOTA 2023), the flora is the most species-rich of any oceanic archipelago (Triantis et al. 2018), contains many evolutionary radiations (Cerca et al. 2023), and holds the highest number of insular woody species and shifts towards insular woodiness globally (Lens et al. 2013; Hooft van Huysduynen et al. 2021; Zizka et al. 2022).

Checklist of native Canary Island angiosperm species

We compiled a list of all known native Canary Island angiosperms based on the most up-to-date published checklists (Beierkuhnlein et al. 2021; BIOTA 2023). From these, we included all species in the following categories: Surely Native and Possibly Native as defined in Beierkuhnlein et al. (2021); Nativo Seguro [Native with Certainty] and Nativo Probable [Probably Native] as defined in BIOTA (2023); duplicate species were subsequently removed. We further refined the resulting list using expert knowledge to remove species that were deemed to have been erroneously considered native (Brandes and Fritzsch 2002; Beierkuhnlein et al. 2021; BIOTA 2023; Sauerbier et al. 2023; Supplementary Data S1). For each species, we recorded their endemicity status (endemic (occurring only in the Canary Islands) or non-endemic (occurring natively in the Canary Islands and elsewhere)) and their growth form (annual herbaceous, perennial herbaceous, insular woodiness, derived woodiness, or ancestral woodiness) using published literature and floras (Kattge et al. 2020; Weigelt et al. 2020; Díaz et al. 2022; Hanz et al. 2022; Barajas Barbosa et al. 2023; Denelle et al. 2023; Sauerbier et al. 2023). Biennial herbaceous species were reclassified to 'perennial herbaceous' to give an annual and non-annual herbaceous category. For the specific reference used for each species, see Supplementary Data S10.

Estimating angiosperm colonisations of the Canary Islands

Using our updated checklist of Canary Island angiosperms, we estimated the number of times each natively occurring genus colonised the Canary Islands and gave a written description for each inferred colonisation event. This step was essential for designing the sampling strategy for our phylogenetic reconstruction (see 'Sampling Strategy for our Phylogenetic Tree' below), but was also used to compute metrics for lineages that were not included in the phylogeny (see 'Statistical Analysis of Colonisation Data' below). Whenever prior

phylogenetic information was available, we used phylogenetic trees to infer the closest mainland relative of each native Canary Island angiosperm lineage. For a list of the phylogenies used for each genus, see Supplementary Data S1, and for a list of the mainland relatives used per Canary Islands lineage, see Supplementary Data S2. A genus was considered to have colonised the Canary Islands several times if it had multiple species native to the Canaries which do not cluster together in the phylogeny. For example, if two Canarian species are more closely related to different mainland species than each other, they were counted as two colonisations. It is important to note that our method does not account for multiple colonisations of the Canary Islands by the same species. While these are known to occur (Caujapé-Castells et al. 2011, 2017; Valtuena et al. 2016), this information is not available for most species and can be difficult to infer from molecular data. Therefore, we chose to focus on colonisations at the genus level. Our final estimate is, therefore, an underestimate of the total number of angiosperm colonisations on the Canary Islands, but a good estimate of the number of colonisations that gave rise to extant Canary Island lineages.

For genera for which no well-sampled phylogenies were available (e.g. did not include all Canarian species, or did not sample mainland species presumed to be close relatives), we applied a measure of uncertainty for the number of colonisations. A minimum bound counted how many distinct colonisations were possible to estimate using the available literature, and a maximum bound assumed that every species colonised the Canary Islands separately. A 'likely' number of colonisations for each genus was also estimated using available phylogenetic literature, distribution data, and the endemicity status of the native Canary Island species. For example, if a genus had multiple native Canarian species, with several endemics and several island species that are also widespread on the continent, the endemic species were assumed to result from a single colonisation in the 'likely' estimate and the non-endemics were considered to result from a single colonisation each. Our final estimate of Canary Island angiosperm colonisations thus consisted of a minimum, likely, and maximum number of colonisations. For full explanations for each inferred colonisation event, see Supplementary Data S1, 'Inference Notes'.

Statistical Analysis of Colonisation Data

From the previous steps, we obtained the total number of Canary Island species and likely colonisations for each family of flowering plants represented in the Canary Island flora. We assessed the relationships between the native mainland species richness, the number of Canary Island colonisations for each family, and the number of native Canary Island species. We downloaded native

angiosperm species diversity data for the countries comprising the Mediterranean Basin and Macaronesian enclave (including the Mediterranean islands) from the World Checklist of Vascular Plants using Version 12 (28 September 2023) from the R package 'rWCVPdata' (Brown et al. 2023). Though not all native species and families on the Canaries have colonised from the Mediterranean Basin, 74 of 83 families occurring on the Canary Islands are present in the countries that comprise the Mediterranean Basin. These 74 families comprise 1,358 (possibly) native species on the Canaries (Supplementary Data S1), meaning we capture a nearly complete representation of Canary Island species using the Mediterranean Basin as the comparative mainland. Additionally, the Canarian flora has a close affinity with the Mediterranean flora (Vargas 2020), and as the limits of the Mediterranean Basin fluctuate over time, this explains the close affinity of these floras, justifying the use of the Mediterranean Basin as our choice of mainland system.

We assessed correlations between Canary Islands angiosperm richness, number of colonisations, and native mainland species richness at the family level. All data were log-transformed before computing Kendall's rank correlation coefficient, as the untransformed data were not normally distributed. As the subsequent dataset contained many taxonomic groups with < 5 species present on either the Canary Islands or the mainland, these taxa were removed to avoid biasing the data. All reported values are subsequently given for the filtered dataset containing ≥ 5 species, although these relationships were identical to those using the full dataset. Island disharmony in comparison to the Mediterranean mainland was assessed using binomial tests to determine whether family-level native angiosperm richness on the Canary Islands was significantly under-/over-represented relative to the Mediterranean Basin (Roeble et al. 2024). Binomial tests were run for each family using the 'binom_test()' function in R, where x represents the total number of native angiosperm species, n represents the number of native mainland species, and p represents the proportion of island species relative to the total number of angiosperms native to the Mediterranean Basin (Supplementary Data S4).

Sampling Strategy for Our Phylogenetic Tree

In this section, we describe the sampling of plant specimens used for our time-calibrated phylogeny. A fully sampled tree would include all 1,385 species of Canary Island angiosperms as well as representatives of each of their closest extant mainland relatives, which would allow estimating stem and crown ages for all Canary Island angiosperm lineages. However, such a high level of sampling is currently not feasible using our approach because, for many Canary Island lineages, it is currently not possible to infer their closest mainland rel-

ative based on existing phylogenetic literature (e.g., Canary Island species belonging to genera that are very diverse on the continent, and for which no well-sampled published phylogeny is available). Based on our updated phylogenetically informed checklist of Canary Island angiosperms, we were able to infer the number of colonisations and identify the closest mainland relatives for the majority of Canary Island lineages. We used this information to guide our sampling for the new phylogenomic dataset produced for this study.

For the trait-dependent diversification analysis, we focused on non-monocot angiosperms because monocots never produce a wood cylinder due to the absence of a vascular cambium. In the few monocot lineages developing secondary growth, a monocot cambium or 'master cambium' is formed that is not equivalent to the vascular cambium, because it produces different tissues other than wood and secondary phloem (Carlquist 2012; Lens et al. 2013). Hence, as we are interested in the effects of IW on diversification, monocot lineages were removed from the sample so as not to bias our analyses against the presence of woodiness. Additionally, as monocots represent a distinct monophyletic group within angiosperms, including them may lead to incorrectly inferred differences in diversification rates: differences in diversification rates between woody and herbaceous species may simply reflect that they are a distinct monophyletic group with their own evolutionary rates different from those of the non-monocots, regardless of their woodiness status, and hence this would not answer whether woodiness leads to faster diversification on the Canaries. As a sensitivity analysis, to assess the effect of not considering monocots, we also repeated our trait-dependent diversification analyses, including monocots using the 'missing data' modelling options available in DAISIE (see below). Our sampling group for the new phylogeny thus consisted of every native, non-monocot Canary Islands angiosperm species for which we could identify the nearest mainland ancestor and the number of colonisations could be estimated (see Supplementary Data S1). This consisted of 838 native Canarian angiosperm species (~61% of the total native Canary Islands angiosperm species) comprising 437 colonisation events (50%), and two outgroup species (the most closely related extant mainland taxon) for every native Canarian colonisation event, totalling 1,710 specimens to sample for this study (Supplementary Data S2). A recent simulation study has shown that this sampling fraction and our approach, by including the majority of species-rich lineages (24/30 (80%) of the most species-rich angiosperm genera, or 25/30 (83%) of the most species-rich non-monocot lineages), leads to very little loss of information compared to complete sampling using DAISIE to estimate biogeographical parameters (Dehayem Nanwou et al. 2025). Our sampling was consistent with existing studies interested in obtaining colonisation estimates (Martín-Hernanz et al. 2023),

consisting of one individual per native Canary Island species and two outgroup species per lineage. For the outgroup species, we sampled the two most closely related mainland species for Canary Islands endemics, and for non-endemic, native species, we sampled two different nearby mainland populations.

To generate our phylogeny, we used a combination of existing Angiosperms353 sequence data and newly generated sequences from fresh, herbarium, and DNA-bank material using capture-based hybridisation sequencing. For this study, we obtained dried leaf material specimens from 23 herbaria and botanical gardens throughout Eurasia and the Canary Islands, as well as from colleagues collecting field samples (Supplementary Data S22). We generated new sequences for 1,244 specimens (495 native Canarian species and 749 outgroup taxa; Supplementary Data S2) of the 1,710 target specimens, representing 735 unique species. Previously published sequence data were also obtained for 210 specimens: 25 *Helianthemum* (Martín-Hernan et al. 2025), 22 Brassicaceae (Hendriks et al. 2023, unpublished data), and 163 Asteraceae (Arjona et al. 2025) from co-authors. To maximise the number of nodes that could be calibrated using fossils in our analysis, we complemented our dataset with publicly available data from the Plant and Fungal Tree of Life Project (PAFTOL; 291 samples; Baker et al. 2022; Zuntini et al. 2024), for a total of 1,745 tips. Seven gymnosperm species were used to form the outgroup, one from each recognised order, except Welwitschiales. See Supplementary Data S22 for all specimens included in the study.

DNA Extraction, Target Capture, and Sequencing

For all herbarium or field collected samples, we extracted genomic DNA from an average of 20 mg of leaf material (or branch, stem, and/or flowers if no leaf material was available or if the species did not produce leaves) using the DNeasy PowerPlant Pro Kit (Qiagen, Hilden, Germany), following manufacturers protocols but repeating the final elution in two x 60 ul elutions rather than one 120 ul elution. DNA concentrations were quantified using the Agilent 5300 Fragment Analyzer with the High Sensitivity Genomic DNA Analysis Kit (DNF-488; Agilent Technologies, Santa Clara, CA, USA). Genomic DNA was stored at the Naturalis Biodiversity Center, Leiden, The Netherlands.

DNA extracts with an average fragment length > 400 bp were sheared using sonication in an M220 Focused-ultrasonicator (Covaris, Woburn, Massachusetts, USA) to obtain an average fragment length of 350 bp. Genomic libraries were prepared using the NEBNext Ultra II DNA Library Prep Kit (New England Biolabs, Ipswich, Massachusetts, USA). This process consisted of end repair of sheared or fragmented DNA, adaptor ligation, size selection, and PCR enrichment of adaptor-ligated DNA over 8-14 cycles. The quality of the libraries was

then assessed using the Agilent 5300 Fragment Analyzer HS Small Fragment DNF-477 kit. Indexing was conducted using 384 unique combinations of IDT10 primers (Integrated DNA Technologies, Coralville, Iowa, USA), following the methodology of Hendriks et al. (2021).

The libraries were pooled into groups of 7-40 specimens based on total DNA quantity (ng), fragment length distribution, and taxonomy (to the order level). Pools were enriched using the Angiosperms353 bait set (Arbor Biosciences myBaits® Target Sequence Capture Kit, Arbor Biosciences, Ann Arbor, Michigan, USA; hereafter 'A353'; Johnson et al. 2019), targeting 353 genes, following the manufacturer's recommendations. An experimental pool of 62 specimens was also enriched using the A353 bait set to test the efficacy of enriching large pools of specimens distantly related by taxonomy. Hybridisation was performed at 65°C for 24 hours for the first 176 specimens, but was reduced to 62°C for the remaining specimens, as we found that this increased the percentage of loci on target recovered. Post-hybridisation libraries were amplified using KAPA HiFi HotStart Ready Mix (Roche Sequencing) for 13-32 cycles, followed by a bead clean-up. Amplified libraries were then 'spiked' with 0.2 ul of unenriched library pool (pre-capture pools). To reduce sequencing costs, amplified libraries were then pooled equimolarly into 19 endpools, avoiding duplicated index combinations. Sequencing was performed using a NovaSeq 6000 sequencer (Illumina, San Diego, California, USA) at BaseClear, the Netherlands, producing 150 bp paired-end reads at 100x targeted coverage.

Sequence Assembly and Paralog Filtering

Raw, demultiplexed sequence data, in the form of FASTQ files, were quality-controlled and trimmed using Trimmomatic v0.39 (Bolger et al. 2014), using the settings of Baker et al. (2022). Trimmed reads were then mapped using HybPiper version 2.1.6 using the 'mega353' target file (Johnson et al. 2016; McLay et al. 2021). The BWA algorithm was preferentially used over BLASTX as this resulted in more reads mapping for our samples (Camacho et al. 2009; Li and Durbin 2009, but see Murphy et al. (2020)). Subsequently, each gene was assembled using SPAdes version 3.15.3 and coding sequences were extracted using Exonerate version 2.4.0 (Slater and Birney 2005; Bankevich et al. 2012).

Three summary functions provided by HybPiper were run: 'hybpiper stats', 'retrieve_sequences', and 'retrieve_paralogs'. HybPiper stats provides summary statistics related to mapping success, including the percentage of reads on target, the number of genes with sequences, and the number of genes with sequences of 25-150% of the target length. This was run for all samples to gauge the success of mapping and contributed to whether a specimen needed to be resampled (see Supplementary Data S22 for HybPiper stats summary of the

final specimens used in this study). Retrieve_sequences pulls the best (unfiltered) contig sequence fasta files (in .FNA format) for each locus, regardless of the paralogs warning. Similarly, retrieve_paralogs pulls the mapped sequence fasta files flagged with a paralog warning; if no paralog warning was detected, the best contig is retrieved from the HybPiper output. Phylogenetic analyses including paralogous genes (homologous genes that share a common ancestor as a product of gene duplication) can produce incorrect topologies (Morales-Briones et al. 2022; Frost et al. 2024). Given the prevalence of paralogs in many of the families included in our dataset (i.e., Brassicaceae and Asteraceae; Herrando-Moraira et al. 2018; Hendriks et al. 2023), and the presence of several groups which are known to hybridise in our dataset (thereby increasing the likelihood of paralogs being present via allopolyploidy; Curto et al. 2018; White et al. 2020; Morales-Briones et al. 2022), paralogs were filtered to assess the effect on tree resolution and branch support.

Once the mapped contigs and paralogs had been retrieved, we used ParaGone to filter putative paralogs (Morales-Briones et al. 2022; Yang and Smith 2014). To identify the best-fitting ParaGone algorithm for our dataset, each algorithm was run (Maximum Inclusion (hereafter 'MI'), Rooted Ingroups ('RT'), and Monophyletic Outgroups ('MO'); see Yang and Smith (2014) for further details), and phylogenetic reconstruction was performed using the orthologs output from each algorithm (see below).

Gene Alignments, Tree Reconstructions, and Time Calibration

Once the orthologs had been retrieved from ParaGone, gene alignments were generated separately for each gene using OMM_MACSE (Ranwez et al. 2018; Scornavacca et al. 2019). OMM_MACSE was used over MACSE for its ability to better handle large numbers of sequences whilst still generating codon-aware alignments, by utilising external alignment software that does not produce codon-aware alignments, such as MAFFT. Aligned genes were then trimmed using TrimAL v1.3 (Capella-Gutiérrez et al. 2009), with the parameters resoverlap 0.75, seqoverlap 0.64, and gt 0.70. Gene trees were generated using maximum likelihood methods from the trimmed alignments using ultrafast bootstrapping in IQ-TREE v.2.2.2.7 (Hoang et al. 2018; Minh et al. 2020). These steps were repeated, using the same parameters, for the unfiltered contigs obtained from the retrieve_sequences function.

We used three approaches to compare the effect of paralog filtering on phylogenomic tree reconstruction. First, we used the gene trees from IQ-TREE generated from the unfiltered contigs to build multi-species coalescent species trees using ASTRAL-IV v1.18.4.5 and ASTRAL-Pro3 v1.18.3.5 (Zhang and Mirarab 2022a, 2022b). A preliminary species tree was produced using ASTRAL-IV

to identify specimens with obviously erroneous placement (i.e., recovered completely incorrectly due to specimen misidentification, contamination, or very poor-quality input data); these specimens were then resampled (390 specimens). The wet lab, sequence assembly, and gene alignment steps were repeated for resampled specimens and a new non-paralog filtered multi-species coalescent species tree was produced via ASTRAL-IV to re-assess erroneous placement. Any specimens identified as erroneous after resampling were removed from our dataset to avoid incorrect inference in downstream analyses and the ASTRAL-IV tree was rebuilt (153 specimens). This process was conducted for both the non-paralog-filtered and paralog-filtered datasets.

Second, orthologous gene trees produced by each ParaGone algorithm were used to reconstruct multispecies coalescent species trees in both ASTRAL-IV and ASTRAL-Pro3. As each ParaGone algorithm produces a different number of orthologous groups, and resultantly, produces different topologies and node support values (Jackson et al. 2023), we assessed the relative performance of each approach and compared the produced trees with those of the unfiltered ASTRAL-IV and ASTRAL-Pro3 trees. As the MI and RT algorithms produced too many orthologous groups to feasibly generate a species tree from (at 22,184 and 5,612, respectively), the MO algorithm was used to produce the paralog-filtered tree (which produced 111 orthologous groups). For our dataset, as paralogy resolution resulted in many more polytomies (see Supplementary Data S23), we proceeded with the non-paralog filtered ASTRAL-IV tree for the time calibration.

Third, we used a maximum likelihood (ML) supermatrix approach in RAxML-NG v1.2.1 to construct a phylogeny tree with branch lengths that could be used for divergence-time estimation (Kozlov et al. 2019). We adopted a 'gene shopping' approach, implemented in SortaDate (Smith et al. 2018), where gene trees were filtered first based on similarity to the species tree (the non-paralog filtered ASTRAL-IV tree), second by clock-likeness (root-to-tip variance), and then by tree length. Gene tree concordance was favoured as high gene tree conflict can lead to errors in divergence estimates (Carruthers et al. 2022). The best 25 loci were identified and used for the subsequent dating analysis. The trimmed alignments for these genes were concatenated using CatSequences (https://github.com/ChrisCreevey/catsequences) and RAxML-NG was run using the concatenated supermatrix with a GTR+G model, with 1,000 bootstraps, constraining the topology to the non-paralog filtered ASTRAL-IV tree. Divergence times were estimated using the relative rate framework implemented in RelTime for its increased accuracy relative to TreePL (Barba-Montoya et al. 2021; Costa et al. 2022), especially in the presence of convergent substitution rate shifts among lineages, which we expect within our dataset. Fossil calibrations

were based on 163 unique minimum age calibrations from the AngioCal v.1.1 dataset, two Asteraceae-specific minimum age calibrations, and an updated age for crown Eudicots (Ramírez-Barahona et al. 2020; Palazzesi et al. 2022; Zuntini et al. 2024; Smith and Beaulieu 2024; Supplementary Data S5). A maximum constraint of 154 million years was used at the crown node of angiosperms, as the resulting age estimates were found to match existing divergence estimates of many angiosperm families (Zuntini et al. 2024).

DAISIE Analysis

We used DAISIE to estimate evolutionary and biogeographical rates of Canarian non-monocot angiosperms (Valente et al. 2015). DAISIE utilises phylogenetic information to estimate per lineage rates of colonisation (of an island system), cladogenetic speciation (speciation with lineage splitting within the island, e.g. leading to radiations), anagenetic speciation (evolution of new endemic species with no net lineage splitting, does not lead to increased diversity on the island), and extinction, as well as diversity limits of a system (Valente et al. 2015, 2020). Estimating extinction rates from phylogenies reconstructed with only extant species using birth-death models has been widely debated (Rabosky 2010, 2016; Beaulieu and O'Meara 2015; Louca and Pennell 2021). DAISIE estimates extinction rates differently from other phylogenetic methods not only because it uses a diversity-dependent model (Etienne et al. 2012) (rather than the constant-rate birth-death model as in Rabosky (2010)), but also because it harvests information simultaneously from multiple island colonist lineages rather than from a single phylogenetic tree. Previous simulation studies using DAISIE have shown that extinction rates can be estimated with relatively high precision and little bias (Valente et al. 2015, 2017a, 2017b; Hauffe et al. 2020).

A DAISIE analysis is centred on a focal island assemblage, in our case all species and lineages of non-monocot angiosperms that occur natively in the Canary Islands. For each lineage, DAISIE requires its endemicity status, number of species, as well as its colonisation time and branching times (if any, in case of radiations). Note that the entire focal group needs to be considered, so both sampled and unsampled taxa need to be included. For all lineages sampled in our phylogeny, we extracted their information from our species checklist (endemicity, number of species) and the time-calibrated phylogeny (divergence times). Stem ages for these lineages were manually extracted from the time-calibrated ML tree, along with all branching times within Canary Island lineages. We treated these stem ages as precise ages of colonisation in the DAISIE analyses for endemic lineages, though these stem ages may better reflect divergence times than colonisation times in lineages with high insular extinction rates. For non-endemic lineages, these stem ages are treated as maximum ages, as we

cannot be certain which mainland population colonised the Canary Islands; maximum ages reflect this uncertainty in colonisation estimate. As crown estimates have also been proposed as an adequate proxy for island colonisation time for endemic island clades (García-Verdugo et al. 2019; Martín-Hernanz et al. 2023), we built another dataset with crown ages for Canary Island endemic species of the same monophyletic group, and stem ages for all other lineages (stem+crown dataset). In this dataset, endemic lineages were treated as precise ages of colonisation and non-endemic lineages as maximum ages (Supplementary Data S23). To avoid biasing against insular diversification, Macaronesian endemics were treated as Canarian endemics in the analyses (see Supplementary Data S1 for Macaronesian endemics), under the assumption that diversification happened within the Macaronesian meta-archipelago. Endemicity status in Figures 2 and 3 reflects this. Species that belong to our focal group but which are not sampled in the phylogeny (were not targeted for sequencing, or sequencing was not successful) are also included in the DAISIE analyses using the following approaches: a) if a species is not sampled in the phylogeny but is known to belong to a lineage that is sampled in the phylogeny, we add it as missing species to that clade, using the "missing species" option in DAISIE; b) if an entire lineage is missing from the phylogeny (can be a singleton lineage or a lineage composed of multiple species), we add this to DAISIE by specifying that the lineage could have colonised any time between the island age and the present; c) for lineages for which only the nearest mainland relatives were successfully sequenced, we used the stem age of the mainland relatives as a maximum age of colonisation. For three genera that were recovered as polyphyletic (*Cuscuta, Persicaria*, and *Viola*), the stem age of the oldest genus lineage was used as the maximum age.

DAISIE also requires an estimate of the size of the mainland pool of species that could have colonised the Canary Islands, to infer colonisation rates, and of the age of the insular system. The mainland pool size was estimated using native non-monocot angiosperm species data for the Mediterranean Basin from the World Checklist of Vascular Plants (WCVP) using the R package 'rWCVPdata' (Brown et al. 2023). A species list was built using native angiosperm species from all countries in the Mediterranean Basin. All duplicate instances per species and introduced species were removed to give the number of native angiosperm species in the Mediterranean Basin. For the non-monocot analyses, we removed all monocot species from the list of native angiosperm species in the Mediterranean Basin. Inaccuracies in the mainland pool size will affect inferred colonisation rates, but will not affect inferred speciation and extinction rates and subsequent model choice. The maximum age of the Canary Islands was fixed to 23 My following Campeny et al. (2024).

We fitted two sets of DAISIE models to each stem and stem+crown datasets, using the DAISIE_ML function in the DAISIE R package (Valente et al. 2015; Etienne et al. 2021): background and two-type models. The background models estimate the overall background parameters that apply to the entire focal assemblage; two-type models allow a specified group of species of a certain 'type' to be governed by different values of parameters compared to the background. For the background models, we fitted a set of four models that varied in whether K (carrying capacity) was estimated (diversity-dependent model) or fixed to infinity (diversity-independent model), and in whether anagenesis was estimated or fixed to zero. For the two-type models, we assessed model fit in three separate growth form analyses: herbaceousness at the time of colonisation (i.e., herbaceous and insular woody species on the Canaries) vs woodiness at the time of colonisation (ancestrally woody and derived woody species; Supplementary Data 12); insular woodiness vs all other growth form types (herbaceous, ancestrally woody, and derived woody) on the Canaries; and annual vs perennial species (i.e. annual herbaceous vs combination of perennial herbaceous, insular woody, ancestrally woody, and derived woody). For each of these growth form comparisons, we used the models 5-30 of Michielsen et al. (2023), which vary one or more parameters per model (i.e., cladogenesis, extinction, carrying capacity (K), colonisation, and/or anagenesis). For a full description of each of the 30 models, see Michielsen et al. (2023). Therefore, for each dataset (stem and stem+crown), we fitted a total of 30 models. We ran each of these 30 models 33 times, randomising input starting parameters. We then calculated the Bayesian Information Criterion (BIC) and Akaike Information Criterion (AIC) for each optimisation to select a preferred model. We favoured the BIC results as BIC has been shown to produce lower error rates when choosing among DAISIE models (Valente et al. 2020).

For the two-type analyses in DAISIE, we specified a proportion of species on the mainland that belong to each type. For the herbaceous vs woody analysis, the proportion of the type-2 species was calculated by combining three global datasets on growth form (Weigelt et al. 2020; Brown et al. 2023; Poppenwimer et al. 2023). The combined dataset was then filtered to only include the species in our mainland pool, and the growth form was retained for each dataset. As the WCVP does not record woodiness status, growth forms were reclassified to give binary woodiness/non-woodiness. To reconcile differences in the woodiness status of species between the three datasets, the average number of herbaceous and woody species was used to find the proportion for the herbaceous vs woody analysis (11,749 to 3,569, respectively). For the analyses of IW vs other growth forms, the mainland proportion was estimated by summing the number of herbaceous Mediterranean species in the most recent common

ancestral lineage of the Canarian IW lineages (including those which are not majority IW but have insular woody species; see Supplementary Data S11 for specific lineage used). This method estimated 1,386 potentially IW species, or 7.1% of the mainland pool size for non-monocots (see Supplementary Data S11). To assess how this estimate could affect the results, we reran the analysis by incrementally varying the proportion by half to double our obtained estimate and the model choice remained consistent (results not shown). For the annual vs perennial analysis, the mainland proportion was calculated using life form data from the WCVP (Brown et al. 2023). These steps were conducted for both the non-monocot only and whole angiosperm approaches. As model preference did not differ between the stem or stem+crown datasets, we report results only for the stem dataset, which is more comparable with previous studies using DAISIE, which mostly use the stem as the node for colonisation time proxy. Once the best model was identified for each growth form analysis, we used the DAISIE_sim function to simulate 1000 islands for the stem and stem+crown background and two type models. The results of the 1000 simulations for the herbaceous vs woody analysis in Supplementary Data S15, for the background model and the IW vs other growth forms are given in Supplementary Data S16, and for the annual vs perennial analysis in Supplementary Data S17.

To assess whether or not including monocots affected our results, we reran DAISIE including monocots. Because we did not sequence monocots, we added them to the data by making use of the 'missing data' modelling options available in DAISIE. To achieve this, we first used previously published phylogenetic information to infer the number of lineages and species per lineage of Canary Island monocots (see Section "Estimating angiosperm colonisations of the Canary Islands"). We then used this information to specify how many endemic and non-endemic species and lineages of monocots are present on the archipelago and to specify that these lineages could have colonised any time between the archipelago age and the present (Supplementary Data S21). Therefore, in the DAISIE analyses with monocots, we included diversity and endemicity status data for all monocot lineages, and treat branching and colonisation times as missing data (Valente et al. 2018).

To assess how uncertainty in the number of colonisations can affect our results, we reran DAISIE using our maximum number of inferred colonisations (Supplementary Data S1, S25), instead of our 'likely' number of colonisations. Again, our overall results remain unchanged. IW lineages have higher diversification rates (Supplementary Data S26), and woody species have lower extinction and colonisation rates than herbaceous species, but woody species were also found to have lower cladogenesis and higher anagenesis rates than herbaceous species (Supplementary Data S27). See Supplementary Data S26 and

S27 for the results of the 1000 simulations for the IW vs other growth forms and herbaceous vs woody analysis, respectively.

To assess the goodness of fit of the best fitting model for each analysis, we simulated 500 datasets using the best fitting parameters using the function 'DAISIE_sim' in the DAISIE R package (14, 131). We then compared six summary statistics for the background model, and eight for the two-type models, between these simulated datasets and the empirical datasets.

ACKNOWLEDGEMENTS

We thank Lizzie Roeble (Naturalis Biodiversity Center) and Lucas Miller (Technical University of Munich) for their contributions to target group identification, and the following institutes/people who helped us with the sampling: Universidad de La Laguna (TFC); Jardín Botánico Canario Viera y Clavijo-Unidad Asociada de I+D+i al CSIC (JBCVC-UACSIC); Banco de ADN de la Flora Canaria of the JBCVC-UACSIC; Flora Iberica (CGL2005-05471-C04-01); Javier Morente-López (Goethe-University Frankfurt); Universidad de Sevilla (SEV); Steven Janssens (Meise Botanic Gardens, BR); Georgian Academy of Sciences; Royal Botanic Gardens, Kew (K); Museum National d'Histoire Naturelle (P), P gives access to the collections in the framework of the RECOLNET national Research Infrastructure; Suzanne Cubey (Royal Botanical Gardens Edinburgh E); Ranee Prakash (Natural History Museum, London); Roderick Bouman (Hortus Botanicus, Leiden); Leopoldo Medina (Real Jardín Botánico de Madrid, MA); Institute of Biotechnology of the National Academy of Sciences of the Krygyz Republic; Andreas Berger (Naturhistorisches Museum Wien, W); Henry Väre (The Botanical Museum of the Finnish Museum of Natural History, H); Patrik Frödén and Arne Thell (Lund University Botanical Museum, LD); Filipe Covelo (Herbarium of the University of Coimbra, COI); Reto Nyffeler (United Herbaria, Z & ZT); and Roxali Bijmoer, Susana Arias Guerrero and Marnel Scherrenberg (Naturalis Biodiversity Center). We also thank our Naturalis colleagues Kasper Hendriks, Sander van Zon, Marina Ventayol Garcia and Elza Duijm for their assistance with sequencing and wet lab work, and Arjan Schieven, Kasper Hendriks, Lizzie Roeble, Nina Kerdiles, and Edwin van Haas for their assistance with bioinformatics. We thank Abelardo Aparicio, Rafael González Albaladejo, and Juan Viruel for sharing the *Helianthemum* sequencing data (BioProject PRJNA1219060). The collection and sequencing of *Helianthemum* were funded through the following projects: PID2020-116355GB-I00 and CGL2017-82465-P. Most of the Asteraceae family was sampled and sequenced via the support of the following Spanish Ministry of Science and Innovation (MICINN)'s projects:

ASTERALIEN - PID2019-110538GA-I00 and DecodAdapt - PID2023-147122NB-I00. We appreciate help from the Center for Information Technology of the University of Groningen to access the Hábrók high-performance computing cluster, from the Dutch Research Council for financial support through a grant awarded to F.L. and R.S.E (OCENW.KLEIN.498).

AUTHOR CONTRIBUTIONS

Conceptualisation: R.F.A.B., L.V., R.S.E., O.D.N., F.L.; Target Group Identification: R.F.A.B., L.V., R.S.E., O.D.N., F.L., J.M.F.-P., J.C.-C., S.M.-H., S.S., P.V., J.P., Y.A.F., L.S.J.-G.; Sampling: R.F.A.B., O.D.N., N.M.S.v.d.B., M.d.A.-A., C.G.-M., R.J.-M., J.C.-C., A.M., J.P., Y.A.F., L.S.J.-G., S.M.-H., S.S., P.V.; Sequencing: R.F.A.B., L.v.H., S.M.-H., J.P., Y.A.F.; Bioinformatics: R.F.A.B.,L.V., L.J.; Modelling: R.F.A.B., L.V., R.S.E., O.D.N.; Writing first draft: R.F.A.B.; L.V., R.S.E., O.D.N., F.L.; Supervision: F.L., L.V., R.S.E. All co-authors contributed to editing and reviewing the manuscript.

DATA AVAILABILITY

All scripts and publicly available data are available from a Zenodo repository (DOI: 10.5281/zenodo.16565207). Raw sequence data will be made available on NCBI upon acceptance at NCBI BioProject PRJNA1284034.

CODE AVAILABILITY

All scripts required to run the analyses herein are available from a Zenodo repository (DOI: 10.5281/zenodo.16565207) and archived in GitHub https://github.com/RFABrewer/Canarian_Angio_Biogeography/.

REFERENCES

Angiosperm Phylogeny Group, Chase, M. W., Christenhusz, M. J., Fay, M. F., Byng, J. W., Judd, W. S., & Stevens, P. F. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. *Botanical journal of the Linnean Society* **181**, 1-20 (2016).

Arjona, Y., Jay-Garcia, L. S., Reyes-Betancort, J. A., Morente-Lopez, J., Salas-Pascual, M., Naranjo-Cigala, A., & Patino, J. Taming Darwin's conundrum: On the role of phylogenetic and functional differentiation in plant naturalization on oceanic islands. Preprint at: https://doi.org/10.1101/2025.04.09.648018 (2025).

Arjona, Y., Nogales, M., Heleno, R., & Vargas, P. Long-distance dispersal syndromes matter: diaspore–trait effect on shaping plant distribution across the Canary Islands. *Ecography* **41**, 805-814 (2018).

Baker, W. J., Bailey, P., Barber, V., Barker, A., Bellot, S., Bishop, D., & Forest, F. A comprehensive phylogenomic platform for exploring the angiosperm tree of life. *Systematic Biology* **71**, 301-319, (2022).

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., & Pevzner, P. A. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology* **19**, 455-477 (2012).

Barajas Barbosa, M. P., Craven, D., Weigelt, P., Denelle, P., Otto, R., Díaz, S., & Kreft, H. Assembly of functional diversity in an oceanic island flora. *Nature* **619**, 545-550 (2023).

Barba-Montoya, J., Tao, Q., & Kumar, S. Assessing rapid relaxed-clock methods for phylogenomic dating. *Genome Biology and Evolution* **13**, evab251 (2021).

Beaulieu, J. M., & O'Meara, B. C. Extinction can be estimated from moderately sized molecular phylogenies. *Evolution* **69**, 1036-1043 (2015).

Beierkuhnlein, C., Nogales, M., Field, R., Vetaas, O. R., Walentowitz, A., Weiser, F., & Chiarucci, A. Volcanic ash deposition as a selection mechanism towards woodiness. *NPJ Biodiversity* **2**, 14 (2023).

Beierkuhnlein, C., Walentowitz, A., & Welss, W. FloCan—A revised checklist for the flora of the Canary Islands. *Diversity* **13**, 480 (2021).

BIOTA. *Biocan*, Banco del Inventario Natural de Canarias. Available from: https://www.biodiversidadcanarias.es/ [Accessed 20th July 2023].

Bolger, A. M., Lohse, M., & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114-2120 (2014).

Borregaard, M. K., Amorim, I. R., Borges, P. A., Cabral, J. S., Fernández-Palacios, J. M., Field, R., & Whittaker, R. J. Oceanic island biogeography through the lens of the general dynamic model: assessment and prospect. *Biological Reviews* **92**, 830-853 (2017).

Brandes, D., & Fritzsch, K. *Alien plants of Fuerteventura, Canary Islands Plantas extranjeras de Fuerteventura, Islas Canarias* (2002).

Brown, M. J., Walker, B. E., Black, N., Govaerts, R. H., Ondo, I., Turner, R., & Nic Lughadha, E. rWCVP: a companion R package for the World Checklist of Vascular Plants. *New Phytologist* **240**, 1355-1365 (2023).

Burns, K. C. *Evolution in Isolation: The Search for an Island Syndrome in Plants* (Cambridge University Press., 2019).

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. BLAST+: architecture and applications. *BMC Bioinformatics* **10**, 1-9 (2009).

Campeny, M., Menéndez, I., Quevedo, L., Yepes, J., Casillas, R., Ahijado, A., & Mangas, J. Rare Earth element distribution on the Fuerteventura Basal Complex (Canary Islands, Spain): a geochemical and mineralogical approach. *Solid Earth* **15**, 639-656 (2024).

Capella-Gutiérrez, S., Silla-Martínez, J. M., & Gabaldón, T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**, 1972-1973 (2009).

Carine, M. A., Santos-Guerra, A., Guma, I. R., & Reyes-Betancort, J. A. Endemism and evolution of the Macaronesian flora. *Beyond cladistics: The branching of a paradigm*, 101-124 (2010).

Carlquist, S. *Island Biology* (Columbia University Press, New York, 1974).

Carlquist, S. Monocot xylem revisited: new information, new paradigms. *The Botanical Review* **78**, 87-153 (2012).

Carruthers, T., Sun, M., Baker, W. J., Smith, S. A., De Vos, J. M., & Eiserhardt, W. L. The implications of incongruence between gene tree and species tree topologies for divergence time estimation. *Systematic Biology* **71**, 1124-1146 (2022).

Carvajal-Endara, S., Hendry, A. P., Emery, N. C., & Davies, T. J. Habitat filtering not dispersal limitation shapes oceanic island floras: species assembly of the Galápagos archipelago. *Ecology Letters* **20**, 495-504 (2017).

Caujapé-Castells, J. *Jesters, Red Queens, Boomerangs and Surfers: A Molecular Outlook on the Canarian Endemic Flora*. Pp. 284-324 In: Bramwell D, Caujapé-Castells J (eds) *The Biology of Island Floras* (Cambridge University Press, London, 2011).

Caujapé-Castells, J., García-Verdugo, C., Marrero-Rodríguez, Á., Fernández-Palacios, J. M., Crawford, D. J., & Mort, M. E. Island ontogenies, syngameons, and the origins and evolution of genetic diversity in the Canarian endemic flora. *Perspectives in Plant Ecology, Evolution and Systematics* **27**, 9-22 (2017).

Cerca, J., Cotoras, D. D., Bieker, V. C., De-Kayne, R., Vargas, P., Fernández-Mazuecos, M., & Martin, M. D. Evolutionary genomics of oceanic island radiations. *Trends in Ecology & Evolution* **38**, 631-642 (2023).

Costa, F. P., Schrago, C. G., & Mello, B. Assessing the relative performance of fast molecular dating methods for phylogenomic data. *BMC Genomics* **23**, 798 (2022).

Curto, M., Schachtler, C., Puppo, P., & Meimberg, H. Using a new RAD-sequencing approach to study the evolution of Micromeria in the Canary islands. *Molecular Phylogenetics and Evolution* **119**, 160-169 (2018).

Darwin, C. *On the Origin of Species by Means of Natural Selection* (John Murray, London, 1859).

Dehayem, O. N., Brewer, R. F.A., Valente, L., Lens, F., & Etienne, R. S. Impact of sampling strategy on inference of community assembly processes in phylogenetic island biogeography. *Methods in Ecology and Evolution* **16,** 1507-1520 (2025).

del Arco Aguilar, M. J., & Rodríguez Delgado, O. *Vegetation of the Canary Islands* (Springer International Publishing, 2018).

Denelle, P., Weigelt, P., & Kreft, H. GIFT—An R package to access the Global Inventory of Floras and Traits. *Methods in Ecology and Evolution* **14**, 2738-2748 (2023).

Díaz, S., Kattge, J., Cornelissen, J. H., Wright, I. J., Lavorel, S., Dray, S., & Zotz, G. The global spectrum of plant form and function: enhanced species-level trait dataset. *Scientific Data* **9**, 755 (2022).

Etienne R. S., Valente, L., Phillimore, A. B., Haegeman, B., Lambert, J. W., Neves, P., Xie, S., Bilderbeek, R. J. C., & Hildenbrandt, H. DAISIE: Dynamical Assembly of Islands by Speciation, Immigration and Extinction. R package version 4.0.2. https://cran.r-project.org/package=DAISIE. https://doi.org/10.5281/zenodo.4054058 (2021).

Etienne, R. S., Cabral, J. S., Hagen, O., Hartig, F., Hurlbert, A. H., Pellissier, L., & Storch, D. A minimal model for the latitudinal diversity gradient suggests a dominant role for ecological limits. *The American Naturalist* **194**, 122-133 (2019).

Etienne, R. S., Haegeman, B., Dugo-Cota, Á., Vilà, C., Gonzalez-Voyer, A., & Valente, L. The phylogenetic limits to diversity-dependent diversification. *Systematic Biology* **72**, 433-445 (2023).

Etienne, R. S., Haegeman, B., Stadler, T., Aze, T., Pearson, P. N., Purvis, A., & Phillimore, A. B. Diversity-dependence brings molecular phylogenies closer to agreement with the fossil record. *Proceedings of the Royal Society B: Biological Sciences* **279**, 1300-1309 (2012).

Fernández-Palacios, J. M., De Nascimento, L., Otto, R., Delgado, J. D., García-del-Rey, E., Arévalo, J. R., & Whittaker, R. J. 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 (2011).

Florencio, M., Patiño, J., Nogué, S., Traveset, A., Borges, P. A., Schaefer, H., & Santos, A. M. Macaronesia as a fruitful arena for ecology, evolution, and conservation biology. *Frontiers in Ecology and Evolution* **9**, 718169 (2021).

Frost, L. A., Bedoya, A. M., & Lagomarsino, L. P. Artifactual orthologs and the need for diligent data exploration in complex phylogenomic datasets: a museomic case study from the Andean flora. *Systematic Biology*, syad076 (2024).

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

García-Verdugo, C., Caujapé-Castells, J., & Sanmartín, I. Colonization time on island settings: lessons from the Hawaiian and Canary Island floras. *Botanical Journal of the Linnean Society* **191**, 155-163 (2019).

Gillespie, R. G. Island time and the interplay between ecology and evolution in species diversification. *Evolutionary Applications* **9**, 53-73 (2016).

Gillespie, R. G., Benjamin, S. P., Brewer, M. S., Rivera, M. A. J., & Roderick, G. K. Repeated diversification of ecomorphs in Hawaiian stick spiders. *Current Biology* **28**, 941-947 (2018).

Gitzendanner, M. A., Soltis, P. S., Wong, G. K. S., Ruhfel, B. R., & Soltis, D. E. Plastid phylogenomic analysis of green plants: a billion years of evolutionary history. *American Journal of Botany* **105**, 291-301 (2018).

Givnish, T. J. Ecology of plant speciation. *Taxon* **59**, 1326-1366 (2010).

Gotelli, N. J., Anderson, M. J., Arita, H. T., Chao, A., Colwell, R. K., Connolly, S. R., & Willig, M. R. Patterns and causes of species richness: a general simulation model for macroecology. *Ecology Letters* **12**, 873-886 (2009).

Hanz, D. M., Cutts, V., Barajas-Barbosa, M. P., Algar, A. C., Beierkuhnlein, C., Fernández-Palacios, J. M., ... & Irl, S. D. Climatic and biogeographical drivers of functional diversity in the flora of the Canary Islands. *Global Ecology and Biogeography* **31**, 1313-1331 (2022).

Hauffe, T., Delicado, D., Etienne, R. S., & Valente, L. Lake expansion elevates equilibrium diversity via increasing colonization. *Journal of Biogeography* **47**, 1849-1860 (2020).

Helmstetter, A. J., Zenil-Ferguson, R., Sauquet, H., Otto, S. P., Méndez, M., Vallejo-Marin, M., & Käfer, J. Trait-dependent diversification in angiosperms: Patterns, models and data. *Ecology Letters* **26**, 640-657 (2023).

Henao Diaz, L. F., Harmon, L. J., Sugawara, M. T., Miller, E. T., & Pennell, M. W. Macroevolutionary diversification rates show time dependency. *Proceedings of the National Academy of Sciences* **116**, 7403-7408 (2019).

Hendriks, K. P., Kiefer, C., Al-Shehbaz, I. A., Bailey, C. D., van Huysduynen, A. H., Nikolov, L. A., & Lens, F. Global Brassicaceae phylogeny based on filtering of 1,000-gene dataset. *Current Biology* **33**, 4052-4068 (2023).

Hendriks, K. P., Mandáková, T., Hay, N. M., Ly, E., Hooft van Huysduynen, A., Tamrakar, R., & Bailey, C. D. The best of both worlds: Combining lineage-specific and universal bait sets in target-enrichment hybridization reactions. *Applications in Plant Sciences* **9**, (2021).

Hernández-Hernández, T., & Wiens, J. J. Why are there so many flowering plants? A multiscale analysis of plant diversification. *The American Naturalist* **195**, 948-963 (2020).

Herrando-Moraira, S., Calleja, J. A., Carnicero, P., Fujikawa, K., Galbany-Casals, M., Garcia-Jacas, N., & Vilatersana, R. Exploring data processing strategies in NGS target enrichment to disentangle radiations in the tribe Cardueae (Compositae). *Molecular Phylogenetics and Evolution* **128**, 69-87 (2018).

Hoang, D. T., Chernomor, O., Von Haeseler, A., Minh, B. Q., & Vinh, L. S. UFBoot2: improving the ultrafast bootstrap

approximation. *Molecular Biology and Evolution* **35**, 518-522 (2018).

Hooft van Huysduynen, A., Janssens, S., Merckx, V., Vos, R., Valente, L., Zizka, A., & Lens, F. Temporal and palaeoclimatic context of the evolution of insular woodiness in the Canary Islands. *Ecology and Evolution* **11**, 12220-12231 (2021).

Illera, J. C., Rando, J. C., Melo, M., Valente, L., & Stervander, M. Avian Island Radiations Shed Light on the Dynamics of Adaptive and Nonadaptive Radiation. *Cold Spring Harbor Perspectives in Biology*, a041451 (2024).

Jackson, C., McLay, T., & Schmidt-Lebuhn, A. N. hybpiper-nf and paragone-nf: Containerization and additional options for target capture assembly and paralog resolution. *Applications in Plant Sciences* **11**, e11532 (2023).

Jiménez-Moreno, G., Burjachs, F., Expósito, I., Oms, O., Carrancho, Á., Villalaín, J. J., & van der Made, J. Late Pliocene vegetation and orbital-scale climate changes from the western Mediterranean area. *Global and Planetary Change* **108**, 15-28 (2013).

Jiménez-Moreno, G., Fauquette, S., & Suc, J. P. Miocene to Pliocene vegetation reconstruction and climate estimates in the Iberian Peninsula from pollen data. *Review of Palaeobotany and Palynology* **162**, 403-415 (2010).

Johnson, M. G., Gardner, E. M., Liu, Y., Medina, R., Goffinet, B., Shaw, A. J., & Wickett, N. J. HybPiper: Extracting coding sequence and introns for phylogenetics from high-throughput sequencing reads using target enrichment. *Applications in Plant Sciences* ***4***, 1600016 (2016).

Johnson, M. G., Pokorny, L., Dodsworth, S., Botigué, L. R., Cowan, R. S., Devault, A., & Wickett, N. J. A universal probe set for targeted sequencing of 353 nuclear genes from any flowering plant designed using k-medoids clustering. *Systematic Biology* **68**, 594-606 (2019).

Kattge, J., Bönisch, G., Díaz, S., Lavorel, S., Prentice, I. C., Leadley, P., & Cuntz, M. TRY plant trait database–enhanced coverage and open access. *Global Change Biology* **26**, 119-188 (2020).

König, C., Weigelt, P., Taylor, A., Stein, A., Dawson, W., Essl, F., & Kreft, H. Source pools and disharmony of the world's island floras. *Ecography* **44**, 44-55 (2021).

Koutroumpa, K., Warren, B. H., Theodoridis, S., Coiro, M., Romeiras, M. M., Jiménez, A., & Conti, E. Geo-climatic changes and apomixis as major drivers of diversification in the Mediterranean sea lavenders (Limonium Mill.). *Frontiers in Plant Science* **11**, 612258 (2021).

Kozlov, A. M., Darriba, D., Flouri, T., Morel, B., & Stamatakis, A. RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics* **35**, 4453-4455 (2019).

Lens, F., Davin, N., Smets, E., & del Arco, M. Insular woodiness on the Canary Islands: a remarkable case of convergent evolution. *International Journal of Plant Sciences* **174**, 992-1013 (2013).

Li, H., & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754-1760 (2009).

Losos, J. B., & Ricklefs, R. E. Adaptation and diversification on islands. *Nature* **457**, 830-836 (2009).

Louca, S., & Pennell, M. W. Why extinction estimates from extant phylogenies are so often zero. *Current Biology* **31**, 3168-3173 (2021).

MacArthur, R. H., & Wilson, E. O. An equilibrium theory of insular zoogeography. *Evolution*, 373-387 (1963).

MacArthur, R.H. & Wilson, E.O. *The Theory of Island Biogeography* (Princeton University Press, Princeton, 1967).

Magallón, S., Sánchez-Reyes, L. L., & Gómez-Acevedo, S. L. Thirty clues to the exceptional diversification of flowering plants. *Annals of Botany* **123**, 491-503 (2019).

Martín-Hernanz, S., Nogales, M., Valente, L., Fernández-Mazuecos, M., Pomeda-Gutiérrez, F., Cano, E., & Vargas, P. Time-calibrated phylogenies reveal Mediterranean and pre-Mediterranean origin of the thermophilous vegetation of the Canary Islands. *Annals of Botany* **131**, 667-684 (2023).

Martín-Hernanz, S., Viruel Sánchez, J., Aparicio Martínez, A., Gonzalez Albaladejo, R. Subset of Helianthemum sequences used in Brewer et al. dataset. Zenodo

https://doi.org/10.5281/zenodo.15343037 (2025).

McLay, T. G., Birch, J. L., Gunn, B. F., Ning, W., Tate, J. A., Nauheimer, L., & Jackson, C. J. New targets acquired: Improving locus recovery from the Angiosperms353 probe set. *Applications in Plant Sciences* **9,** (2021).

McPeek, M. A., & Brown, J. M. Clade age and not diversification rate explains species richness among animal taxa. *The American Naturalist* **169**, 97-106 (2007).

Michielsen, N. M., Goodman, S. M., Soarimalala, V., van der Geer, A. A., Dávalos, L. M., Saville, G. I., & Valente, L. The macroevolutionary impact of recent and imminent mammal extinctions on Madagascar. *Nature Communications* **14** (2023).

Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., Von Haeseler, A., & Lanfear, R. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Molecular Biology and Evolution* **37**, 1530-1534 (2020).

Morales-Briones, D. F., Gehrke, B., Huang, C. H., Liston, A., Ma, H., Marx, H. E., & Yang, Y. Analysis of paralogs in target enrichment data pinpoints multiple ancient polyploidy events in Alchemilla sl (Rosaceae). *Systematic Biology* **71**, 190-207 (2022).

Murphy, B., Forest, F., Barraclough, T., Rosindell, J., Bellot, S., Cowan, R., & Cheek, M. A phylogenomic analysis of Nepenthes (Nepenthaceae). *Molecular Phylogenetics and Evolution* **144**, 106668 (2020).

Palazzesi, L., Pellicer, J., Barreda, V. D., Loeuille, B., Mandel, J. R., Pokorny, L., ... & Hidalgo, O. Asteraceae as a model system for evolutionary studies: from fossils to genomes. *Botanical Journal of the Linnean Society* **200**, 143-164 (2022).

Patiño, J., Whittaker, R. J., Borges, P. A., Fernández-Palacios, J. M., Ah-Peng, C., Araújo, M. B., & Emerson, B. C. A roadmap for island biology: 50 fundamental questions after 50 years of The Theory of Island Biogeography. *Journal of Biogeography* **44**, 963-983 (2017).

Poppenwimer, T., Mayrose, I., & DeMalach, N. Revising the global biogeography of annual and perennial plants. *Nature* **624**, 109-114 (2023).

Price, J. P., Otto, R., Menezes de Sequeira, M., Kueffer, C., Schaefer, H., Caujapé-Castells, J., & Fernández-Palacios, J. M. Colonization and diversification shape species–area relationships in three Macaronesian archipelagos. *Journal of Biogeography* **45**, 2027-2039 (2018).

Rabosky, D. L. Challenges in the estimation of extinction from molecular phylogenies: a response to Beaulieu and O'Meara. *Evolution* **70**, 218-228 (2016).

Rabosky, D. L. Extinction rates should not be estimated from molecular phylogenies. *Evolution* **64**, 1816-1824 (2010).

Rabosky, D. L., & Glor, R. E. Equilibrium speciation dynamics in a model adaptive radiation of island lizards. *Proceedings of the National Academy of Sciences* **107**, 22178-22183 (2010).

Rabosky, D. L., & Hurlbert, A. H. Species richness at continental scales is dominated by ecological limits. *The American Naturalist* **185**, 572-583 (2015).

Ramírez-Barahona, S., Sauquet, H., & Magallón, S. The delayed and geographically heterogeneous diversification of flowering plant families. *Nature Ecology & Evolution* **4**, 1232-1238 (2020).

Rangel, T. F., Edwards, N. R., Holden, P. B., Diniz-Filho, J. A. F., Gosling, W. D., Coelho, M. T. P., & Colwell, R. K. Modelling the ecology and evolution of biodiversity: Biogeographical cradles, museums, and graves. *Science* **361**, eaar5452 (2018).

Ranwez, V., Douzery, E. J., Cambon, C., Chantret, N., & Delsuc, F. MACSE v2: toolkit for the alignment of coding sequences accounting for frameshifts and stop codons. *Molecular Biology and Evolution* **35**, 2582-2584 (2018).

Ricklefs, R. E. A comprehensive framework for global patterns in biodiversity. *Ecology Letters* **7**, 1-15 (2004).

Roeble, L., van Benthem, K. J., Weigelt, P., Kreft, H., Knope, M. L., Mandel, J. R., & Valente, L. Island biogeography of the megadiverse plant family Asteraceae. *Nature Communications* **15** (2024).

Ruhfel, B. R., Gitzendanner, M. A., Soltis, P. S., Soltis, D. E., & Burleigh, J. G. From algae

to angiosperms–inferring the phylogeny of green plants (Viridiplantae) from 360 plastid genomes. *BMC Evolutionary Biology* **14**, 1-27 (2014).

Sauerbier, H., Calixto, F. C., & Muer, T. *Flora vascular de Canarias* (Turquesa, 2023).

Schrader, J., Weigelt, P., Cai, L., Westoby, M., Fernández-Palacios, J. M., Cabezas, F. J., & Kreft, H. Islands are key for protecting the world's plant endemism. *Nature*, 1-7 (2024a).

Schrader, J., Wright, I. J., Kreft, H., Dixon, K., Gallagher, R. V., Andrew, S. C., & Westoby, M. Trait filtering in island floras: A conceptual framework. *Journal of Biogeography* **51**, 1596-1606 (2024b).

Scornavacca, C., Belkhir, K., Lopez, J., Dernat, R., Delsuc, F., Douzery, E. J., & Ranwez, V. OrthoMaM v10: scaling-up orthologous coding sequence and exon alignments with more than one hundred mammalian genomes. *Molecular Biology and Evolution* **36**, 861-862 (2019).

Slater, G. S. C., & Birney, E. Automated generation of heuristics for biological sequence comparison. *BMC Bioinformatics* **6**, 1-11 (2005).

Smith, S. A., & Beaulieu, J. M. Ad fontes: divergence-time estimation and the age of angiosperms. *New Phytologist* **244**, 760-766 (2024).

Smith, S. A., & Donoghue, M. J. Rates of molecular evolution are linked to life history in flowering plants. *Science* **322**, 86-89 (2008).

Smith, S. A., Brown, J. W., & Walker, J. F. So many genes, so little time: a practical approach to divergence-time estimation in the genomic era. *PloS One* **13**, e0197433 (2018).

Thomson, F. J., Moles, A. T., Auld, T. D., & Kingsford, R. T. Seed dispersal distance is more strongly correlated with plant height than with seed mass. *Journal of Ecology* **99**, 1299-1307 (2011).

Tietje, M., Antonelli, A., Baker, W. J., Govaerts, R., Smith, S. A., & Eiserhardt, W. L. Global variation in diversification rate and species richness are unlinked in plants. *Proceedings of the National Academy of Sciences* **119**, e2120662119 (2022).

Triantis, K. A., Economo, E. P., Guilhaumon, F., & Ricklefs, R. E. Diversity regulation at macro-scales: species richness on oceanic archipelagos. *Global Ecology and Biogeography* **24**, 594-605 (2015).

Upham, N. S., Esselstyn, J. A., & Jetz, W. Inferring the mammal tree: species-level sets of phylogenies for questions in ecology, evolution, and conservation. *PLoS Biology* **17**, e3000494 (2019).

Valente, L. M., & Vargas, P. Contrasting evolutionary hypotheses between two mediterranean-climate floristic hotspots: the Cape of southern Africa and the Mediterranean Basin. *Journal of Biogeography* **40**, 2032-2046 (2013).

Valente, L. M., Phillimore, A. B., & Etienne, R. S. Equilibrium and non-equilibrium dynamics simultaneously operate in the Galápagos islands. *Ecology Letters* **18**, 844-852 (2015).

Valente, L. M., Phillimore, A. B., & Etienne, R. S. Using molecular phylogenies in island biogeography: It's about time. *Ecography* **41**, 1-3 (2018).

Valente, L., Etienne, R. S., & Dávalos, L. M. Recent extinctions disturb path to equilibrium diversity in Caribbean bats. *Nature Ecology & Evolution* **1**, 0026 (2017b).

Valente, L., Illera, J. C., Havenstein, K., Pallien, T., Etienne, R. S., & Tiedemann, R. Equilibrium bird species diversity in Atlantic islands. *Current Biology* **27**, 1660-1666 (2017a).

Valente, L., Phillimore, A. B., Melo, M., Warren, B. H., Clegg, S. M., Havenstein, K., & Etienne, R. S. A simple dynamic model explains the diversity of island birds worldwide. *Nature* **579**, 92-96 (2020).

Valtuena, F. J., López, J., Álvarez, J., Rodríguez-Riaño, T., & Ortega-Olivencia, A. Scrophularia arguta, a widespread annual plant in the Canary Islands: a single recent colonization event or a more complex phylogeographic pattern? *Ecology and Evolution* **6**, 4258-4273 (2016).

Vargas, P. *The Mediterranean Floristic Region: High Diversity of Plants and Vegetation Types*. In: Goldstein, M.I., DellaSala, D.A. (Eds.), *Encyclopedia of the World's Biomes* (Elsevier, 2020).

Vasconcelos, T. A trait-based approach to determining principles of plant biogeog-

raphy. *American Journal of Botany* **110**, e16127 (2023).

Veron, S., Kondratyeva, A., Robuchon, M., Grandcolas, P., Govaerts, R., Haevermans, T., & Mouchet, M. High evolutionary and functional distinctiveness of endemic monocots in world islands. *Biodiversity and Conservation* **30**, 3697-3715 (2021).

Von Humboldt, A., & Bonpland, A. *Essay on the Geography of Plants* (University of Chicago Press., 1807)

Wallace, A. R. XVIII.—On the law which has regulated the introduction of new species. *Annals and Magazine of Natural History* **16**, 184-196 (1855).

Warren, B. H., Simberloff, D., Ricklefs, R. E., Aguilée, R., Condamine, F. L., Gravel, D., & Thébaud, C. Islands as model systems in ecology and evolution: Prospects fifty years after MacArthur-Wilson. *Ecology Letters* **18**, 200-217 (2015).

Weigelt, P., Kissling, D.W., Kisel, Y., Fritz, S. A., Karger, D. N., Kessler, M., & Kreft, H. Global patterns and drivers of phylogenetic structure in island floras. *Scientific Reports* **5**, 12213 (2015).

Weigelt, P., König, C., & Kreft, H. GIFT–A Global Inventory of Floras and Traits for macroecology and biogeography. *Journal of Biogeography* **47**, 16-43 (2020).

White, O. W., Reyes-Betancort, J. A., Chapman, M. A., & Carine, M. A. Geographical isolation, habitat shifts and hybridisation in the diversification of the Macaronesian endemic genus Argyranthemum (Asteraceae). *New Phytologist* **228**, 1953-1971 (2020).

Whittaker, R. J., Triantis, K. A., & Ladle, R. J. A general dynamic theory of oceanic island biogeography. *Journal of Biogeography* **35**, 977-994 (2008).

Wiens, J. J. What explains patterns of biodiversity across the Tree of Life? New research is revealing the causes of the dramatic variation in species numbers across branches of the Tree of Life. *BioEssays* **39**, 1600128 (2017).

Yang, Y., & Smith, S. A. Orthology inference in nonmodel organisms using transcriptomes and low-coverage genomes: improving accuracy and matrix occupancy for phylogenomics. *Molecular Biology and Evolution* **31**, 3081-3092 (2014).

Yu, D., & Wiens, J. J. The causes of species richness patterns among clades. *Proceedings of the Royal Society B* **291**, 20232436 (2024).

Zhang, C., & Mirarab, S. ASTRAL-Pro 2: ultrafast species tree reconstruction from multi-copy gene family trees. *Bioinformatics* **38**, 4949-4950 (2022b).

Zhang, C., & Mirarab, S. Weighting by gene tree uncertainty improves accuracy of quartet-based species trees. *Molecular Biology and Evolution* **39**, msac215 (2022a).

Zizka, A., Onstein, R. E., Rozzi, R., Weigelt, P., Kreft, H., Steinbauer, M. J., ... & Lens, F. The evolution of insular woodiness. *Proceedings of the National Academy of Sciences* **119**, e2208629119 (2022).

Zuntini, A. R., Carruthers, T., Maurin, O., Bailey, P. C., Leempoel, K., Brewer, G. E., & Knapp, S. Phylogenomics and the rise of the angiosperms. *Nature* **629**, 1-8 (2024).

For the supplementary material associated with this chapter, please see the Zenodo repository accompanying this thesis (DOI 10.5281/zenodo.22093870).





Chapter 3

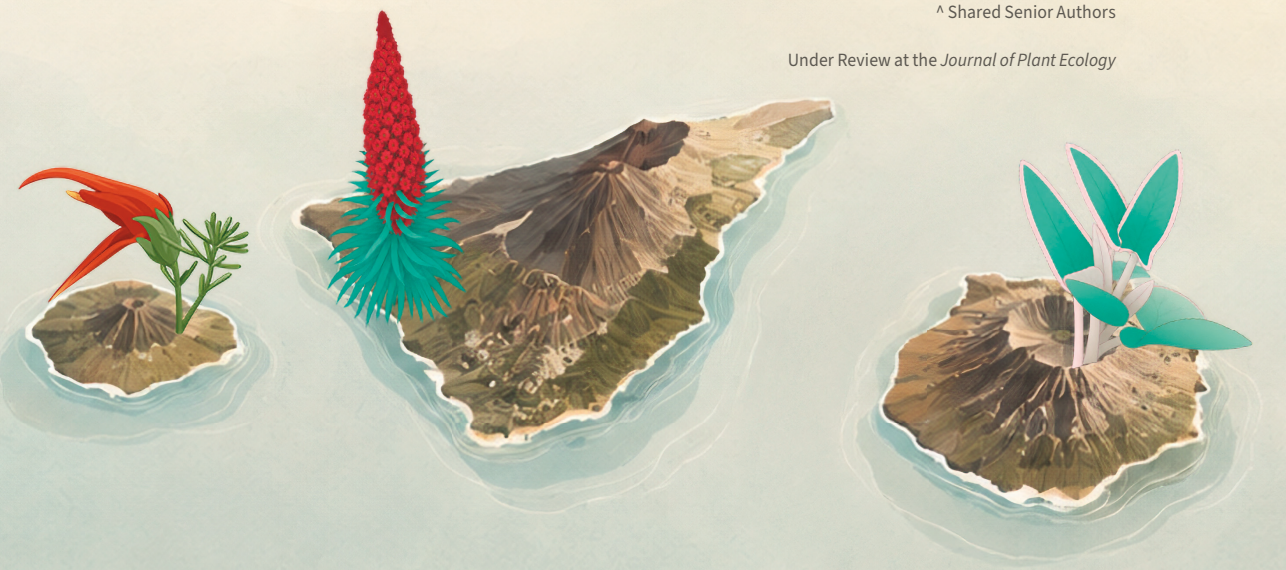
Climatic seasonality and topographic complexity shape Canary Islands plant growth form distributions

Lucas S. Jansen[1,*], Ryan F.A. Brewer[1,2,*,], Juli Caujapé-Castells[3], Luis Valente[1,2,^], Frederic Lens[1,4,^]**

[1] Naturalis Biodiversity Center, Darwinweg 2, 2333 CR Leiden, The Netherlands;
[2] Groningen Institute for Evolutionary Life Sciences, University of Groningen, Box 11103, 9700 CC Groningen, The Netherlands;
[3] Department of Molecular Biodiversity & DNA Bank of the Canarian Flora, Jardín Botánico Canario 'Viera y Clavijo' –Unidad Asociada CSIC, Cabildo de Gran Canaria, Las Palmas, Gran Canaria, Spain;
[4] Institute of Biology Leiden, Leiden University, Sylviusweg 72, Leiden 2333 BE, The Netherlands

* Shared First Authors
** Corresponding author: Ryan F. A. Brewer (R.F.Brewer@outlook.com)
^ Shared Senior Authors

Under Review at the *Journal of Plant Ecology*



ABSTRACT

Plant growth forms reflect the environmental conditions under which they evolved and persist. Yet, previous studies linking environmental conditions with plant growth form have overlooked that different growth form types arose through distinct evolutionary pathways. For instance, some woody angiosperm lineages evolved woodiness from a herbaceous ancestor, whereas others have retained their woodiness throughout evolutionary history. The Canary Islands are a global hotspot of insular woodiness (woodiness evolved on islands from herbaceous colonisers), raising the question of whether insular woody species occur in different environmental conditions than other growth forms. Here, we use a novel pipeline to extract, filter, and analyse publicly available spatial data to identify the environmental correlates of plant growth forms in the Canary Islands non-monocot angiosperm flora, accounting for their distinct evolutionary histories and differences in life span. To assess whether and how five growth forms (annual herbaceous, perennial herbaceous, ancestral woodiness, derived woodiness, and insular woodiness) differ in environmental niche space, we applied phylogenetic principal component analysis to a dataset of ~600 native non-monocot Canary Island angiosperm species. Unexpectedly, we find that herbaceous and woody growth forms highly overlap in niche space. Subtle differences are explained mostly by climatic seasonality (temperature and precipitation) and topographic complexity. Using phylogenetic ANOVAs, we identified significant pairwise differences between growth forms across five environmental variables, with differentiation between combined woody and herbaceous niches, and between different evolutionary histories across woody growth forms. Our phylogenetically-inferred macroecological approach can be applied to any categorical trait of interest in plants or animals.

Keywords: Canary Islands; Environmental niche; growth form; phylogenetic ANOVA; phylogenetic principal component analysis

INTRODUCTION

Since transitioning from ocean to land ~500 million years ago (Morris et al. 2018), terrestrial plant life has evolved into a diverse array of growth forms: from the structurally simple extinct early land plants to the complex forms of extant herbaceous and woody seed plants (Stewart and Rothwell 1993). This myriad of growth forms reflects the wide range of environmental conditions and biological interactions that have given rise to the extant global flora (Rowe & Speck 2005). Yet, despite centuries of study (von Humboldt 1807), much remains unknown about the relationship between environmental factors and different forms of plant life.

Many different ways to classify growth forms exist, from the classification system of Raunkiaer (e.g., chamaephyte, phanerophyte, etc.; Raunkiaer 1934; Irl et al. 2020), to a simple binary trait state describing the aboveground stem (herbaceous vs woody; Poppenwimer et al. 2023), a more complex multistate that takes into account life span (annuality vs perenniality) and evolutionary history (Hooft van Huysduynen et al. 2021), or based on descriptions of the entire plant (e.g., trees, shrubs, epiphytes, terrestrial herbs; Taylor et al. 2023). Here, we adopt the multistate growth form classification to look into the relationship with environmental factors. This relationship is clearly visible. For example, at the global scale, a recent study found that the extant distributions of annual herbaceous species — i.e., non-woody species that complete their life cycle within one year — are driven by prolonged high temperatures and seasonal drought (Poppenwimer et al. 2023), with these species surviving the harshest periods of the year as seeds. In contrast, perennial species (those with longer-lived belowground and/or aboveground organs that are either herbaceous or woody) are more frequently encountered in tropical conditions and at the highest latitudes, creating a non-linear global latitudinal relationship in the distribution of annuals and perennials (Friedman 2020; Poppenwimer et al. 2023; Taylor et al. 2023).

However, an often overlooked factor in macroecological studies of plant growth form is that extant growth forms have evolved through different evolutionary trajectories; by grouping growth forms with different evolutionary origins, based on extant phenotype, their true environmental correlates may be obscured. For example, because angiosperms originated from woody gymnosperm-like ancestors, it is generally accepted that the most recent common ancestor of angiosperms was woody, giving rise to today's "ancestrally woody" species that retained their woody growth form throughout evolutionary history (Gerrienne et al. 2011; Doyle 2012). Many woody angiosperm lineages later independently lost woodiness, evolving into (annual or perennial) herbaceous

plant taxa, accounting for nearly 50% of extant angiosperm species (Klimeš et al. 2022). Interestingly, multiple herbaceous lineages have re-evolved woodiness, an evolutionary reversal known as phylogenetically derived woodiness (Zizka et al. 2022). However, despite the massive number of transitions from herbaceousness towards derived woodiness in angiosperms on islands and continents (e.g., *Lobelia* (Campanulaceae), *Veronica* (Plantaginaceae; Zizka et al. 2022, 2025), the majority of woody angiosperm species we see today are ancestrally woody, i.e. have retained woodiness from woody ancestors (Hooft van Huysduynen et al. 2021).

Within derived woody growth types, a further distinction based on geographical context can be made. "Insular woodiness" refers to when phylogenetic inference shows that woodiness evolution occurred on an island after being colonised by a herbaceous continental species (Carlquist 1974; Zizka et al. 2022). Insular woodiness differs from (continental-)derived woodiness, which are island species which shifted from herbaceousness to woodiness on a nearby continent before colonising the island as a derived woody island species. This distinction recognises insular woodiness as a typical island phenomenon, considered as the most conspicuous part of the classic island syndrome set of plant traits (Darwin 1859; Carlquist 1974; Lens et al. 2013a; Zizka et al. 2022). Interestingly, many iconic island lineages are insular woody, such as the Hawaiian silverswords (*Argyroxiphium, Dubautia, Wilkesia*, Asteraceae; Baldwin & Sanderson 1998) and Canary Island viper buglosses (*Echium,* Boraginaceae; Böhle et al. 1996).

Several hypotheses have been proposed to explain the evolution of insular woodiness, such as competition for capturing sunlight in dense herbaceous populations growing in open environments after volcanic islands rise above sea level (Darwin 1859), and longer life spans either due to a lack of large native herbivores on islands (Carlquist 1974), due to longer flowering times as a result of pollinator scarcity on isolated islands (Wallace 1880), or due to a favourable aseasonal insular climate for (sub)tropical islands (Carlquist 1974; Zizka et al. 2025). More recently, a theory proposed by Lens et al. (2013b), stated that drought could have been an important global driver of woodiness evolution. Experimental evidence for this theory, provided by, among others, Dória et al. (2018), found that stems of insular woody daisies on the Canary Islands (*Argyranthemum*, Asteraceae) were better suited to avoid drought-induced gas bubbles in their water-conducting vessels compared to their herbaceous relatives. This drought hypothesis was further supported by a global macroecological analysis showing repeated transitions toward woodiness on islands (Zizka et al. 2022), and even more strongly on continents, where the drought signal is particularly pronounced (Zizka et al. 2025).

Research on the macroecology of growth form in island plants has hitherto overlooked the distinct evolutionary histories associated with these growth forms (such as insular woodiness, ancestral woodiness, and derived woodiness). Consequently, if we aim to disentangle the relationship between environment and growth form on islands, we need comprehensive analyses addressing fine-scale environmental correlates and considering the different evolutionary origins of growth forms (Sauquet & Magallón 2018). Here we perform such an analysis, focusing on the Canary Islands. The Canary Islands archipelago has one of the most diverse and best-studied insular floras globally and is a hotspot for insular woodiness, in both numbers of species and shifts towards insular woodiness (Beierkuhnlein et al. 2021; Zizka et al. 2022; Barajas Barbosa et al. 2023). To investigate fine-scale processes underlying plant growth forms, we categorised the non-monocot angiosperm flora of the Canary Islands into two herbaceous (annual and perennial herbaceous) and three woody (ancestral, derived, and insular woodiness) growth forms. We do not include the exclusively herbaceous monocots as they never develop woodiness due to the lack of a wood-producing vascular cambium, even if the environmental conditions would favour woodiness (Carlquist 2012; Zizka et al. 2022).

We test the following hypotheses: (1) woody and herbaceous species on the Canaries occur in localities with different environmental characteristics, detectable through differences in niche space. Within the two herbaceous and three woody growth forms, we expect additional differences in niche space due to differences in life span and evolutionary history, respectively. (2) As insular woody lineages are, by definition, the only ones that evolved woodiness *in situ*, we expect the insular woody niche to deviate most from that occupied by herbaceous growth forms, reflecting a possible adaptive benefit of this shift in growth form on the islands. (3) In line with previous studies addressing the drivers of insular woodiness, we expect that variables related to aridity and aseasonality impact the niche of insular woody species (Carlquist 1974; Lens et al. 2013a; Hooft van Huysduynen et al. 2021; Zizka et al. 2022).

To assess these hypotheses, we follow four steps. First, we develop a novel bioinformatics pipeline to extract environmental variables based on open-source plant occurrence data and use it to compile a dataset of environmental variables for the non-monocot angiosperm Canarian flora. Secondly, we use a phylogenetic principal component analysis (Revell 2009) to disentangle the variables that delineate environmental niche space for the angiosperm flora of the Canaries, identifying the environmental correlates for each growth form whilst accounting for phylogenetic ancestry. Third, we apply phylogenetic ANOVAs to test for pairwise differences in the association of each variable with the distribution of each growth form. Finally, we highlight the impact of the

variable that was most significant in the phylogenetic ANOVAs: rugosity, i.e. topographic complexity, measured using a terrain ruggedness index.

METHODS

Locality Data Extraction

Study Region & Target Species

The Canary Islands are a volcanic archipelago 96 km to the west of Point Stafford (Western Sahara) and consist of seven main islands (El Hierro, La Palma, La Gomera, Tenerife, Gran Canaria, Fuerteventura and Lanzarote) and four small islands (<30 km2) off the coast of Lanzarote (La Graciosa, Alegranza, Lobos, and Montaña Clara; Figure 1). The Canarian archipelago has one of the most species-rich oceanic island floras globally, as well as the highest number of insular woody species globally (Triantis et al. 2015; Zizka et al. 2022).

The Canary Islands host ~1,375 angiosperm species, of which approximately a third are endemic to the archipelago (Canary Islands Biodiversity Database 2024). This study focuses on angiosperms given their well-resolved taxonomic backbone (Sauquet & Magallón 2018), their extensive representation in earlier evolutionary research (Hernández-Hernández & Wiens 2020; Florencio et al. 2021), and the broad range of herbaceous (annual, perennial) and woody growth forms (insular woody, derived woody, ancestrally woody).

We focus exclusively on non-monocot angiosperms as monocots never produce a wood cylinder due to the lack of a wood-forming vascular cambium. In a few monocot lineages, there is secondary growth through the formation of a monocot cambium (or 'master cambium') that is not equivalent to the vascular cambium, as different tissues are produced (Lens et al. 2013b; Carlquist 2012). Therefore, as we are interested in testing the theories on the development of insular woodiness, monocots are removed due to their inability to produce wood, and we retain only the angiosperm lineages with the capacity to develop woodiness. In total, we targeted all 1,160 non-monocot angiosperm species native to the Canary Islands (of which several were removed in the final analyses after filtering (see 'GBIF Extraction & Locality Data Filtering' below) and sampling regimes (see 'Sampling Regimes' below); see Table 2).

GBIF Extraction & Locality Data Filtering

We developed an approach to extract, filter, and analyse publicly available occurrence and spatial data, whilst accounting for phylogeny, and to subsequently visualise environmental niche space. Details on our novel pipeline are covered in section 2.2 below. To obtain distribution data on the native

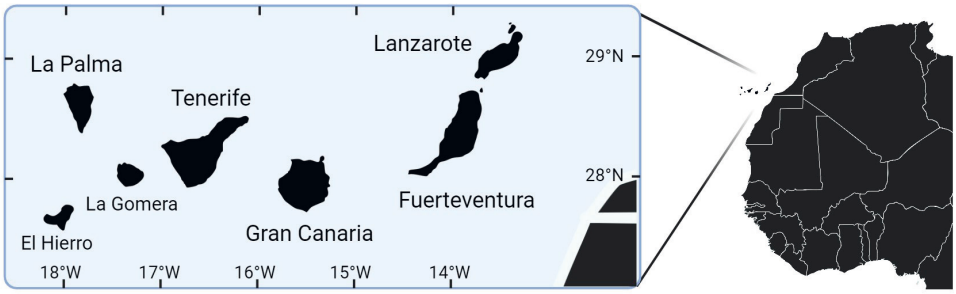


Figure 1. Map of the Canary Islands archipelago. The names of the seven main islands are included, as well as latitude and longitude (degrees) ticks.

non-monocot angiosperm species of the Canary Islands, we extracted occurrence data from the Global Biodiversity Information Facility (GBIF) using the 'rgbif' R package (Chamberlain et al. 2024; GBIF 2025), following an updated version of Beierkuhnlein et al. (2021). For 186 plant species, no GBIF occurrences were found, and thus these were removed from the subsequent analyses, resulting in a total of 974 plant species for which at least one GBIF occurrence was found on the Canary Islands (Table S1). For a list of the changes made to Beierkunhlein et al. (2021), see Table S2.

Occurrence data was filtered in four steps. First, we filtered locality data using the 'clean_coordinates' function of the 'CoordinateCleaner' package (Zizka et al. 2019). After evaluating the results of these initial cleaning criteria, further filtering steps were implemented to remove occurrences that still displayed erroneous patterns. We implemented a custom function to check whether the occurrence coordinates were assigned to islands where the respective plant species are known to natively occur (following Beierkuhnlein et al. 2021). Coordinates on smaller islets and islands were categorised under the name of the island to which they were closest in proximity (e.g. coordinates on La Graciosa and Alegranza were extracted as occurring on Lanzarote).

We also wrote a custom function to remove reprojected occurrences (i.e., from raster to vector data). Due to the reprojection, these occurrences no longer corresponded to exact localities where the species occurs, but rather to the central point in the raster grid prior to reprojection. As such, including these occurrences would bias our dataset due to the inaccuracies stemming from reprojection. These occurrences were removed by filtering out occurrences for which the code 'cdrep' was assigned in the extracted GBIF data. After this, we performed a final filtering step based on the reported coordinate uncertainty; occurrences with a coordinate uncertainty above a certain threshold were filtered out. To do this, three different sampling regimes were used.

Sampling Regimes

To find a balance between high species representation in terms of number of occurrence records and high coordinate certainty, we compared the results for three different uncertainty threshold resolutions: 1000, 500 and 100 meters, and minimum number of occurrences per species (after filtering): at least 3, 13 and 25 occurrences per species. These occurrence thresholds were based on van Proosdij et al. (2016), who suggested several thresholds to accurately model species distributions based on different geographic scenarios; fewer occurrences are needed when modelling species with small distributions (i.e., narrow endemics) than more widely distributed species.

These thresholds were combined into three sampling regimes: lenient, balanced, and strict. The lenient sampling regime has a minimum of three occurrences per species with a maximum uncertainty limit of 1000 meters; the balanced sampling regime has at least 13 occurrences with a maximum uncertainty of 500 meters; and the strict sampling strategy has at least 25 occurrences and a maximum uncertainty of 100 meters.

Environmental Dataset Compilation

Variable Extraction

We extracted 16 environmental variables for the (filtered) occurrences for each species from various environmental (raster) layers from multiple sources. An overview of these variables and their sources is presented in Table 1. All environmental layers had a resolution of 100 meters (~3-4 arc-seconds) and were assigned the same coordinate reference system. See the 'Code Availability' section for the source files.

We selected nine environmental variables from the CanaryClim v1.0 dataset (Table 2; Patino et al. 2023), comprising both topographic and climatic variables. To investigate the potential drivers of insular woodiness on the Canary Islands, we included variables associated with some of the published hypotheses: diurnal and seasonal temperature variability, temperature of the coldest quarter, and precipitation seasonality (calculated as the coefficient of variation of the monthly precipitation estimates) to assess the aseasonal favourable climate hypothesis (Carlquist 1974); and the air temperature of the driest quarter, annual precipitation amount, and the precipitation in the driest quarter to assess the drought hypothesis (Zizka et al. 2022). Additionally, data on evapotranspiration and interception (hereafter referred to as evapotranspiration) were included (FAO, 2020). Using the evapotranspiration and precipitation data, an (adjusted Thornthwaite) aridity index was calculated (Herschy et al. 1998), quantifying the ratio of water entering (through precipitation) to water leaving

(through evapotranspiration) the environment via plant species, as described in equation (1):

$$AI_s = \frac{P_s}{ET_s}$$

where, for species *s*, *AIs* is the aridity index, *Ps* is the median annual precipitation (mm/year), and *ETs* is the median annual actual evapotranspiration and interception (mm/year).

Furthermore, we used the Digital Elevation Model (DEM) of the CanaryClim dataset to extract the elevation for each of the filtered occurrences. Elevation can serve as a proxy for multiple different parameters related to plant growth, such as the amount of solar radiation, air temperature, and precipitation (Körner 2007). Also, topographic heterogeneity has been recognised as a key variable linked to plant species richness (Liu et al. 2019; Nürk et al. 2020; Chang et al. 2023). As such, we included a Terrain Ruggedness Index (TRI), calculated following the methodology of Riley et al. (1999). TRI quantifies topographic heterogeneity by considering the difference in elevation between a cell in a grid and the neighbouring cells. The TRI raster was calculated using the 'tri' function in the 'spatialEco' package (Evans & Murphy 2023), using a value of 3 as the scale window parameter and the CanaryClim DEM as input. Finally, we used the cloud cover data from the Climatic Atlas of the Canaries (Luque Sollheim et al. 2024). Cloud cover is driven largely by the north-east trade winds on islands with sufficient elevation, and generates climatic zones particularly associated with the laurisilva forest (del Arco-Aguilar & Rodríguez Delgado 2018).

Variable Aggregation

After extracting the 16 environmental variables for each species, we aggregated the occurrence-level data to the species level. To summarise the information from occurrences, we used the median value for most environmental variables (Table 1), because of its robustness against outlying values (Müller 2000). For elevation, besides the median, we also used the median absolute deviation (MAD) to investigate the relationship between the elevational range at which a species occurs and its growth form.

Growth Form Determination

For each species, we scored the growth form in one of five categories: annual herbaceous (AH), perennial herbaceous (PH), insular woodiness (IW), derived woodiness (DW), and ancestral woodiness (AW), following Hooft van Huys-

duynen et al. (2021). We determined the growth form of the species through several methods. First, for IW lineages, we used IW species lists from Lens et al. (2013a) and Zizka et al. (2022). For DW and AW lineages, we used Hooft van Huysduynen et al. (2021), Zizka et al. (2025), and an unpublished dataset by F. Lens. If no information on growth form was found for a species in these studies, the protologue of the species was used, primarily accessed through the International Plant Names Index (IPNI 2024). In case the protologue could not be found or did not resolve the growth form, a literature search was performed using Google Scholar and the Web of Science (WoS) for the species name and a reference to the growth form (e.g. 'woody', 'woodiness', or 'herbaceous'). If no information on growth form was available through the aforementioned steps, information from herbarium sheets, pictures of herbarium sheets, or individuals in the field were used. Lastly, if a species was identified as woody and not in any of the aforementioned studies, the available phylogenetic literature was used to reconstruct the evolutionary history of the woodiness in the clade. The list of all species with the assigned growth forms is provided in Table S1.

Environmental Data Analysis

Correlation Analysis

We performed Kendall's correlation tests to remove strongly correlated environmental variables ($\tau > 0.7$; Figure S1), and ran the analyses using the balanced sampling approach. To allow for comparison between results, we used the same subset of ten variables that were retained for subsequent analyses using the strict and lenient approaches.

Phylo-PCA & Phylo-ANOVA

To construct a phylogeny for our analyses, we used V.PhyloMaker2 (Jin and Qian 2022), a software for generating plant phylogenies based on the vascular plant supertree of Smith and Brown (2018) and the pteridophyte phylogeny of Zanne et al. (2014). For each sampling regime, we produced a separate phylogeny using a list of all the species which remained after locality data filtering, with the 'GBOTB.extended.TPL' tree selected, with tree and nodes arguments set to default. All species in our input list were recovered in the phylogeny. This phylogeny can also be replaced by an input tree (newick file) from the user if plants are not the group of interest.

To visualise the major axes of variation across the environmental variables, while accounting for the shared evolutionary history among taxa, we performed a phylogenetic principal component analysis (hereafter 'phylo-PCA') as described by Revell (2009; Figure 2), using the 'phyl.pca' function from the 'phytools' package (Revell 2024). Before the PCA analysis, the data were cen-

Table 1. Bioclimatic and topographic variables used in this study.

Variable	Source	Temporal coverage	Unit
Elevation	CanaryClim	-	m
Mean annual air temperature	CanaryClim	1997 - 2013	°C
Mean diurnal annual air temperature range	CanaryClim	1997 - 2013	°C
Temperature seasonality	CanaryClim	1997 - 2013	°C
Mean daily air temperature of the driest quarter	CanaryClim	1997 - 2013	°C
Mean daily air temperature of the coldest quarter	CanaryClim	1997 - 2013	°C
Annual precipitation amount	CanaryClim	1997 - 2013	mm
Precipitation amount of the driest quarter	CanaryClim	1997 - 2013	mm
Precipitation seasonality	CanaryClim	1997 - 2013	%
Actual evapotranspiration and interception	FAO - WaPOR	2018 - 2023	mm
Evaporation	FAO - WaPOR	2018 - 2023	mm
Transpiration	FAO - WaPOR	2018 - 2023	mm
Interception	FAO - WaPOR	2018 - 2023	mm
Terrain Ruggedness Index (TRI)	-	-	-
(Adjusted Thornthwaite) aridity index	*	*	-
Cloud Cover	Atlas Climático de Canarias	2000 - 2014	%

*The adjusted Thornthwaite aridity index was calculated using the median annual precipitation and evapotranspiration as described in equation (1); as such, it has mixed sources and temporal coverages.

tred and scaled to ensure equal contribution to the model. Additionally, we colour-coded the points based on their growth form.

To test for differences in the environmental variables across the five different growth forms, while accounting for their shared evolutionary history, we used phylogenetic Analysis of Variance (phylo-ANOVA; Figure 3). The 'phylANOVA' function in the 'phytools' R package (Revell 2024) was used, which simulates trait evolution under a Brownian motion model along the provided phylogenetic tree to generate a null distribution of F-statistics (Garland Jr et al. 1993). The significance of group differences was assessed via 1,000 simulations, and p-values were computed as the proportion of simulated F-values exceeding the observed value. P-values were adjusted for multiple comparisons using the correction procedure by Holm (1979).

Rugosity (La Palma case study)

To further explore the differences between growth forms found by the phylo-ANOVA, we plotted the base terrain ruggedness index (TRI) raster layer against the proportion of the occurrences of each growth form, relative to the total occurrences, per grid cell for the island of La Palma (Figure 4). We used La Palma as a case study due to the clarity of the valleys and ravines on the rugos-

ity raster layer compared to the other islands. We used the 'geom_sf' package in ggplot2, where each grid cell corresponds to ~300m^2 (Wickham 2016). Visualising the proportion of each growth form per grid cell, relative to the total occurrences within the grid cell, highlights regions where one growth form predominates. Visualising the raw raster file also provides context for how gradients observed over topographically complex regions relate to the distribution of each growth form. For the same analyses repeated across the other Canarian islands, see Figures S2-7.

RESULTS

Locality Data Extraction

Of our 1,160 target species, we obtained growth form data for 1,155 species, and 974 of these had locality data in GBIF (Table 2, Table S1). Table 2 shows the number of species retained after filtering for each of the three sampling regimes across all five growth forms. The percentage of species remaining from the total number of species for which GBIF locality data were extracted is also reported. As expected, the lenient sampling regime has the most species retained after filtering, with 803 of the 974 species retained (82%); the strict sampling regime has the fewest species retained, at 396 (41%); and the balanced sampling regime retains 548 species (56%).

The results and figures in the following sections are shown for the balanced sampling regime, unless stated otherwise (for results from strict and lenient sampling regimes, see Figures S10 & S11). A minimum number of 13 occurrences (as in the balanced regime) implies both narrow endemic species as well as more widespread species are included, following van Proosdij et al. (2016), and an uncertainty limit of 500 m balances spatially accurate distributions with high species representation. However, we acknowledge that given the high environmental variability within the Canary Islands, specimens with the maximum uncertainty of 500 m resolution may be too coarse to capture the true climatic conditions of species.

Environmental Niche Space (Phylo-PCA)

To investigate the environmental variables associated with different plant growth forms in the Canary Islands and address our first and second hypotheses, we used a phylo-PCA to map the environmental niche space of five growth forms (Figure 2, Table S3). We find a broad overlap between the growth forms' environmental niche spaces (confidence ellipses), aggregating in the centre of island environmental niche space (Figure 2a). However, there are some dis-

Table 2. The number of species in each growth form category targeted in this study, with locality data in GBIF, for each sampling approach (lenient, balanced, and strict), with the percentage of species with locality data remaining after filtering for each sampling regime.

Growth Form	Total in target list	Total with GBIF data	Lenient approach	% after filtering (lenient)	Balanced approach	% after filtering (balanced)	Strict approach	% after filtering (strict)
Annual Herbaceous	371	297	231	77.8	153	51.5	91	30.6
Perennial Herbaceous	320	263	222	84.4	161	61.2	126	47.9
Ancestral Woodiness	126	105	89	84.8	63	60.0	55	52.4
Derived Woodiness	97	85	67	78.8	47	55.3	33	38.8
Insular Woodiness	241	224	194	86.6	124	55.4	91	40.6
Total	**1155**	**974**	**803**	**82.4**	**548**	**56.3**	**396**	**40.7**

tinctions in the direction of environmental niche space occupied by each of the growth forms (Figure 2a), suggesting that the environmental variables have differing importance for each growth form. For example, IW species are placed mostly along a climatic seasonality axis, linked to both precipitation and temperature (PC1 in Figure 2), and to a minor degree, rugosity (TRI_median). The IW environmental niche space shows strong convergence with the AW niche, though the latter is less linked to increased seasonality and more related to increased interception, i.e. the precipitation caught by vegetation that evaporates before reaching the ground. In contrast, the niche space of herbaceous species is mostly dispersed along the temperature-seasonality-elevation-precipitation axis (PC2 in Figure 2a), with precipitation seasonality and cloud cover (PC1 in Figure 2) more important in PH than AH. Unexpectedly, DW niche space shows greater congruence with the herbaceous growth forms than with the other woody growth forms; a finding likely attributed to DW species occurring more prevalently in coastal regions than other woody growth forms (Figure 2a, 2b). For how each variable contributes to the principal component axes, see Figure S8.

To assess the validity of the extracted environmental variables (Table S3), we compared our recovered environmental niche space to the main vegetation zones present in the Canaries following Pérez (1999) and del Arco-Aguilar & Rodríguez Delgado (2018): (dry) coastal region, the thermo-sclerophyllous forests, laurisilva forest, pine forest region, and the high-altitude region. We highlight ~100 species that are typical of a single vegetation zone in niche space (Figure 2b) to show the phylo-PCA largely clusters species into coherent regions (Pérez 1999; del Arco-Aguilar & Rodríguez Delgado 2018): coastal species cluster in arid regions of low precipitation and low rugosity, given by smaller PC1 values and larger PC2 values (Figure 2a, S2, Table S3), and laurisilva specialists

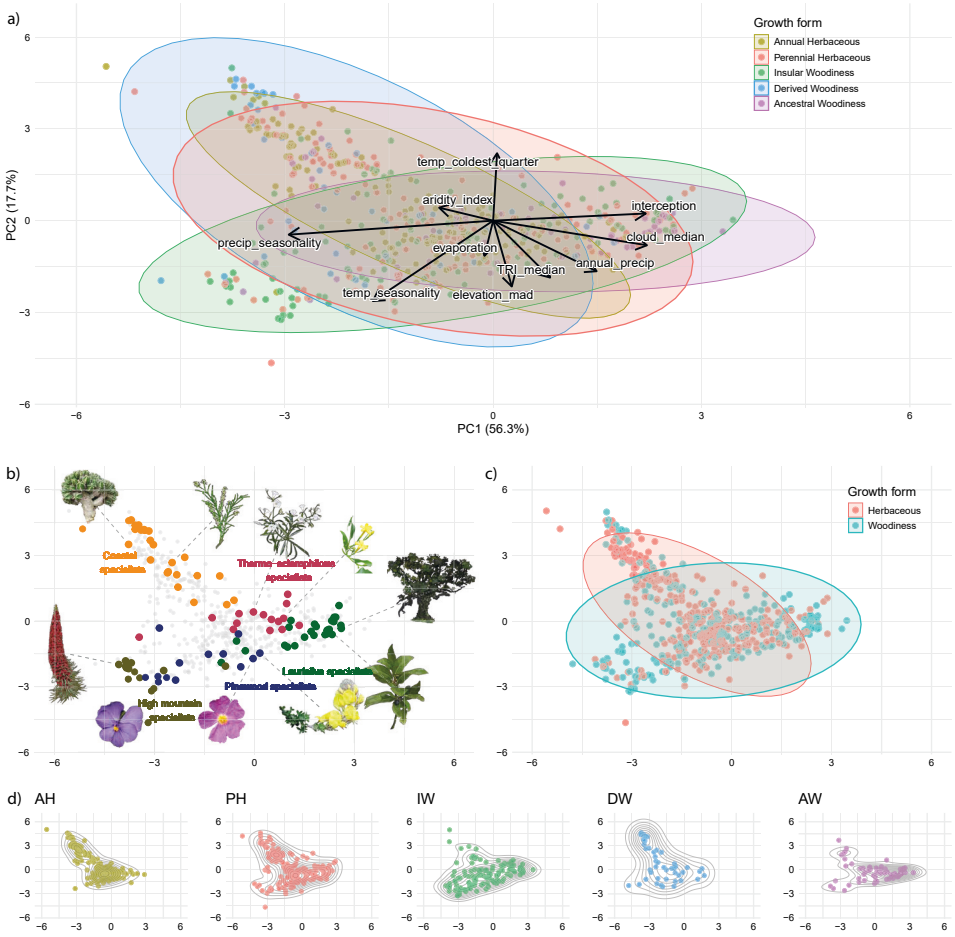


Figure 2. Environmental niche space of the growth forms of the Canary Island non-monocot angiosperm flora. a) Phylogenetic PCA of the different environmental variables labelled per growth form for the 548 species retained using the balanced sampling regime. Each point represents a species, shaded ellipses represent 95% confidence intervals around the points for each growth form, and the arrows represent the loadings of the phylo-PCA to highlight the contribution of each variable. b) Major Canarian vegetation zones for 99 species plotted onto the phylo-PCA, with a selection of specialists highlighted, using Pérez (1999) and del Arco-Aguilar & Rodríguez Delgado (2018). Plant images from top-left, clockwise: *Euphorbia handiensis*, *Tamarix canariensis*, *Convolvulus floridus*, *Chrysojasminum odoratissimum*, *Ocotea [Mespilodaphne] foetens*, *Laurus novocanariensis*, *Adenocarpus foliolosus*, *Cistus symphytifolius*, *Viola cheiranthifolia*, *Echium wildpretii*. For the highlighted species labelled on the phylo-PCA, see Figure S9. c) Environmental niche space labelled in groups: herbaceous versus woody; and d) separated into five distinct growth forms: AH = Annual Herbaceous, PH = Perennial Herbaceous, IW = Insular Woody, DW = Derived Woody, AW = Ancestral Woody. Contour lines represent 2D kernel density estimation. For the sources of the plant images, see Table S4.

are found in areas of high interception (indicative of closed canopy cover), high cloud cover, and low seasonality, given by more positive PC1 values (Figures 2a, 2b, S2, Table S3).

To further explore broad differences in niche space and test our first hypothesis, we aggregated our growth form data into binary woody and herbaceous growth forms (Figure 2c). As with Figure 2a, overlap in niche space is observed centrally, but the niche space can be separated along the temperature-seasonality-elevation-precipitation axis (PC2), likely resulting from the dominance of herbaceous species in coastal regions.

When subdividing our phylo-PCA into five plots by growth form from AH towards AW, a slight step-wise shift away from a predominance in arid and aseasonal niche space typical of AH species becomes visible (top-left of niche space in Figure 2d, given by high PC2 values). Though all growth forms are present throughout niche space, progressively, the density of occurrences in niche space gently shifts towards rugose, more seasonal regions typical of IW species (bottom-left in Figure 2d, given by low PC1 values), or aseasonality and high interception (bottom-right in Figure 2d, given by high PC1 values) in AW species, whilst DW species are widely distributed across niche space.

The results in Figure 2 (using the balanced sampling regime) are largely congruent with those using other sampling regimes, suggesting 500 m coordinate uncertainty may not be too coarse to capture climatic preferences of species. Across sampling regimes, overlapping clusters are observed centrally, with differences between niche spaces found largely in the directionality of the confidence ellipses. In the strict sampling regime, the directionality of the niche space for each growth form was further pronounced, with greater convergence between AH and PH niche space, and between IW and AW niche space (Figure S10). DW niche space remains widespread throughout niche space, but with increased overlap with the AH niche. For the lenient sampling regime, there is greater overlap between the growth forms, with little distinction between the herbaceous and woody niche in niche space (Figure S11), but the relative importance of each variable on each growth form remains similar.

Pairwise Comparisons (Phylo-ANOVA)

To statistically analyse the differences between the localities where the growth forms occur, in terms of the environmental variables, and test our third hypothesis, we used a phylo-ANOVA performing pairwise comparisons between each growth form for each environmental variable (Figure 3). We find significant differences between at least two growth forms in five of the ten variables analysed. Rugosity has the most pairwise significant differences between growth forms, with significance found in four pairwise tests (Figure 3b). Generally, AH species have the lowest rugosity, with rugosity being higher in woody growth forms than herbaceous ones, and IW species having the highest observed median rugosity. Despite both temperature and precipitation seasonality shap-

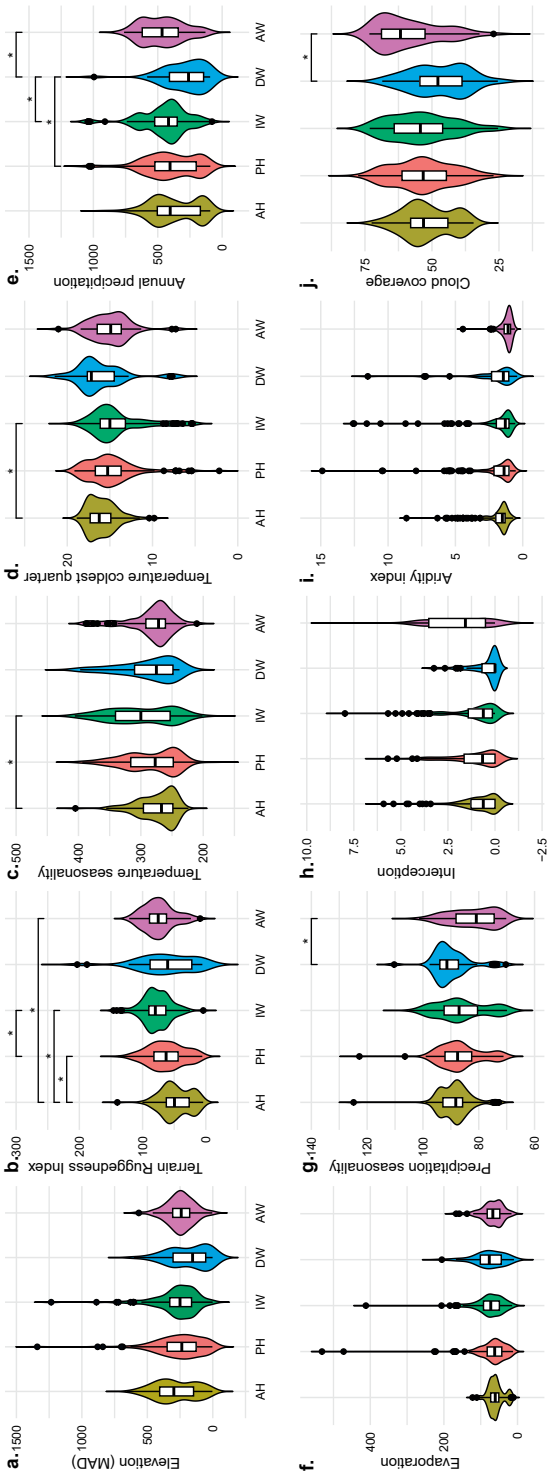


Figure 3. Results of the phylo-ANOVA. Boxplots show the distribution of different environmental variables across each growth form: a) Elevation (MAD), b) Terrain ruggedness index, c) Temperature seasonality (°C*100), d) Temperature in the coldest quarter (°C), e) Annual precipitation (mm), f) Evaporation (mm), g) Precipitation seasonality (%), h) Square root of the interception (mm), i) Adjusted Thornthwaite aridity index, and j) Cloud Coverage (%). The violin plots show the distribution of each variable per growth form and the median value for each environmental variable for each growth form. AH = annual herbaceousness, PH = perennial herbaceousness, IW = insular woodiness, DW = derived woodiness, AW = ancestral woodiness. Significant group differences (Holm-adjusted p-value < 0.05; phylo-ANOVA) are indicated with *. The square root of the interception is used due to the spread of the raw data.

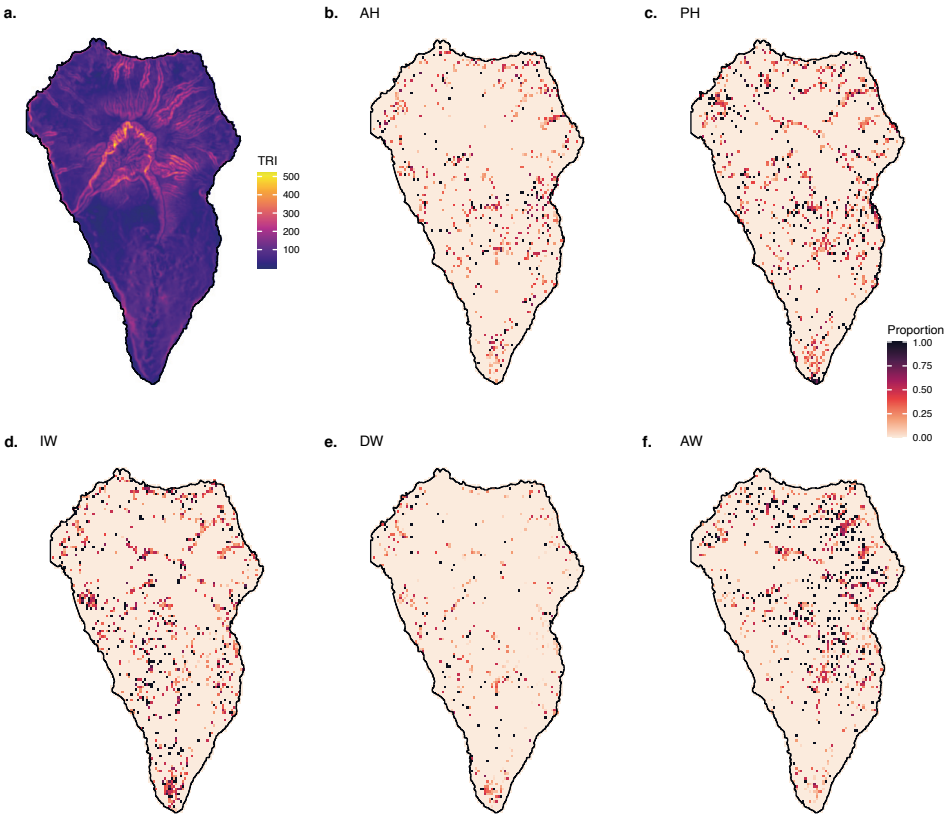


Figure 4. Rugosity in La Palma Island and the distribution of different growth forms, using the proportion of each growth form relative to the total occurrences per grid cell. a) Raster of terrain ruggedness index, b) AH = annual herbaceousness, c) PH = perennial herbaceousness, d) IW = insular woodiness, e) DW = derived woodiness, and f) AW = ancestral woodiness. Peach grid squares can refer to zero proportion of a growth form or a grid cell with no data. For the distribution of the growth forms on the other Canary Islands, see Figures S2-7.

ing environmental space (Figure 2a), there is only one significant pairwise difference for each of these variables (Figure 3c, g). IW species have significantly higher temperature seasonality than AH species, with the highest median temperature seasonality across all growth forms (Figure 3c). AW species have significantly lower precipitation seasonality than DW species, the former having the lowest observed median precipitation seasonality and the latter the highest (Figure 3g). Additionally, AW species occur in places with significantly higher cloud coverage than DW species, with the former having the highest observed median cloud coverage and the latter the lowest (Figure 3j). Three significant pairwise differences are found for annual precipitation (Figure 3e), between DW and PH, DW and IW, and DW and AW. IW species are also found in significantly colder regions (in the coldest quarter) than AH species (Figure 3d).

Rugosity (La Palma Case Study)

As the phylo-ANOVA revealed rugosity can best separate the five growth forms (Figure 3b), we explored the differences between growth forms along the rugosity variable further (Figure 4). By plotting the base TRI raster alongside the proportion of each growth form per grid cell, we can visualise the relationship between rugosity and areas where each growth form is particularly prevalent. We find that high proportions of IW species occurrence are found in high rugosity areas, such as around the central Caldera de Taburiente and the ravines and narrow valleys with steep flanks scattered in the northern and central part of the island (Figure 4d). The high proportion of PH species in highly rugose areas (Figure 4c), as well as across the lower rugosity southern regions of La Palma, reflects the wide range of TRI values obtained for this growth form (Figure 3b). Figure 4 also highlights the predominance of AW species within the (rugose) laurisilva forest of La Palma (Figure 4f), and the preference of AH species for drier, coastal areas of lower rugosity (Figure 4b), in agreement with the findings of the phylo-PCA (Figure 2). Rugosity, therefore, emerges as an important variable for the distribution of growth forms in the Canary Islands.

DISCUSSION

Here, we present a comprehensive analysis of the associations between phylogenetically inferred growth forms and environmental variables for the Canary Islands' (Figure 1) non-monocot angiosperm community, comprising in total 1,160 native species (Figure 2). Using a novel bioinformatics pipeline, we build on existing locality data-filtering methods by incorporating functions to remove occurrences on islands to which a given species is not native or to which occurrence records are reprojected from raster data into vector data, and allowing for selecting pre-defined thresholds of coordinate uncertainty and number of occurrences. After extracting and filtering GBIF occurrence data for 974 species of the non-monocot flora (84%), we extracted and aggregated environmental data for each species (Table S3), and visualised environmental niche space for the 548 species that passed data filtering using a phylo-PCA based on the balanced sampling approach (Figure 2). For the final steps of our approach, we used phylo-ANOVAs to statistically analyse the pairwise differences between the environmental variables for each growth form to establish which variables are important for segregating the niche space of these growth forms (Figures 2, 3).

We highlighted ~100 species on the phylo-PCA that occur in a single major Canarian vegetation zone (Figure 2b), following Pérez (1999) and del Arco-Aguilar & Rodríguez Delgado (2018), mapping the species' names, which suggests our pipeline produces results consistent with known habitat occurrences (Figure S9). For example, known specialists of high montane systems such as *Echium wildpretii* and *Viola cheiranthifolia* are recovered well within regions of niche space linked to high elevation and seasonality, given by small PC1 and PC2 values (Figures 2b, S2, Table S3; Rodríguez-Rodríguez et al. 2019; Graham et al. 2021). The iconic cactiform *Euphorbia* species typical of the xerophytic lowland communities ("cardonal-tabaibal"), such as *E. canariensis* and *E. handiensis* (Coello et al. 2024), are mapped in high aridity, and low elevation and rugosity areas of niche space. Endemic species of the laurisilva subtropical forest, such as *Laurus novocanariensis* and *Ocotea foetens,* occur in regions of niche space dominated by high precipitation, cloud coverage and interception (precipitation caught by vegetation that evaporates before reaching the ground), with low seasonality (Figure 2a, Table S3; Betzin et al. 2016). Our results also resemble dos Santos et al. (2022), who found a triangular shape with coastal regions, high montane systems, and the laurisilva forests on the corners of the niche space.

Though there are several existing pipelines to clean and analyse occurrence data (e.g., Zizka et al. 2019), and subsequently extract environmental variables from spatial data (van den Hoogen et al. 2021; Velazco et al. 2022; Coca-de-la-Iglesia et al. 2023; Jemeļjanova et al. 2024), our approach goes beyond these first steps by incorporating existing phylogenetically-corrected visualisation and statistical methods. This means our approach allows us to visualise environmental niche space for any predefined set of categorical traits (e.g., growth form, taxonomy/clade), rather than only exploring the extracted environmental data in a species distribution modelling context to predict occurrence as implemented in existing pipelines (van den Hoogen et al. 2021; Velazco et al. 2022). Our method also allows for greater precision in delimiting the native range of a species than in existing methods based on the geographical codes of the Taxonomic Databases Working Group (TDWG; Brummitt 2001). For example, by using island-level native status across the Canary Islands from published floras (Beierkuhnlein et al. 2021; Canary Islands Biodiversity Database 2024), we can more accurately map the native range of a species across the archipelago than methods that use TDWG codes that would treat the Canary Islands as a single region. Finally, as we use floras as input in our approach, we provide an alternative method for accounting for taxonomy and synonymy than the previously existing 'scrubr' R package (Chamberlain 2025).

Substantial overlap in environmental niche space for growth form types in the Canary Islands flora

We hypothesised that the marked differences in life span, woodiness, and evolutionary history across the five growth forms should lead to pronounced differences in environmental niche space. However, we found considerable niche overlap across the five growth forms centrally in niche space (Figure 2), e.g. as can be seen by the confidence ellipses (Figure 2a). For example, annual and perennial herbaceous growth forms (AH, PH) show marked overlap, despite differences in life span. Furthermore, derived woody species (that acquired their woodiness before they colonised the islands; DW) overlap considerably with AH species (Figure 2a), likely explained by the predominance of both growth forms in coastal areas (Figure 2a, b). Nevertheless, the niche space of the growth forms does show some distinctions, providing mixed support for our first hypothesis, suggesting that there are detectable differences in niche space between combined herbaceous and combined woody category and the five separate growth form categories. Interestingly, in support of the second hypothesis, indicating that the insular woody (IW) niche will deviate most from those of the herbaceous growth forms, IW species indeed exhibit distinctions from the herbaceous growth forms (Figure 2a), and show strong convergence to occupy similar environmental space as ancestrally woody (AW) species, although differing in the effect of seasonality. This marked convergence in niche space between IW and AW, despite woodiness evolving differently across evolutionary history (i.e., de-novo wood evolution on the island versus woodiness conserved over evolutionary time), hints at the existence of a 'woody niche' in niche space (Lens et al. 2013a). However, following this line of thought, we would expect the woody niche to be more distinct as postulated in our first hypothesis (Figure 2c). We do find some support for this when the niches of the three grouped woody growth forms are compared with the two herbaceous growth forms, but the woody-herbaceous niche overlap remains considerable (Figure 2c). These findings are also robust under different sampling regimes (Figures S10, S11).

Unexpectedly, and contrasting previous studies and our third hypothesis (that variables related to aridity and seasonality will impact the niche of IW species; Carlquist 1974; Zizka et al. 2022), IW species are found in areas of higher climatic seasonality (Figures 2a, d), given by smaller PC1 values (Figure 2a, S2, Table S3). Most variation in the niche space of IW species is explained by the precipitation-seasonality-cloud-cover-interception axis (PC1 in Figure 2, Figure S8), and IW species exhibit the highest overall median temperature seasonality (Figure 3c), as well as the lowest overall median temperature in the coldest quarter (Figure 3d). This suggests greater tolerance to seasonally lower

temperatures than herbaceous species, contrasting global studies (Klimeš et al. 2022). However, none of the IW species analysed occur in areas where frost occurs, as none grow in regions where the temperature in the coldest quarter reaches the frost point (Figure 3d). An aseasonal climate, defined by Carlquist (1974) as growth uninterrupted by frost, is therefore still characteristic for the IW Canary Islands species, and could still serve as a potential driver favouring woodiness, despite seasonality being important in shaping the IW environmental niche space (Figures 2a, 3g). Another unexpected finding of our study is that the IW niche is not defined by parameters indicating increased drought, contrary to previous studies (Lens et al. 2013b; Doria et al. 2018; Hooft van Huysduynen et al. 2021; Zizka et al. 2022). This does not exclude the possibility, however, that drought did play an important role in promoting the evolution of woodiness on the Canary Islands for at least some lineages (e.g., Doria et al. 2018), and it matches the outcome of the global derived woodiness review suggesting that drought is a more important global woodiness driver on continents compared to islands (Zizka et al. 2025).

Compared to previous studies, we recover several established relationships. For instance, AH species favour arid regions of low precipitation and elevation, and consistently high temperatures (Figures 2a, d, 3c, d, Table S3), consistent with previous analyses (Irl et al. 2020; Poppenwimer et al. 2023; Taylor et al. 2023). Also, PH species, previously assessed as hemicryptophytes or subshrubs, show similar correlates as AH species but with greater importance of increasing precipitation and elevation, and lower temperatures (Irl et al. 2020; Taylor et al. 2023). This relationship is also recovered here (Figure 2a), as given by the PH niche space being shaped more by the temperature-seasonality-elevation-precipitation PC2 axis compared to AH species (Figure S8). Likewise, high precipitation, cloud cover, and low climatic seasonality (both precipitation and temperature) are correlated with the distribution of AW species (Figures 2, 3c, e, g, j), explaining nearly all variation across PC1 (Figure 2), in agreement with previous studies who classify AW species as phanerophytes or trees (Poppenwimer et al. 2023; Taylor et al. 2023).

The trait that best shapes plant growth form in the Canaries is topographic complexity, included in our approach as rugosity (TRI variable). IW species are those that are found to occur generally in areas with the highest rugosity, significantly higher than both herbaceous growth forms and significantly higher in AW than AH species (Figure 3b). Additionally, PH species are found in more topographically complex areas than AH species (Figure 3b). These findings demonstrate that woody species, and to a lesser extent perennials, are abundant in more topographically complex regions, which could be explained by two factors. First, the current distributions of woody species in highly rugose

areas may be relict populations, reflecting grazing pressure from large mammal herbivores introduced to the archipelago over the last ~2,000 years (Castilla-Beltrán et al. 2021). Alternatively, over longer evolutionary timescales, topographically complex regions may create microclimates and/or micro-habitats, which could spur speciation events due to limited seed dispersal, suppressed gene flow between isolated populations, and increased climatic diversity along steep elevational gradients (Vitales et al. 2014; Chang et al. 2023; Graham et al. 2025). Therefore, the recovered relationship between (insular) woodiness and rugosity may stem from a combination of a lack of herbivory in these potentially relict populations affecting extant distribution patterns, and the presence of IW species in areas of increased opportunity for allopatric speciation. As most IW species on the Canaries belong to evolutionary radiations (Carine et al. 2010; Lens et al. 2013b), other factors that promote diversification within these lineages may allow for rapid diversification along the steep elevational gradient found in these highly rugose areas, thereby creating the correlation between rugosity and IW.

In conclusion, based on our approach, spanning the extraction of specimen records to phylogenetically-inferred comparison and visualisation of environmental niche space, we find substantial overlap across the niches of five plant growth forms on the Canaries, with significant differences in the environmental correlates of various woody and herbaceous growth forms across the Canary Islands. Climatic seasonality (both temperature and precipitation) and topographic complexity emerge as key distinguishing factors in driving differences in environmental niche space. Although the environmental niche space overlap across the growth forms is unexpectedly high given the differences in life span and evolutionary history, our results for AH, PH and AW species confirm previous results based on global analyses. Surprisingly, for IW species, the environmental niche space is shaped by seasonal colder (but frost-free) temperatures without any correlation to drought-related parameters. Our results also highlight the importance of separating species into growth forms congruent with their evolutionary history (e.g., ancestral woodiness versus insular woodiness), as derived woody island species show greater similarity in their environmental niche to herbaceous species than to other woody growth forms, with the former primarily linked to temperature, elevation, and precipitation (or lack thereof). Overall, we hope our approach, which can be applied to any categorical trait of interest, will promote phylogenetically-inferred macroecological studies across the world, thereby contributing to how traits have shaped environmental niche space for any plant-animal-fungi lineage.

ACKNOWLEDGEMENTS

We thank the Center for Information Technology of the University of Groningen for their support and for providing access to the Hábrók high-performance computing cluster.

AUTHOR CONTRIBUTIONS

Conceptualisation: L.J., R.F.A.B., L.V., F.L.; Environmental variable selection: L.J., R.F.A.B., J.C.-C., L.V., F.L.; Analyses: L.J., R.F.A.B.; Writing first draft: L.J., R.F.A.B., L.V., F.L.; Supervision: R.F.A.B., L.V., F.L. All co-authors contributed to editing and reviewing the manuscript.

CODE AVAILABILITY

The code and data required to recreate all analyses, as well as the pipeline used for extraction and filtering of occurrences, and the supplementary tables, can be found in the supporting GitHub repository at https://github.com/ljn261/canary_islands_project.

FUNDING

This work was supported by the Dutch Research Council [OCENW.KLEIN.498] to F.L.

REFERENCES

Baldwin, B. G., & Sanderson, M. J. (1998). Age and rate of diversification of the Hawaiian silversword alliance (Compositae). *Proceedings of the National Academy of Sciences*, 95(16), 9402-9406.

Barajas Barbosa, M. P., Craven, D., Weigelt, P., Denelle, P., Otto, R., Díaz, S., ... & Kreft, H. (2023). Assembly of functional diversity in an oceanic island flora. *Nature*, 619(7970), 545-550.

Beierkuhnlein, C., Walentowitz, A., & Welss, W. (2021). FloCan—A revised checklist for the flora of the Canary Islands. *Diversity*, 13(10), 480.

Betzin, A., Thiv, M., & Koch, M. A. (2016). Diversity hotspots of the laurel forest on Tenerife, Canary Islands: a phylogeographic study of Laurus and Ixanthus. *Annals of Botany*, 118(3), 495-510.

Böhle, U. R., Hilger, H. H., & Martin, W. F. (1996). Island colonization and evolution of the insular woody habit in Echium L.(Boraginaceae). *Proceedings of the National Academy of Sciences*, 93(21), 11740-11745.

Brummitt, R. K. (2001). World Geographical Scheme for Recording Plant Distributions.

Canary Islands Biodiversity Database, 2024. Government of the Canary Islands. (https://www.biodiversidadcanarias.es/biota) [10th October 2023].

Carine, M. A., Santos-Guerra, A., Guma, I. R., & Reyes-Betancort, J. A. (2010). *Endemism and evolution of the Macaronesian flora*. Beyond cladistics: The branching of a paradigm, 101-124.

Carlquist, S. (1974). *Island Biology*. Columbia University Press.

Carlquist, S. (2012). Monocot xylem revisited: New information, new paradigms. *The Botanical Review*, 78, 87–153.

Castilla-Beltrán, A., De Nascimento, L., Fernández-Palacios, J. M., Whittaker, R. J., Willis, K. J., Edwards, M., & Nogué, S. (2021). Anthropogenic transitions from forested to human-dominated landscapes in southern Macaronesia. *Proceedings of the National Academy of Sciences*, 118(40), e2022215118.

Chamberlain S (2025). _scrubr: Clean Biological Occurrence Records_. R package version 0.4.0, commit 4c02b568e81ea16ef60f913ced88a42c155d8f19, <https://github.com/ropensci/scrubr>.

Chamberlain, S., Barve, V., Mcglinn, D., Oldoni, D., Desmet, P., Geffert, L., & Ram, K. (2024). *rgbif*: Interface to the Global Biodiversity Information Facility API. https://CRAN.R-project.org/package=rgbif.

Chang, Y., Gelwick, K., Willett, S. D., Shen, X., Albouy, C., Luo, A., Wang, Z., Zimmermann, N. E., & Pellissier, L. (2023). Phytodiversity is associated with habitat heterogeneity from Eurasia to the Hengduan Mountains. *New Phytologist*, 240(4), 1647–1658.

Coca-De-La-Iglesia, M., Valcárcel, V., & Medina, N. G. (2023). A protocol to retrieve and curate spatial and climatic data from online biodiversity databases using R. *Bio-protocol*, 13(20), e4847.

Coello, A. J., Vargas, P., Cano, E., Riina, R., & Fernández-Mazuecos, M. (2024). Phylogenetics and phylogeography of Euphorbia canariensis reveal an extreme Canarian-Asian disjunction but limited inter-island colonization. *Plant Biology,* 26(3), 398-414.

Darwin, C. (1859). On the origin of species: Facsimile of the first edition.

del Arco Aguilar, M. J. & Rodríguez Delgado, O. (2018). *Vegetation of the Canary Islands* (pp. 83-319). Springer International Publishing.

Dória, L. C., Podadera, D. S., del Arco, M., Chauvin, T., Smets, E., Delzon, S., & Lens, F. (2018). Insular woody daisies (*Argyranthemum*, Asteraceae) are more resistant to drought-induced hydraulic failure than their herbaceous relatives. *Functional Ecology*, 32(6), 1467–1478.

dos Santos, P., Brilhante, M. Â., Messerschmid, T. F., Serrano, H. C., Kadereit, G., Branquinho, C., & De Vos, J. M. (2022). Plant growth forms dictate adaptations to the local climate. *Frontiers in Plant Science*, 13, 1023595.

Doyle, J. A. (2012). Molecular and fossil evidence on the origin of angiosperms. *Annual Review of Earth and Planetary Sciences*, 40, 301–326.

Evans, J. S., & Murphy, M. A. (2023). *spatialEco*. https://github.com/jeffreyevans/spatialEco.

FAO. (2020). WaPOR Database Methodology: Version 2 Release, April 2020. FAO Rome, Italy.

Florencio, M., Patiño, J., Nogué, S., Traveset, A., Borges, P. A., Schaefer, H., Amorim, I. R., Arnedo, M., Ávila, S. P., Cardoso, P., & others. (2021). Macaronesia as a fruitful arena for ecology, evolution, and conservation biology. *Frontiers in Ecology and Evolution*, 9, 718169.

Friedman, J. (2020). The evolution of annual and perennial plant life histories: Ecological correlates and genetic mechanisms. *Annual Review of Ecology, Evolution, and Systematics*, 51, 461–481.

Garland Jr, T., Dickerman, A. W., Janis, C. M., & Jones, J. A. (1993). Phylogenetic analysis of covariance by computer simulation. *Systematic Biology*, 42(3), 265–292.

GBIF. (2025). Global Biodiversity Information Facility. https://www.gbif.org/what-is-gbif.

Gerrienne, P., Gensel, P. G., Strullu-Derrien, C., Lardeux, H., Steemans, P., & Prestianni, C. (2011). A simple type of wood in two Early Devonian plants. *Science*, 333(6044), 837–837.

Graham, R. E., Reyes-Betancort, J. A., Chapman, M. A., & Carine, M. A. (2021). Inter-island differentiation and contrasting patterns of diversity in the iconic Canary Island sub-alpine endemic Echium wildpretii (Boraginaceae). *Systematics and Biodiversity*, 19(5), 507-525.

Graham, C. H., Araujo, M. L., Barreto, E., Dambros, C. S., Diniz-Filho, J. A. F., Zimmermann, N. E., ... & Coelho, M. T. P. (2025). Biodiversity patterns redefined in environmental space. *Ecology Letters*, 28(2), e70008.

Hernández-Hernández, T., & Wiens, J. J. (2020). Why are there so many flowering plants? A multiscale analysis of plant diversification. *The American Naturalist*, 195(6), 948–963.

Herschy, R., Herschy, R., Singh, V., Aravamuthan, V., Singh, V., Church, M. R., Oliver, J. E., Herschy, R., Simpson, J. R., Griffiths, J. F., & others. (1998). Aridity indices. *Encyclopedia of Hydrology and Water Resources*, 78–83.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. Scandinavian *Journal of Statistics*, 65–70.

Hooft van Huysduynen, A., Janssens, S., Merckx, V., Vos, R., Valente, L., Zizka, A., Larter, M., Karabayir, B., Maaskant, D., Witmer, Y., & others. (2021). Temporal and palaeoclimatic context of the evolution of insular woodiness in the Canary Islands. *Ecology and Evolution*, 11(17), 12220–12231.

IPNI. (2024). *International Plant Names Index*. http://www.ipni.org.

Irl, S. D., Obermeier, A., Beierkuhnlein, C., & Steinbauer, M. J. (2020). Climate controls plant life-form patterns on a high-elevation oceanic island. *Journal of Biogeography*, 47(10), 2261-2273.

Jemeljanova, M., Kmoch, A., & Uuemaa, E. (2024). Adapting machine learning for environmental spatial data-A review. *Ecological Informatics*, 81, 102634.

Jin, Y., & Qian, H. (2022). V. PhyloMaker2: An updated and enlarged R package that can generate very large phylogenies for vascular plants. *Plant Diversity*, 44(4), 335-339.

Klimeš, A., Šímová, I., Zizka, A., Antonelli, A., & Herben, T. (2022). The ecological drivers of growth form evolution in flowering plants. *Journal of Ecology*, 110(7), 1525–1536.

Körner, C. (2007). The use of 'altitude' in ecological research. *Trends in Ecology & Evolution*, 22(11), 569–574.

Lens, F., Davin, N., Smets, E., & del Arco, M. (2013a). Insular woodiness on the Canary Islands: a remarkable case of convergent evolution. *International Journal of Plant Sciences*, 174(7), 992-1013.

Lens, F., Tixier, A., Cochard, H., Sperry, J. S., Jansen, S., & Herbette, S. (2013b). Embolism resistance as a key mechanism to understand adaptive plant strategies. *Current Opinion in Plant Biology*, 16(3), 287–292.

Liu, H., Jiang, Z., Dai, J., Wu, X., Peng, J., Wang, H., Meersmans, J., Green, S. M., & Quine, T. A. (2019). Rock crevices determine woody and herbaceous plant cover in the karst critical zone. *Science China Earth Sciences*, 62, 1756–1763.

Maldonado, C., Molina, C. I., Zizka, A., Persson, C., Taylor, C. M., Albán, J., Chilquillo, E., Rønsted, N., & Antonelli, A. (2015). Estimating species diversity and distribution in the era of Big Data: To what extent can we trust public databases? *Global Ecology and Biogeography*, 24(8), 973–984.

Morris, J. L., Puttick, M. N., Clark, J. W., Edwards, D., Kenrick, P., Pressel, S., Wellman, C. H., Yang, Z., Schneider, H., & Donoghue, P. C. (2018). The timescale of early land plant evolution. *Proceedings of the National Academy of Sciences*, 115(10), E2274–E2283.

Müller, J. W. (2000). Possible advantages of a robust evaluation of comparisons. *Journal of Research of the National Institute of Standards and Technology*, 105(4), 551.

Nürk, N. M., Linder, H. P., Onstein, R. E., Larcombe, M. J., Hughes, C. E., Piñeiro Fernández, L., Schlüter, P. M., Valente, L., Beierkuhnlein, C., Cutts, V., & others. (2020). Diversification in evolutionary arenas–Assessment and synthesis. *Ecology and Evolution*, 10(12), 6163–6182.

Patiño, J., Collart, F., Vanderpoorten, A., Martin-Esquivel, J. L., Naranjo-Cigala, A., Mirolo, S., & Karger, D. N. (2023). Spatial resolution impacts projected plant responses to climate change on topographically complex islands. *Diversity and Distributions*, 29(10), 1245–1262.

Pérez, M. Á. C. (1999). *Native flora of the Canary islands*. Everest.

Poppenwimer, T., Mayrose, I., & DeMalach, N. (2023). Revising the global biogeography of annual and perennial plants. *Nature*, 624(7990), 109–114.

R Core Team. (2023). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. https://www.R-project.org/.

Raunkiær, C.C. (1934). *The Life Forms of Plants and Statistical Plant Geography*. Oxford University Press.

Revell, L. J. (2009). Size-correction and principal components for interspecific comparative studies. *Evolution*, 63(12), 3258–3268.

Revell, L. J. (2024). phytools 2.0: An updated R ecosystem for phylogenetic comparative methods (and other things). *PeerJ*, 12, e16505.

Riley, S. J., DeGloria, S. D., & Elliot, R. (1999). Index that quantifies topographic heterogeneity. *Intermountain Journal of Sciences*, 5(1–4), 23–27.

Rodríguez-Rodríguez, P., G. Fernández de Castro, A., Seguí, J., Traveset, A., & Sosa, P. A. (2019). Alpine species in dynamic insular ecosystems through time: conservation genetics and niche shift estimates of the endemic and vulnerable Viola cheiranthifolia. *Annals of Botany*, 123(3), 505-519.

Rowe, N., & Speck, T. (2005). Plant growth forms: An ecological and evolutionary perspective. *New Phytologist*, 166(1), 61–72.

Sauquet, H., & Magallón, S. (2018). Key questions and challenges in angiosperm macroevolution. *New Phytologist*, 219(4), 1170–1187.

Smith, S. A., & Brown, J. W. (2018). Constructing a broadly inclusive seed plant phylogeny. *American Journal of Botany*, 105(3), 302-314.

Söllheim, Á. L. L., Suarez, P. M., & Hernández, F. G. (2024). The digital climate atlas of the Canary Islands: A tool to improve knowledge of climate and temperature and precipitation trends in the Atlantic islands. *Climate Services*, 34, 100487.

Stewart, W. N., & Rothwell, G. W. (1993). *Paleobotany and the evolution of plants*. Cambridge University Press.

Taylor, A., Weigelt, P., Denelle, P., Cai, L., & Kreft, H. (2023). The contribution of plant life and growth forms to global gradients of vascular plant diversity. *New Phytologist*, 240(4), 1548-1560.

Triantis, K. A., Economo, E. P., Guilhaumon, F., & Ricklefs, R. E. (2015). Diversity regulation at macro-scales: species richness on oceanic archipelagos. *Global Ecology and Biogeography*, 24(5), 594-605.

van den Hoogen, J., Robmann, N., Routh, D., Lauber, T., van Tiel, N., Danylo, O., & Crowther, T. W. (2021). A geospatial mapping pipeline for ecologists. *BioRxiv*, 2021-07.

van Proosdij, A. S., Sosef, M. S., Wieringa, J. J., & Raes, N. (2016). Minimum required number of specimen records to develop accurate species distribution models. *Ecography*, 39(6), 542–552.

Velazco, S.J.E., Rose, M.B., Andrade, A.F.A., Minoli, I., Franklin, J. (2022). flexsdm: An R package for supporting a comprehensive and flexible species distribution modelling workflow. *Methods in Ecology and Evolution*, 13(8) 1661–1669. https://doi.org/10.1111/2041-210X.13874.

Vitales, D., Garnatje, T., Pellicer, J., Vallès, J., Santos-Guerra, A., & Sanmartín, I. (2014). The explosive radiation of *Cheirolophus* (Asteraceae, Cardueae) in Macaronesia. *BMC Evolutionary Biology*, 14, 1-15.

von Humboldt, A. (1807). *Voyage de Humboldt et Bonpland. Première partie. Physique général, et relation historique du voyage. Premier volume, contenant un essai sur la géographie des plantes, accompagné d'un tableau physique des régions equinoxiales, et servant d'introduction à l'ouvrage*.

Wallace, A. R. (1880). *Island Life: Or, the Phenomena and Causes of Insular Faunas and Floras, Including a Revision and Attempted Solution of the Problem of Geological Climates*. Macmillan.

Wickham H (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. ISBN 978-3-319-24277-4, https://ggplot2.tidyverse.org.

Zanne, A. E., Tank, D. C., Cornwell, W. K., Eastman, J. M., Smith, S. A., FitzJohn, R. G., ... & Beaulieu, J. M. (2014). Three keys to the radiation of angiosperms into freezing environments. *Nature*, 506(7486), 89-92.

Zizka, A., Silvestro, D., Andermann, T., Azevedo, J., Ritter, C. D., Edler, D., Farooq, H., Herdean, A., Ariza, M., Scharn, R., Svanteson, S., Wengstrom, N., Zizka, V., & Antonelli, A. (2019). CoordinateCleaner: Standardized cleaning of occurrence records from biological collection databases. *Methods in Ecology and Evolution*, 10, 7. https://doi.org/10.1111/2041-210X.13152.

Zizka, A., Carvalho, F. A., Calvente, A., Baez-Lizarazo, M. R., Cabral, A., Coelho, J. F. R., Colli-Silva, M., Fantinati, M. R., Fernandes, M. F., Ferreira-Araújo, T., & others. (2020). No one-size-fits-all solution to clean GBIF. *PeerJ*, 8, e9916.

Zizka, A., Onstein, R. E., Rozzi, R., Weigelt, P., Kreft, H., Steinbauer, M. J., Bruelheide, H., & Lens, F. (2022). The evolution of insular woodiness. *Proceedings of the National Academy of Sciences*, 119(37), e2208629119.

Zizka, A., Grenié, M., Conradi, T., Onstein, R. E., Bruelheide, H., & Lens, F. (2025). Global convergence in wood evolution is driven by drought on continents and frost-free temperatures on islands. *bioRxiv*, 2025-09. DOI: https://doi.org/10.1101/2025.09.15.676260.

For the supplementary material associated with this chapter, please see the Zenodo repository accompanying this thesis (DOI 10.5281/zenodo.22093870).



Chapter 4

A Model for Trait-Dependent Colonisation and Diversification on Islands

Ryan F.A. Brewer[1,2*], Ornela Dehayem Nanwou[1,2*], Thijs Janzen[2], Bart Haegeman[3], Luis Valente[1,2,], Rampal S. Etienne[2,**], Frederic Lens[1,4,**]**

[1] Naturalis Biodiversity Center, Darwinweg 2, 2333 CR Leiden, The Netherlands
[2] Groningen Institute for Evolutionary Life Sciences, University of Groningen, Box 11103, 9700 CC Groningen, The Netherlands
[3] CNRS/Sorbonne Universite, Laboratory of Microbial Oceanography, 1 avenue ´ 11 Pierre Fabre, 66650 Banyuls-sur-Mer, France
[4] Institute of Biology Leiden, Leiden University, Sylviusweg 72, Leiden 2333 BE, The Netherlands

* Shared First Authors
** Shared Senior Authors



ABSTRACT

Understanding why some clades harbour more species than others is a fundamental question in evolutionary biology. Islands seem ideally suited for addressing this question because they are host to many clades composed of a single species as well as spectacular radiations. State-dependent speciation and extinction (SSE) models can detect the dependence of diversification on traits but typically require a relatively large phylogeny for a single clade and do not capture community dynamics on islands. In contrast, the island biogeography framework DAISIE (Dynamic Assembly of Island biota through Speciation, Immigration, and Extinction) can capture the dynamics of multiple clades within an island assemblage but does not consider trait evolution. Here, we develop TRAISIE (TRait-dependent Assembly of Island biota through Speciation, Immigration and Extinction), a likelihood-based framework specifically tailored to island communities. TRAISIE combines the DAISIE approach, which models island assemblages as emerging from colonisation, speciation and extinction processes, with the SSE approach, which links diversification dynamics to species traits. This combination enables testing and quantifying how (much) species traits influence the overall diversification dynamics (including, among others, colonisation, speciation and extinction) in an insular setting. We first apply TRAISIE to simulated data to show that it can accurately estimate parameters and perform model selection on sufficiently large datasets under several conditions. We next apply TRAISIE to a large phylogenomic dataset for the angiosperm assemblage of the Canary Islands, which is composed of multiple clades of various sizes with a combination of herbaceous (annual and perennial) and woody (exclusively perennial) species. TRAISIE estimates that woody and perennial species have faster cladogenesis rates and lower extinction rates than herbaceous and annual species, with perennial woody lineages showing the highest cladogenesis rates, annual herbaceous species the lowest, and perennial herbaceous species the lowest extinction rates. Overall, growth form (woody vs. herbaceous) and life span (annual vs. perennial) appear to influence diversification in distinct ways across lineages. Nevertheless, model comparison suggests that a model in which diversification and colonisation rates depend solely on a concealed trait provides a better fit than models based on the observed traits. This implies that growth form and life span alone do not fully explain the observed diversification patterns, and that additional, unmeasured factors could also play an important role. However, goodness-of-fit analyses of the preferred model reveal that it provides a poor fit to the data. Thus, our results regarding trait-dependent diversification in the Canary Islands should be interpreted with caution. We argue that the poor fit may

partly be explained by missing species that may not be optimally accounted for, as well as by the possibility that none of the models tested closely approximates the true generating process. We suggest that future studies should better account for missing species and consider models that integrate both observed and concealed traits. Additionally, trait classifications should be more refined, particularly for woodiness, as grouping all woody species into a single category may mask important biological differences, for example, between insular woody and non-insular woody lineages.

INTRODUCTION

The heterogeneous distribution of extant diversity across the tree of life has been at the forefront of evolutionary biology since its inception (Von Humboldt and Bonpland 1807; Wallace 1855; Darwin 1859). Explaining why some clades harbour more species than others remains today one of the most important endeavours in biodiversity research (Antonelli et al. 2023), with three prevailing hypotheses: differential clade ages (McPeek and Brown 2007; Henao Diaz et al. 2019; Yu and Wiens 2024), differential diversification rates (speciation minus extinction) across various lineages due to inherent lineage-specific traits (Wiens 2017), or differential ecological limits to diversity (constrained by resource limitations or carrying capacity of a system/clade; Wiens 2017; Rabosky and Hurlbert 2015). As traits (including phenotypic as well as genomic traits) affect how species interact with their environment, they can have an effect on speciation and extinction rates. For example, suppose a particular trait state promotes dispersal to new environments of species holding that trait. Such a trait may promote speciation or reduce extinction, resulting in an over-representation of the trait within an area/clade (Anderson et al. 2023).

Islands are perfectly suited to study this topic (Warren et al. 2015; Gillespie 2016; Patino et al. 2017; Bellvert et al. 2025). Their small size, distinct boundaries, and multiple colonisations resulting in clades of varying ages and diversity offer an opportunity to explore the causes of heterogeneity across the tree of life, acting as natural laboratories for patterns observed globally (Losos and Ricklefs 2009). Islands also contribute disproportionately to global biodiversity, with oceanic islands containing ~20% of global terrestrial biodiversity, despite covering only 3.5% of global land area (Whittaker et al. 2017). They are home to remarkable examples of evolutionary radiations, such as the Macaronesian *Aeonium* alliance (Crassulaceae) with 63 species (Mort et al. 2002), Hawaiian honeycreepers (Fringillidae) with 60 species (Lerner et al. 2011; Illera et al. 2024), and the *Anolis* (Dactyloidae) radiations across multiple islands of the West Indies with ~150 species (Losos 2009). Yet, the drivers responsible for why some clades diversify widely whilst others remain depauperate remain poorly understood (Patino et al. 2017; Helmstetter et al. 2023; Yu and Wiens 2024).

Over the past decades, attempts to understand how traits influence diversification have led to the development of statistical models that integrate reconstructed phylogenies and trait data to examine the relationship between trait evolution and clade diversification. Under the SSE (State-dependent Speciation and Extinction) framework, species are modelled as evolving under a birth-death process, where speciation and extinction rates at any given time depend

on the value of a species' trait, which itself can change over time. These rates can be estimated using maximum likelihood, and comparisons across trait states then allow inference of the effects of traits on speciation and extinction. The first of this series of models was introduced by Maddison et al. (2007), which jointly models the evolution of a trait with two categorical states and their effect on the diversification of a taxonomic group (BiSSE-Binary State-dependent Speciation and Extinction), given the complete reconstructed phylogenetic tree of the group of interest and the traits or character states of the species at the tips. The BiSSE model was later extended to 1) account for the often incomplete sampling in empirical phylogenies (FitzJohn et al. 2009), 2) to simultaneously test the effect of multiple traits or a character with multiple states on diversification (MuSSE; FitzJohn 2012), 3) to assess how diversification is impacted by quantitative traits (QuaSSE; FitzJohn 2010), 4) to different geographic environments (GeoSSE; Goldberg et al. 2011), and 5) to test whether change happens at speciation events or along branches (claSSE; Goldberg and Igic 2012).

This first group of SSE models implicitly constrains clade diversification heterogeneity to be explained by the observed traits included in the model, thereby excluding alternative explanations of the variation. Maddison et al. (2007) acknowledged this as a limitation, but argued that incorporating further parameters would lead to low power in detecting a signal of trait dependence. While low power may indeed be an issue (Herrera-Alsina et al. 2019), it was later pointed out that comparison of this first group of models to constant-rate models can lead to false positive inferences, where neutral traits are incorrectly identified as drivers of diversification when the true causal trait is not included in the analysis (Rabosky and Goldberg 2015). Recent advances addressed this issue by allowing diversification rates to depend on observed traits, as well as on traits which, although unobserved, may have influenced species diversification (Beaulieu and O'Meara 2016; Herrera-Alsina et al. 2019). The inclusion of these so-called concealed traits allows the framework to distinguish diversification that depends on observed (also called examined) traits from patterns generated by unmeasured (called hidden or concealed) variables.

Standard SSE approaches are designed for single clades (e.g. a single large phylogeny of a group of interest) and are not well-suited for modelling island community dynamics, which often comprise multiple independent clades arising from separate colonisation events from the mainland. Many of the clades on islands are simply too small (i.e. just a few species) to be analysed using existing SSE approaches. By requiring a single tree, classic SSE approaches are not a good match for analysing island assemblages composed of various clades often from different taxonomic groups (e.g. various plant families). Moreover, colonisation represents a key evolutionary process shap-

ing the composition and structure of island communities; therefore, colonisation and its potential trait dependence need to be explicitly incorporated when modelling biodiversity dynamics on islands (MacArthur and Wilson 1967; Valente et al. 2015).

The DAISIE (Dynamic Assembly of Islands through Speciation, Immigration, and Extinction) model was proposed as a phylogenetic framework specifically tailored for island communities, designed to assess the relative roles of colonisation, speciation, and extinction in shaping island community assemblage (Valente et al. 2015; Dehayem et al. 2025; Etienne et al. 2023). However, DAISIE does not account for trait evolution, as its likelihood computation assumes uniform colonisation, speciation, and extinction parameters across species. Here, we develop TRAISIE (TRait-dependent Assembly of Island biota through Speciation, Immigration and Extinction), a likelihood-based framework that extends the DAISIE approach by incorporating trait dependence in colonisation and diversification rates, while allowing traits to evolve over time. Therefore, TRAISIE can be seen as a merger between DAISIE and SSE models. This extension enables the exploration of how one or multiple traits (examined and concealed) influence colonisation, speciation, and/or extinction in island communities. To assess the accuracy of model inference when estimating parameters, we first apply TRAISIE to simulated island communities of various sizes, generated under various evolutionary scenarios: the constant-rate (CR) model, where there is no trait-dependence, the examined-trait dependent diversification (ETD), and the concealed-trait-dependent diversification (CTD) model. We simulate 100 datasets of three different sizes (small, medium, and large under CR (with trait transitions), under ETD, and under CTD). We assess i) the accuracy and precision of parameter estimates when fitting TRAISIE to these datasets, ii) the accuracy of model selection, and iii) the effect of the size of the dataset on the accuracy of these inferences.

As a case study, we apply TRAISIE to a recently developed phylogenomic dataset for the angiosperm community of the Canary Islands (Brewer et al. 2025). Angiosperms have often been the focus of diversification studies (Helmstetter et al. 2023), and many of the fastest diversifying clades globally are angiosperms (Valente et al. 2010; Lagomarsino et al. 2016). In almost 15 years of SSE model research in angiosperms, over 30 traits have been identified as potentially having affected diversification in clades across the angiosperm tree of life, spanning over 80 families (Helmstetter et al. 2023). We use TRAISIE to address several evolutionary hypotheses concerning two traits previously suggested to affect diversification in angiosperms: growth form (herbaceousness vs woodiness) and life span (annual vs perennial). First, we re-examine the results of Brewer et al. (2025), who used DAISIE to test whether clades

assigned to a pre-selected "type" (e.g. herbaceous or woody) differ in diversification rates. They found that woody clades and perennial clades have faster diversification rates than herbaceous clades and annual clades, respectively, assuming that each clade represents a single trait state. This assumption was a simplification because clades can (and do) contain species with different trait states. Furthermore, because DAISIE is silent on traits, their approach did not account for transitions between trait states. We therefore evaluated whether these shortcomings affect their inferences regarding diversification based on the two examined traits. Second, to disentangle possible relationships being conflated within two-state models (herbaceous vs woody, annual vs perennial), we also test a three-state model (annual herbaceous, perennial herbaceous, perennial woody), to further disentangle how growth form and life span affect diversification within the Canary Islands. Finally, we assess how well the best fitting model reproduces the empirical Canary Island dataset by simulating 200 datasets with the parameter estimates for each best fitting model, and comparing how several summary statistics overlap with the observed values.

METHODOLOGY

Model

We develop a new trait-dependent model called TRAISIE (TRait-dependent Assembly of Island biota through Speciation, Immigration and Extinction). The model combines the approach of DAISIE, which models island biodiversity as emerging from colonisation, speciation, and extinction processes, with the SSE framework, which links diversification dynamics to species traits. The model assumes a mainland species pool, in which species are categorised according to their trait states. Consistent with the original DAISIE model, the mainland itself is considered static, i.e. the number of species and their traits do not change over time. While this assumption is clearly violated in reality, it has several technical methodological advantages for parameter estimation, and parameter estimation in DAISIE has been shown to be robust to this assumption (Lambert et al. 2022). After the emergence of the (oceanic) island, a species with trait i can immigrate from the mainland to what is initially an empty environment, at rate γi. Once established, immigrants may go extinct (at rate μi), while others may persist. The surviving lineages can follow several trajectories: (1) They remain as nonendemic species (island and mainland populations belong to the same species). (2) They give rise to a single endemic species that is different from the ancestral mainland population that colonised the island (through anagenetic speciation, at rate of anagenesis λ^a_i, or cladogenetic speciation, at rate of

cladogenesis λ^c_i followed by extinction; or they may also undergo a trait change from trait state *i* to trait state *j*, at rate q_{ij}, which may trigger anagenesis with probability p*ij* in our model). Naturally, endemic species, whether belonging to singleton lineages (i.e., lineages including a single extant species) or to a clade with at least two species, continue to suffer a risk of extinction. (3) They diversify into multiple endemic species through cladogenesis, at the rate λ^c_i. Therefore, the island community is represented by one or more clades (including non-endemic lineages composed of a single non-endemic species, and endemic lineages (singletons or clades with more than one endemic species)), each corresponding to a separate colonisation event, represented by its own phylogenetic tree. Figure 1 provides an example of a dataset for an island community with information on the trait or character states of the extant species. Each clade represents a separate colonisation event from the mainland that has led to one or more species on the island for which we have information on the trait states. The dating estimates of the colonisation events (marked by circles/triangles) are inferred from the divergence between island species and their mainland closest related species or lineage (or closest related mainland population, in case of a non-endemic species) inferred from molecular phylogenies. For cases where the colonisation times are unknown (e.g. the molecular phylogeny is not available for the given lineage), the model integrates over all possible colonisation times, that is, between the island age and the present for singleton lineages, and between the island age and the crown age for an endemic clade with more than one species.

It is possible that the same mainland species has already colonised the island several times before an observed colonisation event (Caujape-Castells et al. 2011). In this case, the model always assumes that the observed lineage descends from the latest colonisation event. This means that at the observed colonisation event, any trace of these previous colonisations is erased in the TRAISIE/DAISIE framework. Recolonisation of the same mainland ancestors can occur and still remain on the island, but only after the previous colonisation has resulted in speciation. It is important to note that for each trait of interest in TRAISIE, trait data for each species is required for the island community of interest, as well as the distribution of trait values across the mainland pool. For how the distribution of trait values across the mainland pool is used in TRAISIE, see section 'Canary Islands Angiosperm trait & phylogenetic data' below. Trait data at the root of the phylogeny representing each clade of the island community can also be included as input for TRAISIE, but is not mandatory. For endemics, this trait data can be inferred from the closest extant continental relative; for non-endemics, in most cases, the trait state of the island species can be used.

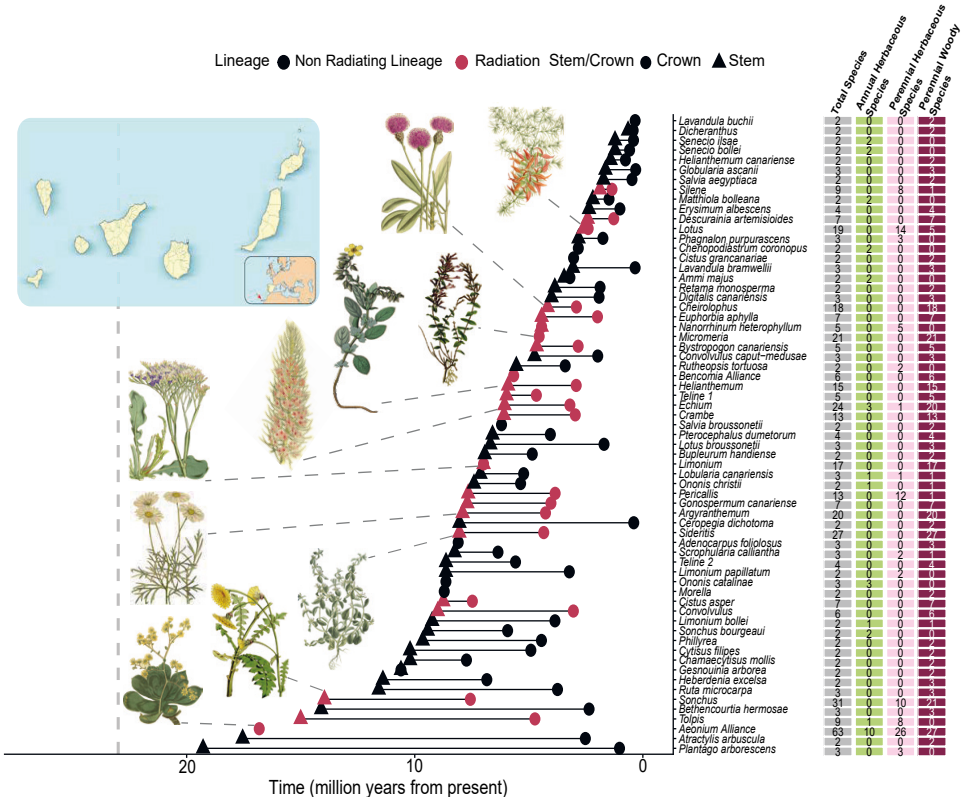


Figure 1: Stem and crown ages of endemic Canary Islands non-monocot angiosperms, for which divergence time estimates are available, adapted from Brewer et al. (2025). Non-endemic lineages are not included here, but their age estimates are available in Brewer et al. (2025). The colonisation of the Canary Islands by each lineage is inferred to have taken place at any time between the stem and the crown age. Radiations refer to monophyletic groups with five or more species in the Canary Islands, with the total number of species and number in each trait state given in columns to the right. Lineages for which only the crown age estimate is shown (e.g., *Aeonium* Alliance (Crassulaceae), *Micromeria* (Lamiaceae), *Bencomia* Alliance (Rosaceae)) did not have a mainland relative successfully sequenced, and therefore, no stem age is available. The inset map shows the Canary Islands and their location west of Western Sahara. The vertical dashed line represents the age of the oldest extant rock in the Canary Islands (23 My). The images represent species from the ten most diverse Canarian non-monocot lineages (from bottom left to top right): *Aeonium glutinosum* (Crassulaceae), *Sonchus gummifer* (Asteraceae), *Sideritis hirsuta* (Lamiaceae), *Argyranthemum foeniculaceum* (Asteraceae), *Limonium bourgeaui* (Plumbaginaeae), *Echium wildpretii* (Boraginaceae), *Helianthemum teneriffae* (Cistaceae), *Micromeria marginata* (Lamiaceae), *Cheirolophus crassifoliius* (Asteraceae), and *Lotus berthelotii* (Fabaceae). Sources of the illustrations are given in Supplementary Data S1.

Likelihood Computation

We assume that the evolutionary dynamics of each clade are independent from one another, allowing the dynamics of each colonising species to be modelled separately. The likelihood of observing the entire community is then calculated

as the product of the probabilities of each independent colonisation plus its potential descendants (which we will refer to as 'clade' hereafter, even if it is a non-endemic or endemic singleton) present on the island, and integrating over all potential colonisation events that occurred before and after the observed colonisation, and subsequently went extinct. The model distinguishes four possible outcomes following a colonisation event, three of which correspond to extant clades observed on the island:

- No surviving lineage: A colonisation event that resulted in a lineage that has gone completely extinct and is not represented among the present-day species on the island.
- Non-endemic singleton: A species present on the island has not undergone speciation since the arrival of the mainland population, while populations of the same species remain on the mainland. The trait state may have changed, but this change in trait state did not lead to a different species (which has probability 1– p*ij* for a change in trait from *i* to *j*).
- Endemic singleton: A species present on the island that underwent anagenetic or cladogenetic speciation after colonisation from a nearby continent, but only one extant species survives on the island, or the island population evolved a trait state change that led to a different island species (with probability p*ij* for a change from trait *i* to *j*).
- Endemic radiation: A clade of two or more species that all descend from a single colonisation event and underwent cladogenetic speciation on the island, resulting in a group of endemic species. The species may have different trait states, depending on whether trait change occurs.

To compute the probabilities of observing a clade, we follow the differential equation-based approach introduced by Maddison et al. (2007) and Dehayem et al. (2025), where probabilities are assigned at the tips of the lineage, and a set of rules is used to iteratively update these values as we trace the lineage backwards in time, from the tips to the colonisation time (stem age of the lineage). Note that this is different from the approach used in the original DAISIE (Valente et al. 2015), which tracks lineages forwards in time, starting at the island age. At the colonisation time of a lineage, we incorporate the probability of colonisation before continuing the computation further back in time to the island's origin. We define the backward time τ, measured from the present to the past. It is related to the usual forward time t by the equation $\tau = \tau p - t + tp$, where τp is the backward time at the present.

In order to capture the state of the system at specific points in time, we introduce E,*i*(τ), and several state-dependent variants of the probability D*(τ)*. These states represent the probabilities of the dynamics being consistent with the observed data. We define these states as follows:

- *E, i(τ)*, the probability that a lineage starting at time τ as an endemic species leaves no descendants at the present;
- $D_{\mathrm{E},i}(\tau)$ is the probability that a lineage present on the island at time τ as an endemic lineage (hence the subscript E) with trait state i evolves into the observed clade;
- $D_{\mathrm{M},i}(\tau)$ is the probability that all lineages descending from the mainland species (hence the subscript M) with trait state *i* present at time τ, as well as those originating from subsequent colonisations, evolve in a manner consistent with the observed data;
- $D_{\mathrm{A},i\mathrm{M}}(\tau)$ is the probability that the species is absent (hence the subscript A) from the island at time τ, and that all lineages resulting from subsequent colonisations evolve in a manner consistent with the observed data

Figure 2 provides a schematic representation of the state probabilities DM(τ) and DA(τ) defined in the model. The times shown in black (τ0, τc, τ2, and τp) are extracted from the data. In contrast, the time point τE shown in blue is unknown and our model integrates over all its possible values. These times are described as follows:

- τ0: time of the island's origin, i.e. the island's age
- τC: time of colonisation event that gives rise to the lineage observed at the present time
- τE: time at which the endemic lineage at the present time becomes endemic
- τp: present time

We now define:

$$D^{(1)}_{\mathrm{A},i_{\mathrm{M}}}(\tau) \quad \text{for } \tau \in [\tau_c, \tau_0] \tag{1}$$

$$D^{(1)}_{\mathrm{M},i}(\tau) \quad \text{for } \tau \in [\tau_c, \tau_0] \tag{2}$$

$$D^{(2)}_{\mathrm{M},i}(\tau) \quad \text{for } t \in [\tau_p, \tau_c] \tag{3}$$

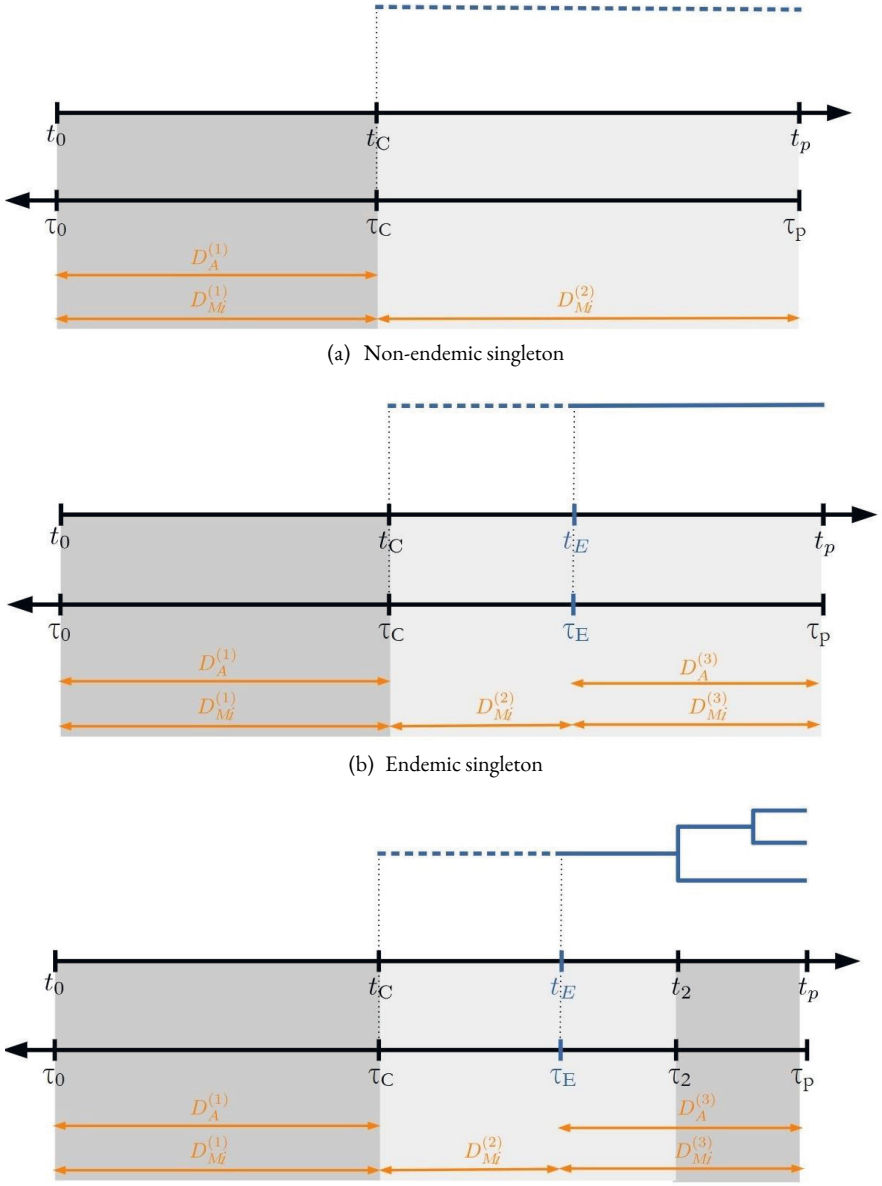


Figure 2: Schematic representations of the temporal structure and definition of the state probabilities $D_{M,i}(\tau)$ and $D_{A,iM}(\tau)$ in the backward pruning algorithm for different cases. t and τ represent the forward and the backward time, respectively. The dashed blue line represents the time interval during which the species is nonendemic, while the solid blue line represents the interval during which the species is endemic. The orange arrows indicate the time intervals over which each probability is computed, going backwards in time. The light and darker grey regions correspond to distinct time intervals, each governed by a different set of differential equations, reflecting changes in model dynamics across the lineage history. See text for definition of the variables plotted.

$$D^{(3)}_{\mathrm{A},i_{\mathrm{M}}}(\tau) \quad \text{for } \tau \in [\tau_p, \tau_E] \tag{4}$$

$$D^{(3)}_{\mathrm{M},i}(\tau) \quad \text{for } t \in [\tau_p, \tau_E] \tag{5}$$

These variables satisfy different equations depending on which part of the tree is considered. We use the superscripts to distinguish the same variable occurring in different systems of equations.

- $D^{(1)}_{\mathrm{M},i}(\tau)$, the probability that the dynamics give a clade consistent with the data, when at time τ we have an island with the non-endemic species with a trait state *i* and which is not represented in the clade. This species originates from a colonisation event between τ0 and τc. The species goes extinct before the present, or a recolonisation occurs by the observed mainland species, erasing the former species' signature on the data;

- $D^{(2)}_{\mathrm{M},i}(\tau)$, the probability that the dynamics give a clade consistent with the data, when at time τ we have an island with the non-endemic species with trait state *i* and which is represented in the clade. This species originates from a colonisation at τc;

- $D^{(3)}_{\mathrm{M},i}(\tau)$, the probability that the dynamics give a clade consistent with the data, when at time τ we have an island with the non-endemic species with trait *i* and which is not represented in the clade. This species originates from a colonisation after the speciation event on the observed lineage or the trait change that led to a different species at τE. The species must become extinct before the present.

Below, we describe how to compute the likelihood for different types of lineages with precisely known colonisation time τ_c. The equations can be readily extended to allow for a colonisation time that is not precisely known but lies between τ_{max} and τ_{min}. These equations are given in the supplementary material.

We assume that each trait state consists of an observed trait state and an unobserved trait state, i.e. a state related to unobserved traits, also called concealed traits. This unobserved trait may have a different impact on diversification dynamics, colonisation and transition rates than the observed trait. Each type of trait may have multiple discrete states. Hence trait *i* is indexed by two subscripts, as in Beaulieu and O'Meara (2016):

- o is the index of the observed states, where $o \in \{0,1,...,n\}$,
- c is the index of the concealed states, where $c \in \{A,B,...,\beta\}$.

Thus, the rates of the model are presented as $\gamma_i = \gamma_{oc}$, $\lambda^c_i = \lambda^c_{oc}$, $\lambda^a_i = \lambda^a_{oc}$, $\mu_i = \mu_{oc}$, and $qij = qo_ic_io_jc_j$.

Probability of Observed Colonisations and Potential Radiations

No Surviving Descendants

Here, we consider the scenario that a mainland species with trait state i_M has not colonised the island, or it has, but after establishment and potential speciation or trait change, the species or any of its descendants did not survive to the present time τ_p. The dynamical equations describing the change in the probabilities between the present time τ_p and the island age τ_0 are given by:

$$\frac{dD^{(1)}_{A,i_M}}{d\tau} = -\gamma_M D^{(1)}_{A,i_M}(\tau) \quad \text{(no recolonisation)}$$
$$+\gamma_M D^{(1)}_{M,i}(\tau) \quad \text{(recolonisation of the mainland species)} \qquad (6a)$$

$$\frac{dD^{(1)}_{M,i}}{d\tau} = -\left(\lambda^a_i + \lambda^c_i + \mu_i + \sum_{j\neq i} q_{ij}\right) D^{(1)}_{M,i}(\tau) \quad \text{(no anagenesis, no cladogenesis, no extinction of the mainland species)}$$
$$+\left(\lambda^a_i E_i(\tau) + \lambda^c_i E^2_i(\tau) + \mu_i + \sum_{j\neq i} q_{ij} p_{ij} E_j(\tau)\right) D^{(1)}_{A,i_M}(\tau) \quad \text{(anagenesis/cladogenesis/extinction/trait change followed by anagenesis)}$$
$$+\sum_{j\neq i} q_{ij}(1-p_{ij}) D^{(1)}_{M,j}(\tau) \quad \text{(trait change not followed by anagenesis)} \qquad (6b)$$

$$\frac{dE_i}{d\tau} = \mu_i \quad \text{(extinction of the lineage)}$$
$$-\left(\lambda^c_i + \mu_i + \sum_{j\neq i} q_{ij}\right) E_i(\tau) \quad \text{(no cladogenesis, no extinction, no trait change)}$$
$$+\lambda^c_i E^2_i(\tau) \quad \text{(cladogenesis followed by extinction of both daughter lineages)}$$
$$+\sum_{j\neq i} q_{ij} E_j(\tau) \quad \text{(trait change followed by extinction)} \qquad (6c)$$

where the time τ is measured backwards from the present time τ_p down to the Island age τ_0, and M the trait state of the mainland ancestor. The term $\sum_{j\neq i} q_{ij}$ represents the sum of all possible transition rates from state i to any other state j, with $j \neq i$.
At the present time τ_p, the initial conditions are given by:

$$D^{(1)}_{\mathrm{A},i_\mathrm{M}}(\tau_p) = 1, \quad D^{(1)}_{\mathrm{M},i}(\tau_p) = 0, \quad E_i(\tau_p) = 0. \tag{7}$$

We compute the contribution of this scenario to the likelihood as:

$$\mathcal{L}_0 = D^{(1)}_{\mathrm{A},i_\mathrm{M}}(\tau_0), \tag{8}$$

Singleton Lineage, Endemic or Non-Endemic, with Colonisation at τ_C
Here we consider the scenario that a mainland species colonises the island at time τ_C, and, in the case of an endemic singleton lineage, speciated at an unknown speciation τ_E time between τ_C and τ_p; between τ_C and τ_E, the species is present on the island and is non-endemic; its dynamics are captured by the variable $D_{M,i}$. After speciation and before τ_p, the mainland ancestor may recolonise the island and go extinct again; this is captured by the variable $D_{M,i}$. The dynamical equations describing the change in the probabilities between τ_p and τ_C are given by:

$$\begin{aligned}
\frac{dD_{\mathrm{E},i}}{d\tau} = &-\left(\lambda^c_i + \mu_i + \sum_{j\neq i} q_{ij}\right) D_{\mathrm{E},i}(\tau) && \text{(no cladogenesis, no extinction, no trait change)}\\
&+ 2\lambda^c_i D_{\mathrm{E},i}(\tau) E_i(\tau) && \text{(cladogenesis followed by extinction of one daughter lineage)}\\
&+ \sum_{j\neq i} q_{ij} D_{\mathrm{E},j}(\tau) && \text{(trait change)} \qquad (9a)\\
\frac{dD^{(2)}_{\mathrm{M},i}}{d\tau} = &-\left(\lambda^a_i + \lambda^c_i + \mu_i + \gamma_\mathrm{M} + \sum_{j\neq i} q_{ij}\right) D^{(2)}_{\mathrm{M},i}(\tau) && \text{(no anagenesis, no cladogenesis, no extinction, no recolonisation, no trait change)}\\
&+ \left(\lambda^a_i D_{\mathrm{E},i}(\tau) + 2\lambda^c_i D_{\mathrm{E},i}(\tau) E_i(\tau) + \sum_{j\neq i} q_{ij} p_{ij} D_{\mathrm{E},j}(\tau)\right) D^{(3)}_{\mathrm{A},i_\mathrm{M}}(\tau) && \text{(anagenesis/cladogenesis followed by extinction of one daughter lineage/trait change followed by anagenesis)}\\
&+ \sum_{j\neq i} q_{ij}(1-p_{ij}) D^{(2)}_{\mathrm{M},j}(\tau) && \text{(trait change not followed by anagenesis)}\\
& && (9b)
\end{aligned}$$

$$\frac{dD^{(3)}_{\mathrm{M},i}}{d\tau} = -\left(\lambda_i^a + \lambda_i^c + \mu_i + \sum_{j\neq i} q_{ij}\right) D^{(3)}_{\mathrm{M},i}(\tau) \quad \text{(no anagenesis, no cladogenesis, no extinction, no trait change)}$$
$$+\left(\lambda_i^a E_i(\tau) + \lambda_i^c E_i^2(\tau) + \mu_i + \sum_{j\neq i} q_{ij} p_{ij} E_j(\tau)\right) D^{(3)}_{\mathrm{A},i_\mathrm{M}}(\tau) \quad \text{(anagenesis/cladogenesis/extinction/trait change followed by anagenesis)}$$
$$+\sum_{j\neq i} q_{ij}(1-p_{ij}) D^{(3)}_{\mathrm{M},j}(\tau) \quad \text{(trait change)} \tag{9c}$$

$$\frac{dD^{(3)}_{\mathrm{A},i_\mathrm{M}}}{d\tau} = -\gamma_\mathrm{M} D^{(3)}_{\mathrm{A},i_\mathrm{M}}(\tau) \quad \text{(no recolonisation)}$$
$$+\gamma_\mathrm{M} D^{(3)}_{\mathrm{M},i}(\tau) \quad \text{(recolonisation of the mainland species)} \tag{9d}$$

The dynamics of E_i are governed by equation (6c).
For an endemic lineage, the initial conditions at the present time τ_p are given by:

$$D_{\mathrm{E},i}(\tau_p) = 1, \quad D^{(2)}_{\mathrm{M},i}(\tau_p) = 0, \quad D^{(3)}_{\mathrm{M},i}(\tau_p) = 0, \quad D^{(3)}_{\mathrm{A},i_\mathrm{M}}(\tau_p) = 1, \quad E_i(\tau_p) = 0 \tag{10}$$

while for a non-endemic lineage, the initial conditions are:

$$D_{\mathrm{E},i}(\tau_p) = 0, \quad D^{(2)}_{\mathrm{M},i}(\tau_p) = 1, \quad D^{(3)}_{\mathrm{M},i}(\tau_p) = 0, \quad D^{(3)}_{\mathrm{A},i_\mathrm{M}}(\tau_p) = 0, \quad E_i(\tau_p) = 0 \tag{11}$$

In the interval $[\tau_c,\tau_0]$ there is no clade that leaves descendants at the present. Still, the mainland species can be present on the island, a possibility we describe with the variable $D_{M,i}$. The dynamical equations describing the change in probabilities in this time interval are given by Eqs. (6a-6c), with initial conditions at τ_c given by:

$$D^{(1)}_{\mathrm{M}i}(\tau_C^+) = D^{(1)}_{\mathrm{A},i_M}(\tau)(\tau_C^+) = \gamma_\mathrm{M} D^{(2)}_{\mathrm{M},i}(\tau)(\tau_C^-),$$
$$E_i(\tau_C^+) = E_i(\tau)(\tau_C^-) \tag{12}$$

The contribution to the likelihood of this scenario is given by:

$$\mathcal{L}_S = D^{(1)}_{\mathrm{A},i_\mathrm{M}}(\tau_0), \tag{13}$$

Endemic Clade, with Colonisation at τ_c
Here, we consider the scenario that the mainland species colonises the island at time τ_c and diversifies into several extant descendant species at the present time τ_p. After the first speciation event of the lineage, the mainland ancestor

may recolonise the island and go extinct again; this is captured by the variables $D^{(3)}_{\mathrm{M},i}$ and $D^{(3)}_{\mathrm{A},i_\mathrm{M}}$. Between the crown age τ_2 and the present τ_p, the lineage is present, and has already undergone cladogenetic speciation. The dynamical equations describing the change in the probabilities in this time interval are given by:

$$\begin{aligned}
\frac{dD_{\mathrm{E},i}}{d\tau} = &-\left(\lambda_i^c + \mu_i + \sum_{j\neq i} q_{ij}\right) D_{\mathrm{E},i}(\tau) && \text{(no cladogenesis, no extinction, no trait change)}\\
&+ 2\lambda_i^c D_{\mathrm{E},i}(\tau) E_i(\tau) && \text{(cladogenesis followed by extinction of one daughter lineage)}\\
&+ \sum_{j\neq i} q_{ij} D_{Ej}(\tau) && \text{(trait change)} \qquad (14a)\\
\frac{dD^{(3)}_{\mathrm{M},i}}{d\tau} = &-\left(\lambda_i^a + \lambda_i^c + \mu_i + \sum_{j\neq i} q_{ij}\right) D^{(3)}_{\mathrm{M},i}(\tau) && \text{(no anagenesis, no cladogenesis, no extinction, no trait change)}\\
&+ \left(\lambda_i^a E_i(\tau) + \lambda_i^c E_i^2(\tau) + \mu_i + p\sum_{j\neq i} q_{ij} E_j(\tau)\right) D^{(3)}_{\mathrm{A},i_\mathrm{M}}(\tau) && \text{(anagenesis/cladogenesis/extinction/trait change followed anagenesis)}\\
&+ (1-p_{ij}) \sum_{j\neq i} q_{ij} D^{(3)}_{\mathrm{M},j}(\tau) && \text{(trait change)} \qquad (14b)\\
\frac{dD^{(3)}_{\mathrm{A},i_\mathrm{M}}}{d\tau} = &-\gamma_\mathrm{M} D^{(3)}_{\mathrm{A},i_\mathrm{M}}(\tau) && \text{(no recolonisation)}\\
&+ \gamma_\mathrm{M} D^{(3)}_{\mathrm{M},i}(\tau) && \text{(recolonisation of the mainland species)} \qquad (14c)
\end{aligned}$$

The dynamics of E_i are governed by equation (6c).

The phylogenetic tree for a clade does not need to be complete. Assuming that species in the lineages are sampled at random (i.e. that all species in the clade have the same probability of being sampled or not), incomplete sampling can be allowed by setting the initial conditions for each tip in the tree to reflect ρ_i, the probability of sampling a species with a trait i (following FitzJohn et al. (2009)), as follows:

$$D_{\mathrm{E},i}(\tau_p) = \rho_i, \quad D^{(3)}_{\mathrm{M},i}(\tau_p) = 0, \quad D^{(3)}_{\mathrm{A},i_\mathrm{M}}(\tau_p) = 1, \quad E_i(\tau_p) = 1 - \rho_i \qquad (15)$$

At each branching time τ_k, we have to combine the probabilities $D_{E,i}$ of the two sister branches before continuing our way back to the crown of the lineage. Denoting these probabilities by $D_{E,il}$ and $D_{E,ir}$ we have:

$$D_{\mathrm{E},i}(\tau_k^+) = \lambda_i^c D_{\mathrm{E},il}(\tau_k^-) D_{\mathrm{E},ir}(\tau_k^-) \qquad (16)$$

In the interval $[\tau_c,\tau_2]$, the species has already colonised and may or may not have speciated. As initial conditions for $D_{\mathrm{E},i}(\tau_2^+)$, we take $\lambda_i^c d\tau$ times the product of the $D_{\mathrm{E},i}$ functions of the subclades. For $D^{(2)}_{\mathrm{M},i}(\tau_2^+)$, we use the same quantity multiplied by the probability $D^{(3)}_{\mathrm{A},i_{\mathrm{M}}}(\tau_2^-)$ that the mainland species is not present on the island. This covers both cases in which the speciation event that transforms the mainland species into an endemic one occurs either between τ_{max} and τ_2 or precisely at time τ_2. The initial conditions of the set of equations at τ^+_2 are then given by:

$$\begin{aligned}
D_{\mathrm{E},i}(\tau_2^+) &= \lambda_i^c D_{\mathrm{E},il}(\tau_2^-) D_{Eir}(\tau_2^-),\\
D^{(2)}_{\mathrm{M},i}(\tau_2^+) &= \lambda_i^c D_{\mathrm{E},il}(\tau_2^-) D_{\mathrm{E},ir}(\tau_2^-) D^{(3)}_{\mathrm{A},i_{\mathrm{M}}}(\tau_2^-),\\
D^{(3)}_{\mathrm{M},i}(\tau_2^+) &= D^{(3)}_{\mathrm{M},i}(\tau_2^-),\\
)^{(3)}_{\mathrm{A},i_{\mathrm{M}}}(\tau_2^+) &= D^{(3)}_{\mathrm{A},i_{\mathrm{M}}}(\tau_2^-),\\
E_i(\tau_2^+) &= E_i(\tau_2^-).
\end{aligned} \tag{17}$$

The dynamical equations describing the change in the probabilities in this time interval are given by Eqs. (9).

In the interval $[\tau_c,\tau_0]$ there is no clade, and so the mainland species is not represented. Still, it can be present on the island, a possibility we describe with the variable $D^{(1)}_{\mathrm{M},i}$. As for the previous case, the dynamical equations describing the change in the probabilities in this time interval are given by Eqs. (6a-6c), with initial conditions at τ_c given by:

$$\begin{aligned}
D^{(1)}_{\mathrm{M},i}(\tau_c^+) = D^{(1)}_{\mathrm{A},i_{\mathrm{M}}}(\tau_c^+) = \gamma_{\mathrm{M}} D^{(2)}_{\mathrm{M},i},\\
E_i(\tau_c^+) = E_i(\tau_c^-)
\end{aligned} \tag{18}$$

The contribution to the likelihood of this scenario is given by:

$$\mathcal{L}_{EC} = D^{(1)}_{\mathrm{A},i_{\mathrm{M}}}(\tau_0) \tag{19}$$

At the root, the trait of the ancestor species may be unknown. In these cases, we calculate the likelihood given each possible states, and the overall likelihood is calculated as the weighted average over all the possible root states.

$$\mathcal{L} = \frac{\sum_i \mathcal{L}_i \mathrm{w}_i}{\sum_i \mathrm{w}_i} \tag{20}$$

where L_i represents the likelihood of a lineage conditional on the root state being in state i, and w_i denotes the weight associated with that root state.

We used root-state information when available, assigning weight 1 to the known root state and 0 to the other observed states, and giving equal weights

to concealed states. When root states were unavailable, weights were assigned according to the relative mainland pool sizes, so that states with larger mainland pools received greater weight.

We implemented the model, likelihood estimation framework and simulation code in the R package TRAISIE available on GitHub at https://github.com/thijsjanzen/TRAISIE.

Model Fitting and Performance Analysis

To assess the performance of the TRAISIE likelihood inference method, we used a parametric bootstrap approach. We simulated 200 datasets under three different models and three scenarios (see below), and we repeated the simulations with three different mainland pool sizes to generate datasets of three different sizes (see Table 1 below; Tables S1, S2). Our aim was to assess (i) the accuracy and precision of parameter estimates, (ii) the ability of the method to accurately identify the correct generating model, and (iii) the size of the dataset (in terms of number of species) required for reliable inference (see Table S2 for the set of parameters used for the simulations). The phylogenies of the island communities and trait data were simulated using a simulation function available in the R package DAISIE.

We began simulating datasets with a simple two-state model in which colonisation, extinction, and speciation (cladogenesis and anagenesis) rates (CES rates) do not depend on their trait states and transitions between the two states are allowed (CR ("constant rates") model, six parameters; Table S1). Next, we considered an ETD ("examined trait-dependent") model, also with two observed trait states, in which CES rates were allowed to differ between species with different observed trait states (10 parameters; Table S1). Finally, we considered a CTD ("concealed trait-dependent") model with two observed and two hidden trait states, in which CES rates were allowed to differ only between species with different concealed states (12 parameters; Table S1). The simulated datasets include small datasets with (mean 39–45 lineages, 84–99 species), medium (mean 105–118 lineages, 222–257 species), and large (mean 540–618 lineages, 1146–1349 species) datasets; (see Tables S1 & S2 in the Supplementary Material for details), to assess the impact of dataset size on the accuracy and precision of the parameter estimates and model selection performance. To explore how well the model performs in selecting the true generating model, we fitted TRAISIE under the three models (CR, ETD, and CTD) for 100 simulated datasets described in Table S1. These datasets were generated using the parameter values given in Table S2, which were based on early versions of the parameter estimates inferred from the empirical data (see below, Table S2). For each dataset, we compared models and counted the instances

where the generating model was selected as the best model according to ML, AIC and BIC.

Canary Island Angiosperm Trait & Phylogenetic Data

After testing the method on simulated datasets, we applied TRAISIE to empirical data of the Canary Island angiosperms to examine how growth form (herbaceous vs. woody) and lifespan (annual vs. perennial) shape colonisation and diversification dynamics in the Canary Island angiosperms community. We focus on the non-monocot community of the Canary Islands, following the main approach of Brewer et al. (2025), because a key focus of that study was to study the effect of woodiness on diversification, and the woodiness trait cannot evolve in monocot species. The Canary Islands are home to 1,385 native species of angiosperms, of which 225 are monocots. Therefore, the total diversity of non-monocots is 1,160 species, which resulted from 676 colonisation events, 179 of which are endemic and 497 are non-endemic.

To obtain trait data for the Canary Islands non-monocot community, we extended the trait database of Brewer et al. (2025) to cover all non-monocot species by searching the available literature using Google Scholar and published Canary Island floras (Muer et al. 2016; Kattge et al. 2020; Weigelt et al. 2020; Dıaz et al. 2022; Hanz et al. 2022; Barajas Barbosa et al. 2023; Denelle et al. 2023; Sauerbier et al. 2023). For any missing species, a literature search with the terms "[species] + "herbaceous"/"woody" + "annual"/"perennial" was conducted. Three states were scored: annual herbaceous, perennial herbaceous, and perennial woody. In the very few cases where a species had populations of both annual and perennial herbaceous species, the species was scored as perennial, and biennial species were reclassified as perennial, to give an annual and a 'non-annual' herbaceous classification. The data were reclassified to a two-state or three-state state (annual/perennial; herbaceous/woody; annual herbaceous/perennial herbaceous/perennial woody). Trait data coverage across species for the final dataset was 85%. For specific trait information about each species and data coverage for each trait, see Supplementary Data S2. To identify the root state (the trait state of the closest continental relative of the Canarian species), we selected the traits of the closest extant continental ancestors of each of the 676 Canarian clades identified in Brewer et al. (2025). To briefly explain their methods, a literature search was conducted to identify the closest extant continental ancestor for each native Canarian angiosperm species using the phylogenetics and taxonomic literature. To identify the trait state for each of the ancestral species identified in Brewer et al. (2025), we repeated our literature search with the terms "[species] + "herbaceous"/"woody" + "annual"/"perennial". For the species used to repre-

sent the ancestor of each of the Canarian clades, its trait data, and the source, see Supplementary Data S3. To identify the trait distribution across the mainland pool, we generated a species list for the mainland pool of Brewer et al. (2025), i.e., native non-monocots of the Mediterranean Basin, which comprises colonisations of 74 of the 83 natively colonising Canarian angiosperm families. We then checked this species list against four sources and recorded the trait state: The Global Inventory of Floras and Traits, GIFT (Denelle et al. 2023); the TRY Initiative (Kattge et al. 2020); Poppenwimer et al. (2023); and Brown et al. (2023).

We used the time-calibrated phylogenetic tree of Canary Island non-monocot angiosperms from Brewer et al. (2025), which is based on the Angiosperms353 nuclear gene bait set (Johnson et al. 2019). This dataset comprises ~ 58% of the total number of Canary Island non-monocot angiosperms (669 of ~ 1160 species) and 435 of the 676 estimated non-monocot clades (~ 65%), as well as 771 continental relatives. Brewer et al. (2025) reconstructed the phylogeny using ASTRAL-IV v1.18.4.5 (Zhang and Mirarab 2022), with branch lengths estimated by constraining the ASTRAL-IV topology in RAxML-NG v1.2.1 (Kozlov et al. 2019), using a concatenated supermatrix of the 25 'best' loci for dating as identified by SortaDate (Smith et al. 2018). They estimated divergence times using 166 unique minimum-age fossil calibrations (Ramırez-Barahona et al. 2020; Palazzesi et al. 2022; Smith and Beaulieu 2024; Zuntini et al. 2024), with a maximum constraint of 154 million years at the crown node, using the relative rate framework implemented in RelTime (Barba-Montoya et al. 2021).

As input for our model, we generated a list object for each trait model using a custom R script, containing each of the 676 clades from Brewer et al. (2025) as a sublist, as well as an additional sublist containing two elements:

1. The island age.
2. A vector with the total number of species in the mainland pool in each trait state.

For each of the 676 clades from Brewer et al. (2025), the generated sublist contains six elements:

1. The clade name.
2. The status of the colonising clade: this indicates whether the clade is an endemic singleton, endemic clade, or non-endemic to the island. Further details on the status coding can be found in the DAISIE package documentation.
3. The trait states for all species in the clade: a vector containing the discretised trait value for each species within the clade for which we have phylogenetic data or for a sequenced singleton clade (from Supplemen-

tary Data S2; if the clade contains more than one species for which we do not have phylogenetic data for, no trait state is recorded in the trait vector for that species. Table S3 shows details on how each trait was discretised. The distribution of the discretised trait values across the phylogeny of Canarian endemic non-monocot angiosperms is illustrated in Figure 1, and information on how clade composition varies per trait state is visualised in Figure 3. For the distribution of lineages across the trait states, see Figure 4.

4. The trait state at the root (trait state of the continental ancestor): a vector containing values equal to the number of examined states, e.g. a zero and a one to indicate the discretised trait state of the continental ancestor of the clade using Supplementary Data S3.
5. The sampling fraction of each trait state for each clade: the sampling fraction was calculated by dividing the number of sampled species of each trait state that are included in the tree by the total number of species with that trait state.
6. The phylogeny: if the clade had two or more species sequenced in the phylogeny of Brewer et al. (2025), we extracted the relevant clade from the phylogeny and added them to the sublist as a 'phylo' object. Otherwise, NA was assigned.

Applying TRAISIE to Empirical Data

We fitted a set of TRAISIE models to the Canary Islands dataset. We constructed and fitted nine different models for the two-state trait classifications (Table S4): one constant rate (CR) model, i.e. with colonisation, speciation and extinction constant across all trait states, but transition rates can be different; four examined trait-dependent models (ETD), and four concealed trait-dependent models (CTD), where diversification only depends on concealed traits. For the three-state trait classification, we used five different models (Table S5): one CR model and four ETD models. The CTD models were not run for the three-state trait model because the required number of parameters to be estimated for these models makes the computational demands on likelihood maximisation too large.

To initialise the search values for maximum likelihood optimisation of the parameters of each model, we used 250 sets of randomly generated initial values per model to improve exploration of parameter space. We used a maximum of nine optimisation cycles of the 'simplex' algorithm (available in the R package DDD), and a tolerance of 10–16 to help ensure convergence to global optima. Once the maximum likelihood parameters were identified for each model, we ran another optimisation step by rerunning TRAISIE using the parameters from the previous optimisation step as initial values with 'optimmethod' set

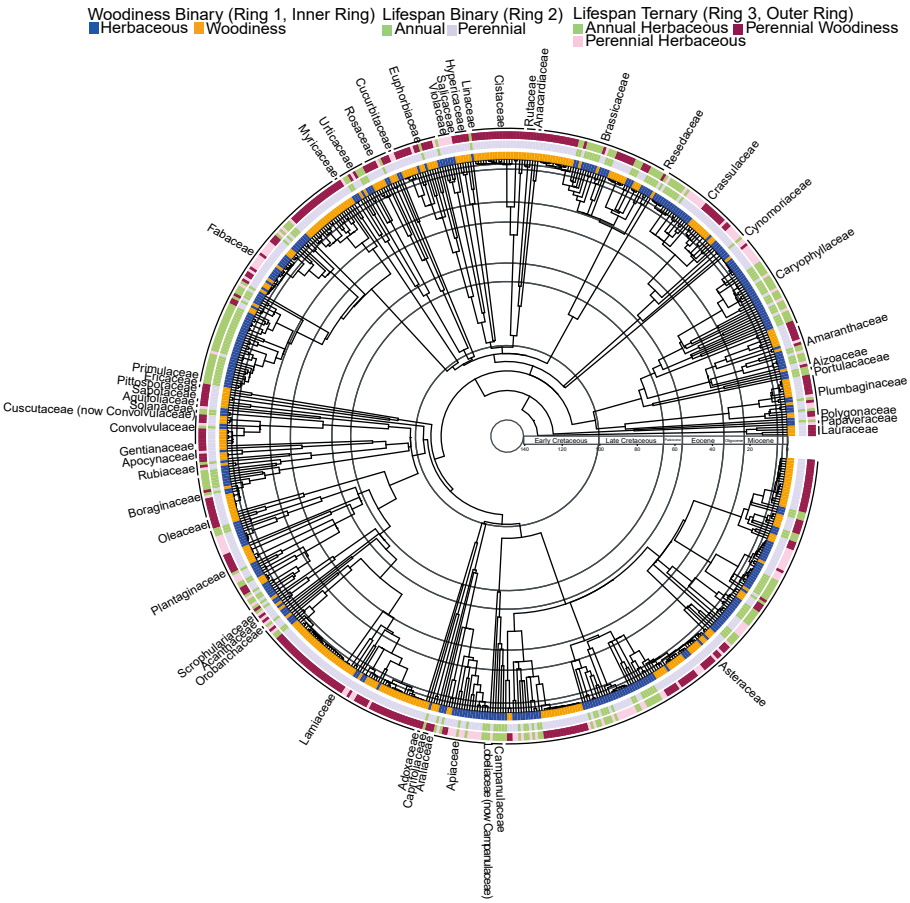


Figure 3: Phylogeny of Canary Islands non-monocot angiosperms, adapted from Brewer et al. (2025), with species at the tips, and their trait state across the three trait models around the tree: growth form (blue = herbaceous, brown = woody; Inner Ring), life span (green = annual, grey = perennial), and growth form + life span (green = annual herbaceous, pink = perennial herbaceous, red = perennial woodiness; Outer Ring). Taxonomic families are given outside the tree. The grey circles within the phylogeny denote geological periods.

to 'subplex' to help ensure convergence. All reported parameter estimates are from this second optimisation step. To assess the relative fit of each model, we used the Akaike Information Criterion (AIC) to select the best performing model for each study.

Finally, to assess whether the best-fitting models faithfully reproduce the observed phylogenetic and trait structure of Canary Island angiosperms, we simulated 200 datasets with the ML parameter estimates of the best-fitting model. We then plotted the distributions of several summary statistics from the simulated dataset, which we then compared with the corresponding values observed in the empirical Canary Islands dataset. To assess the effect of incom-

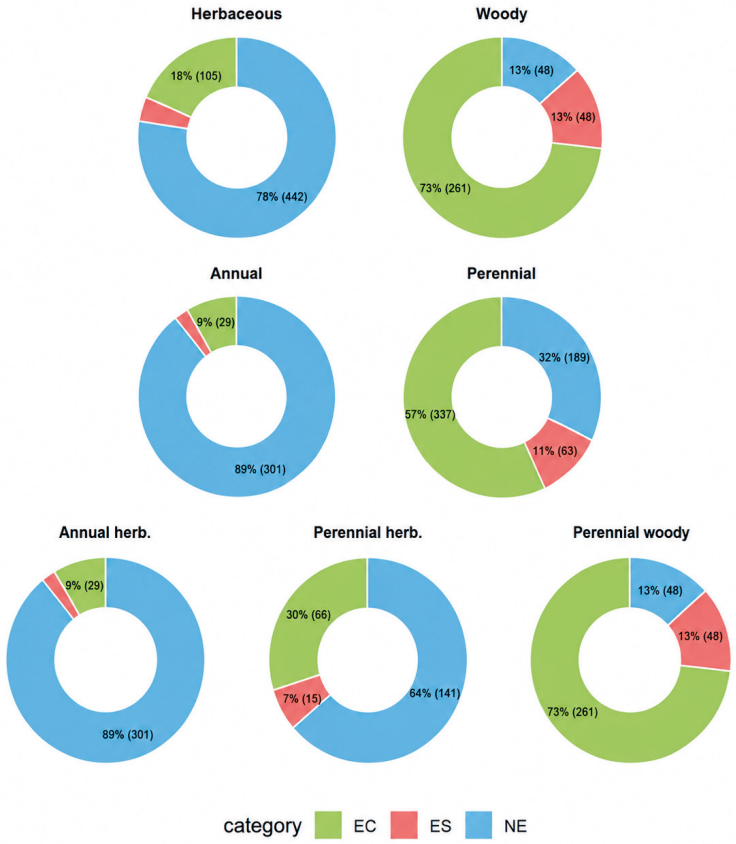


Figure 4: Distribution of non-endemic lineages (NE), endemic singletons (ES), and endemic-clades species (EC) across the different trait states used in our analyses for the Canary Islands non-monocot community.

plete sampling on goodness-of-fit, we then applied the same approach to simulated data generated under the best-fitting two-state growth form CTD model. Specifically, we compared 200 complete simulated datasets with no missing data to 200 simulated datasets containing approximately 32% missing data, in order to evaluate how well TRAISIE performs when data are incomplete.

RESULTS & DISCUSSION

Parameter Estimation on Simulated Data

Using a parametric bootstrap approach with 100 complete simulated datasets, we found that, overall across a broad set of conditions, our inference method (maximum likelihood) accurately recovered the true parameter values: the

median of the parameters estimated is close to the true generating value, and the variances were reasonably small for most parameters (Figures S1-S12). However, the variance for anagenesis and transition rates was noticeably larger than the variance for other parameters. The rate of anagenesis is often difficult to estimate in island diversification models (Valente et al. 2015, 2018). This is because anagenesis is primarily informed by endemic singleton lineages, yet endemic singletons can also be produced by cladogenetic diversification followed by extinction. This is especially true for an island system that seems to have an exceptionally high extinction rate, such as the Canaries (Brewer et al. 2025). This overlap in the processes that generate the same observed pattern likely explains why both TRAISIE and DAISIE models have difficulty estimating the anagenesis rate (Valente et al. 2018). As expected, for the large dataset, the variance of parameter estimates is consistently smaller than for the smaller datasets, across all models tested (see Figure S4, Figure S8, Figure S12). The variance associated with transition rate estimates for the larger dataset is larger than for the other parameters, but it is still smaller than for the smaller datasets.

Model Selection for Simulated Data

Our simulation results show that TRAISIE is often able to identify the true generating model, although performance depends strongly on the generating model, dataset size, and model-selection criterion (see emboldened values in Table 1). For datasets generated under the CR model, the CR model was correctly recovered on average in 50% of cases using AIC and in 94% of cases using BIC, whereas it was never selected under the ML criterion. This poor performance of ML for CR is expected because ML does not penalise model complexity and therefore will always favour the more parameter-rich ETD and CTD models if CR is a special case of them (that is, the ML for CTD or ETD will be the same as or larger than for CR). BIC performs well for CR: it also frequently selects CR when CTD or ETD generated the data, particularly for small datasets, indicating a high false-negative rate for trait dependence. This is because BIC strongly penalises model complexity and tends to favour simpler models.

For datasets generated under the ETD model, ETD was correctly recovered in 93% of simulations using AIC, 79% using BIC, and 93% using ML. Thus, ETD is generally well detected, especially with AIC and ML. This pattern is also reflected in the distribution of AIC weights assigned to ETD across simulated datasets (Figures 5–7). When ETD is the true generating model, the AIC weight of ETD is more concentrated near 1, indicating that ETD often receives strong support. By contrast, when the data are generated under CR or CTD, the AIC weight assigned to ETD is generally lower. CTD was more difficult to correctly

Table 1: Characteristics of the simulated datasets and model-selection performance of TRAISIE. One hundred datasets were simulated under each of the CR (6 parameters), ETD (10 parameters), and CTD (12 parameters) models. Under the CR model, all CES rates are assumed to be identical regardless of the observed and concealed trait states, and we assume non-zero transition rates. Under the ETD model, CES rates differ between species with different observed trait states, and we assume non-zero transition rates. Under the CTD model, CES rates differ only between species with different concealed states. Each dataset was fitted using all three models. Models were compared using the lowest AIC or BIC value, or the highest log-likelihood (ML). Rows show the number of cases in which each model was chosen by TRAISIE as the best-performing model according to each criterion. Bold-faced numbers show the number of datasets for which the generating model was selected as the best-performing model. Models of the same type in datasets Small–Large have identical generating parameter values (see Table S2) and differ only in mainland pool size (second column).

		Data Size				Model used for inference		
Set	Mainland pool size	Mean number of lineages	Mean number of species	Criterion	Generating model	CR	CTD	ETD
Small	(1000,1600)	39	84	ML	CR	**0**	44	56
				AIC	CR	**51**	8	41
				BIC	CR	**94**	0	6
		45	77	ML	CTD	0	**58**	42
				AIC	CTD	32	**23**	45
				BIC	CTD	56	**3**	41
		45	99	ML	ETD	0	18	**82**
				AIC	ETD	19	0	**81**
				BIC	ETD	51	0	**49**
Medium	(3000,4000)	105	222	ML	CR	**0**	29	71
				AIC	CR	**42**	3	55
				BIC	CR	**90**	0	10
		123	214	ML	CTD	0	**72**	28
				AIC	CTD	22	**29**	49
				BIC	CTD	57	**2**	41
		118	257	ML	ETD	0	4	**96**
				AIC	ETD	2	0	**98**
				BIC	ETD	13	0	**87**
Large	(14000,22000)	540	1146	ML	CR	**0**	22	78
				AIC	CR	**58**	5	35
				BIC	CR	**98**	0	2
		630	1096	ML	CTD	0	**83**	17
				AIC	CTD	14	**62**	24
				BIC	CTD	90	**6**	4
		618	1349	ML	ETD	0	0	**100**
				AIC	ETD	0	0	**100**
				BIC	ETD	0	0	**100**

identify. Under CTD-generated data, CTD was correctly selected on average in 38% of cases using AIC, only 4% using BIC, and 71% using ML. The poor performance of BIC for CTD again reflects its tendency to favour the simpler CR model, whereas the stronger performance of ML is likely driven by the larger number of parameters in CTD relative to CR and ETD. Overall, we recommend using AIC as the most appropriate general-purpose model-selection criterion for TRAISIE, as it provides the best compromise across the three generating

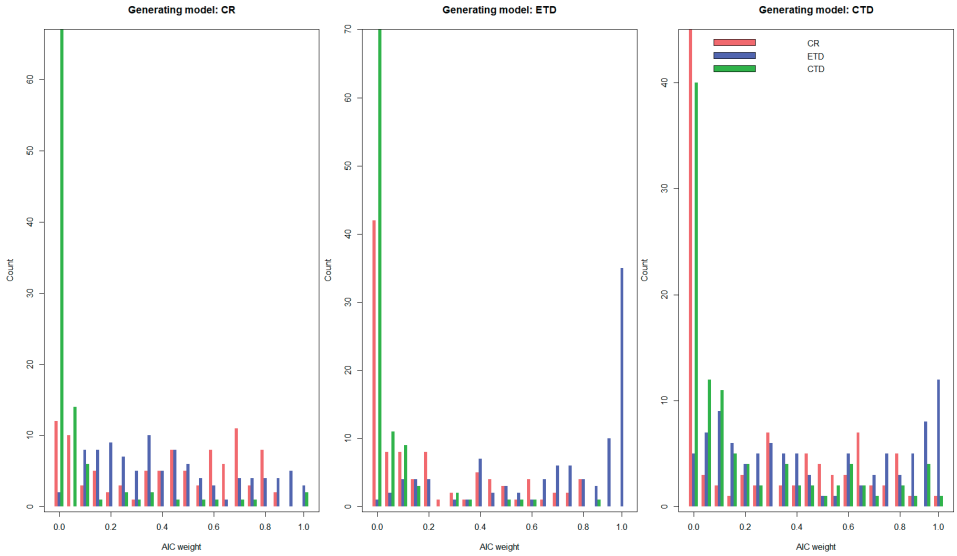


Figure 5: Distribution of AIC weights across 100 simulated datasets under the generating models CR, ETD, and CTD for the small-dataset scenario. Within each panel, grouped bars show the histogram of AIC weights assigned to CR (red), ETD (blue), and CTD (green).

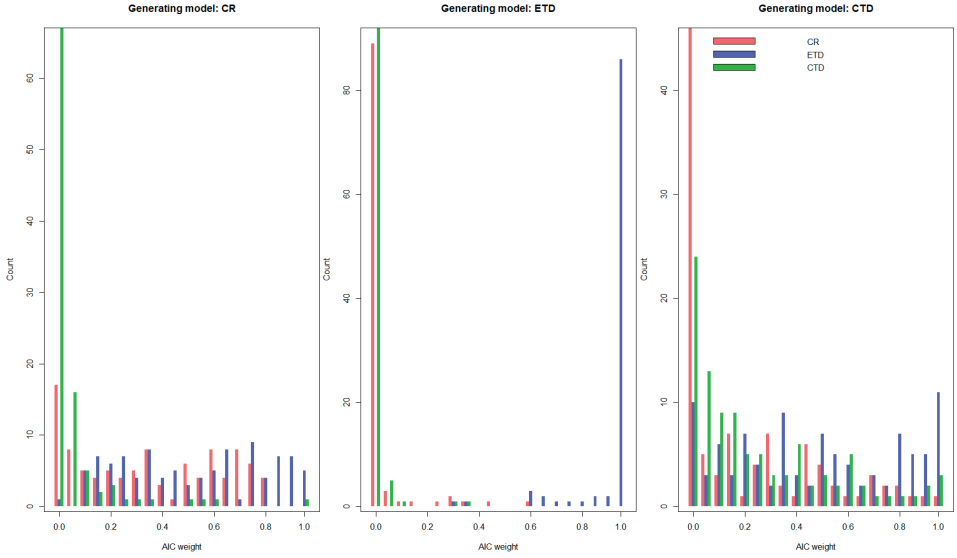


Figure 6: Distribution of AIC weights across 100 simulated datasets under the generating models CR, ETD, and CTD for the medium-dataset scenario. Within each panel, grouped bars show the histogram of AIC weights assigned to CR (red), ETD (blue), and CTD (green).

models. Averaged across all simulation settings, AIC selected the true model more often than either BIC or ML. On average, BIC performed best only for CR-generated data, whereas ML strongly favoured the more complex mod-

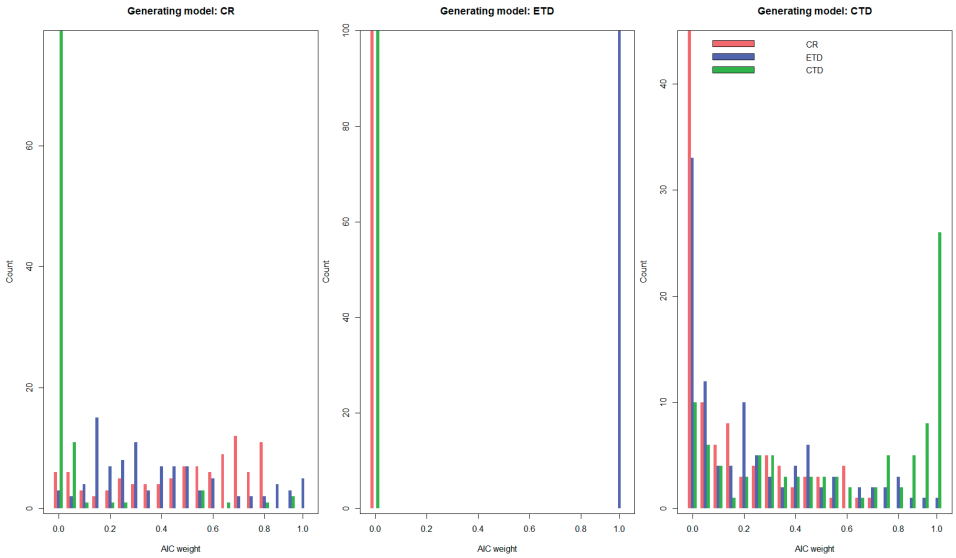


Figure 7: Distribution of AIC weights across 100 simulated datasets under the generating models CR, ETD, and CTD for the large-dataset scenario. Within each panel, grouped bars show the histogram of AIC weights assigned to CR (red), ETD (blue), and CTD (green).

els and therefore failed to recover CR. Nevertheless, CTD remains difficult to distinguish from the other models, especially for small datasets, and smaller datasets are therefore more prone to false negatives. To avoid false negatives, i.e. no signal of ETD when ETD is actually present, then ML or AIC are the best model selection criteria. However, to avoid false positives, i.e. ETD is selected, but the underlying model is CR or CTD, then BIC is the better option. More extensive simulation work, covering a broader range of parameter values and dataset sizes, will be needed to determine under which conditions parameter estimation and model selection are reliable.

Trait-Dependent Diversification in Canary Islands Angiosperms

Applying our approach to the Canary Island angiosperm dataset, we found that models with differential rates are always preferred over the CR models (Tables 2, S4), and CTD models are preferred over ETD models for the binary trait classifications. Under the preferred model for the two-state growth form classification (Model CTD2; Tables 2, S6), cladogenesis is high in one concealed trait state, with near-zero extinction, suggesting a fast diversification rate. For the other concealed state, the converse is found; cladogenesis is low, with high extinction, indicating a low diversification rate. In the high diversification rate trait state, colonisation is very low, whereas colonisation is high in the low diversification rate trait state. These results suggest starkly opposing diversifi-

cation processes between states: one with rampant in-situ diversification, and one reliant on colonisations to offset high extinction. However, there are very high transitions away from the high diversification trait state, with low transitions towards it. This suggests one state that can diversify rapidly when it has evolved, but this state is quite labile. Transitions between examined states are infrequent, and more common from herbaceousness towards woodiness. Although CTD models fit the data better, this does not mean that examined states do not affect diversification at all, so here we report results from the best fitting ETD model (ETD2; Tables 2, S6), but we caution against overinterpretation of these results.

Under the preferred ETD model, woody species in the Canary Islands have higher cladogenesis rates and lower extinction rates than herbaceous species (Table 2), in agreement with Brewer et al. (2025). Likewise, colonisation rates are also lower in woody species than in herbaceous species. Overall, the analysis suggests that species diversity in woody Canary Island non-monocot species is maintained by in-situ diversification, as speciation rates exceed extinction rates. In contrast, herbaceous species rely on new colonisations to maintain diversity. While previous studies have found inconclusive effects of growth form on diversification within specific clades (Soltis et al. 2013; Bouchenak-Khelladi et al. 2015; Neupane et al. 2017; Koutroumpa et al. 2021; Salomon et al. 2025; see Givnish 2010; Helmstetter et al. 2023 for review), analyses across angiosperms generally have found differential diversification rates (Louca and Pennell 2020; Brewer et al. 2025; Givnish 2010). Here, the higher rates attributed to woody species are likely due to the predominance of species that evolved woodiness in the Canaries following colonisation from an herbaceous ancestor (insular woodiness; Zizka et al. 2022) within radiating Canarian clades (Lens et al. 2013). Insular woodiness has been proposed to promote diversification in the Canary Islands (Carine et al. 2010), and 14 of the 20 largest radiations in the Canaries feature insular woody species, and 19 of the 20 include woody species (Supplementary Data S2). For the best fitting parameter estimates per model (using AIC), see Table S6.

Unexpectedly, under the preferred ETD model, the highest transition rates observed are from the woody state towards herbaceousness, while previous phylogenetic studies on the Canary Islands have identified multiple (at least 41; Brewer et al. 2025) transitions from herbaceousness towards insular woodiness (Lens et al. 2013; Zizka et al. 2022; Brewer et al. 2025). We hypothesise that the model explains the presence of large native clades with a predominance of herbaceous species by high transition rates towards the herbaceous state. In the empirical data, more than half of all colonising clades are herbaceous non-endemic singletons (442 of 677), which suggests that herbaceous lineages

Table 2: Parameter estimates from the best fitting two-state and three-state TRAISIE models. λ^c cladogenetic speciation; μ - extinction rate; ɣ - colonisation rate; λ^a anagenetic speciation. Numeric subscripts of the parameter symbols denote trait states (for the two-state growth form model, 0 = herbaceous; 1 = woody. For the two-state life span model, 0 = annual; 1 = perennial. For the three-state model, 0 = annual herbaceous; 1 = perennial herbaceous; 2 = perennial woody). Letters after the parameter symbols refer to examined or concealed traits (A = examined state, B = concealed). The CTD and ETD models shown here are the best fitting CTD and ETD models for the two-state datasets (see Table S6, 7). For the three-state data set, we include the best fitting ETD model (as we did not fit CTD models).

Parameters	Growth Form (two-state) CTD	Growth Form (two-state) ETD	Life Span (two-state) CTD	Life Span (two-state) ETD	Life Span (three-state) ETD
λ^c_{0A}	0.053	0.115	0.048	0.046	0.048
λ^c_{0B}	2.384	0.115	2.259	0.046	—
λ^c_{1A}	0.053	0.378	0.048	0.317	0.195
λ^c_{1B}	2.384	0.378	2.259	0.317	—
λ^c_{2A}	—	—	—	—	0.407
$\mu 0A$	0.627	0.876	0.589	1.029	1.416
$\mu 0B$	1.3E-11	0.876	7.E-24	1.030	—
$\mu 1A$	0.627	0.268	0.589	0.407	0.283
$\mu 1B$	1.3E-11	0.268	7.E-24	0.407	—
$\mu 2A$	—	—	—	—	0.458
$\gamma 0A$	0.039	0.034	0.037	0.062	0.099
$\gamma 0B$	0.001	0.034	0.001	0.063	—
$\gamma 1A$	0.039	0.004	0.037	0.006	0.007
$\gamma 1B$	0.001	0.004	0.001	0.006	—
$\gamma 2A$	—	—		—	0.005
λ^a_{0A}	2.1E-29	5.4E-37	2.E-13	9.0E-25	—
λ^a_{0B}	1.6E-11	5.4E-37	6.E-19	9.0E-25	—
λ^a_{1A}	2.1E-29	6.3E-26	2.E-13	1.3E-27	—
λ^a_{1B}	1.6E-11	6.3E-26	6.E-19	1.3E-27	—
λ^a_{2A}	—	—	—	—	—
$q0A \rightarrow 1A$	0.0456	0.018	0.0543	0.028	0.038
$q0B \rightarrow 1B$	0.0456	0.018	0.0543	0.028	0.038
$q0A \rightarrow 0B$	0.0461	—	0.043	—	—
$q0B \rightarrow 0A$	2.828	—	2.689	—	—
$q1A \rightarrow 0A$	0.008	0.159	0.005	0.048	0.183
$q1B \rightarrow 0B$	0.008	0.159	0.005	0.048	0.183
$q1A \rightarrow 1B$	0.0461	—	0.043	—	0.249
$q1B \rightarrow 1A$	2.828	—	2.689	—	0.249
$q1A \rightarrow 2A$	—	—	—	—	0.053
$q2A \rightarrow 1A$	—	—	—	—	0.025
Log-lik	-4586.175	-4700.85	-4477.472	-4539.59	-4671.911
DF	12	10	12	10	13
AIC	9196.350	9421.700	8978.944	9099.180	9369.822
BIC	9250.562	9466.877	9033.156	9144.357	9428.552

may have low speciation rates. But at the same time, among a few larger native clades, we also observe a predominance of herbaceous species on the Canaries (e.g., *Lotus* (Fabaceae), *Pericallis* (Asteraceae), *Silene* (Caryophyllaceae)). We believe that the model explains the presence of such clades by high transition rates toward the herbaceous state, inferring diversification to have occurred in a woody state and followed by these clades transitioning to herbaceousness. However, existing literature strongly suggests that the ancestral state of the closest mainland relatives is herbaceous (Jaen-Molina et al. 2021; Sandral et al. 2006; Cron et al. 1999; Jones et al. 2014; Naciri et al. 2017). The presence of these herbaceous endemic clades may also indicate that another, unmeasured trait associated with the herbaceous state promoted diversification within these clades.

In the best fitting two-state life span model (CTD2; Tables 2, S7), the same pattern is observed as in the preferred two-state growth form model: one state with a fast diversification rate, but frequent transitions away from this state; and one state with a low diversification rate, reliant on colonisations to maintain diversity. Transitions are also infrequent between examined states, and more common from annuality towards perenniality than the reverse. When considering the preferred ETD model (ETD2; Tables 2, S7), perennial species have higher cladogenesis rates and lower extinction and colonisation rates than annual species (Table 2). Anagenesis is also near zero across all states. However, unlike in the best fitting growth form model, both trait states rely on new colonisation events to maintain in-situ diversity, as extinction rates exceed speciation rates. Notably, the estimated extinction rate is very high in annual species, as is the colonisation rate. Also, unexpectedly, transitions from perenniality towards annuality are the highest. Shifts from perenniality towards annuality are only observed in a few perennial radiating clades (e.g. the *Aeonium* Alliance (Crassulaceae), *Echium* (Boraginaceae), *Tolpis* (Asteraceae); Supplementary Data S2, S3), meaning the inverse was expected. For the best fitting parameter estimates per model (using AIC), see Table S7.

For the three-state life span model, we find the best fitting model (ETD3) allows each trait state and transition rate to vary, except anagenesis, which was fixed to zero (to reduce the number of parameters estimated; Table S5). Under this model, both perennial herbaceous and perennial woody species have higher cladogenesis rates than annual herbaceous species, with woody species having markedly faster cladogenesis rates than perennial herbaceous species (Table 2). This suggests the low cladogenesis rate observed for the herbaceous state in the two-state growth form model is primarily attributed to annual herbaceous species (that are mostly represented by non-endemic clades; see Figure 4), as perennial herbaceous species maintain their high cladogenesis rates,

as observed in the two-state life span model. Annual herbaceous species have a higher extinction rate than both perennial states, with woody species having a higher extinction rate than perennial herbaceous species. These results are in agreement with the following patterns in the empirical data: i) 89% of annual herbaceous clades are singletons, indicative of low speciation and/or high extinction rates (Figure 4); ii) the higher percentage of perennial woody endemic singleton clades than perennial herbaceous endemic singleton clades suggests higher anagenesis or cladogenesis followed by extinction (Figure 4); and iii) the older average age of woody Canarian clades than herbaceous clades (Brewer et al. 2025), which TRAISIE can infer as high extinction rates in the presence of old, singleton clades. As with extinction rates, colonisation rates are also highest in annual herbaceous species, suggesting this state relies upon frequent colonisation to maintain diversity within the Canaries. In contrast, perennial states have considerably lower colonisation rates. Woody species have the lowest colonisation rates, but depend on a combination of their low colonisation rate and low rate of transitions towards woodiness to maintain insular diversity, whereas perennial herbaceous species rely on insular diversification to maintain their diversity. These results align with those of Brewer et al. (2025), who found that both perennial and woody clades (separately) had higher cladogenesis rates than annual and herbaceous clades, respectively, but lower colonisation rates. As discussed above, these results are not unexpected given the predominance of perennial and woody species within Canarian radiating clades, with 19 of the 20 largest radiations comprising mostly of perennial herbaceous or woody species (Supplementary Data S2). The frequent transitions from woodiness to herbaceousness estimated in the two-state model are not observed in the three-state model (Table 2); rather, the higher transition rates are from perennial herbaceousness to annual herbaceousness. The high number of annual herbaceous non-endemic clades, low speciation rate, and presence of annual herbaceous endemic clades (e.g. *Vicia* (Fabaceae), *Carduus* (Asteraceae), *Ononis* (Fabaceae)), may also interact here to infer high transition rates towards the annual herbaceous state. For the best fitting parameter estimates per model (using AIC), see Table S8.

Goodness of Fit of the Models

To assess how well the preferred models reproduce the observed data, we evaluated their goodness of fit by simulating datasets under the corresponding parameter estimates. We simulated 200 datasets with parameters obtained from the preferred models (Table 2), and compared the summary statistics of these simulated datasets with those of the empirical datasets. Overall, the fit varied across the best fitting models of the three datasets.

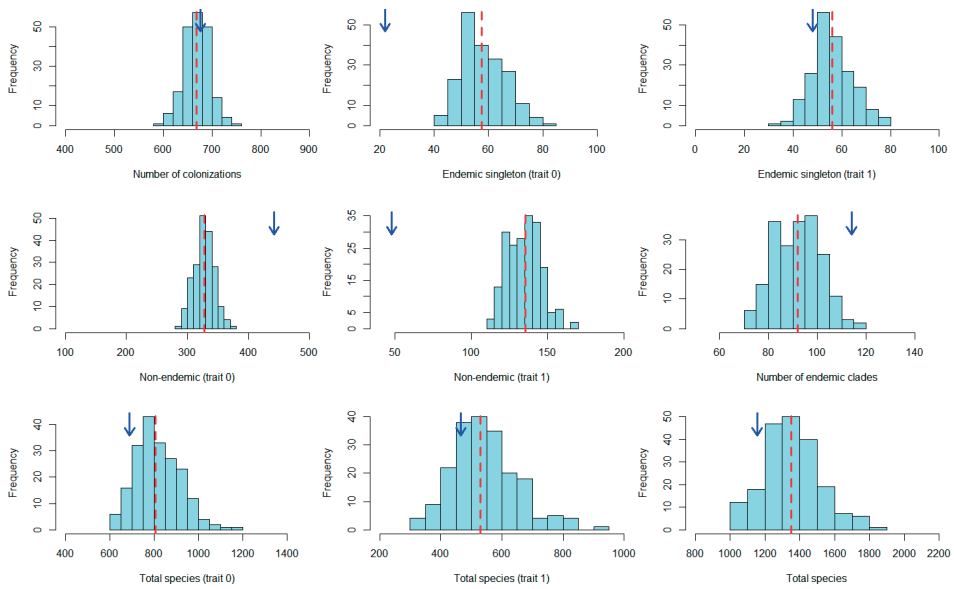


Figure 8: Goodness-of-fit plots for the two-state growth form model. Histograms show the distribution of the summary statistics based on 200 datasets simulated using the parameters from the best fitting model (CTD2; Tables 2 and S4). The dashed red line shows the median of the simulated values, and the blue arrow shows the observed value from the empirical data (see Table S3). Trait 0 = herbaceous, trait 1 = woody.

For the two-state growth form model, the simulations show that four summary statistics are well reproduced: the number of colonisations, number of endemic singletons with trait state one (i.e., woodiness), total species with trait state one, and total species with trait state zero (i.e., herbaceousness), with the median from the simulated datasets falling close to the observed values for these four summary statistics (Figure 8). However, other summary statistics were poorly predicted, with the true values falling outside, or on the fringes of, the distribution of simulated data (such as the number of endemic clades, non-endemic species with trait zero, and non-endemic species with trait one. The total number of species is generally also overestimated. We hypothesise that the poor fit may be partly be explained by missing species in the empirical dataset (see below). We also expect that mixed models (ECTD models), combining aspects of the best fitting ETD and CTD models, may be more biologically meaningful and therefore potentially provide a better fit to the data.

For the preferred two-state life span model, the observed values fell on the fringes or outside the distribution of simulated values for nearly all summary statistics (Figure 9), except the total number of colonisations and total number of species. For the three-state life span model, goodness-of-fit is also poor for most of the plotted summary statistics (Figure 10).

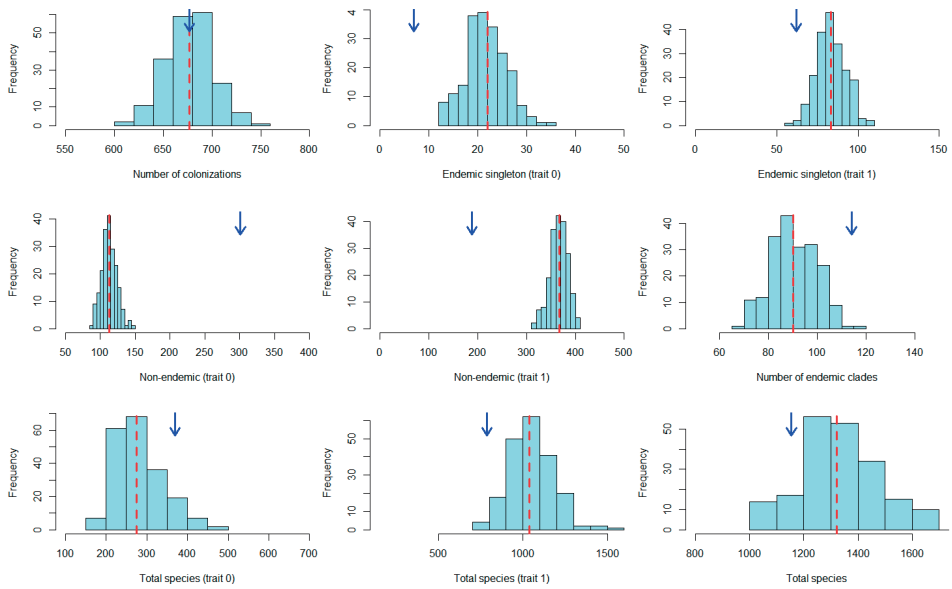


Figure 9: Goodness-of-fit plots for the two-state life span model. Histograms show the distribution of the summary statistics based on 200 datasets simulated using the parameters from the best fitting model (CTD2; Tables 2, S4). The dashed red line shows the median of the simulated values, and the blue arrow shows the observed value from the empirical data (see Table S3). Trait 0 = annual, trait 1 = perennial.

The poor goodness-of-fit across all models could be partly attributed to the missing data in the empirical dataset. Despite being the most complete phylogenomic dataset of any insular plant community (Brewer et al. 2025), 35% of non-monocot lineages, and ~40% of native Canarian non-monocot species are not sampled in the phylogeny. These missing species are indirectly included in the analyses by the specification of the sampling fraction for each trait state. However, this amount of missing data may simply be too high to generate a strong fit of TRAISIE models to the data. Additionally, these poor fits could be due to the absence of endemic clades/singletons with recolonisations of their non-endemic mainland ancestor in our dataset (in other words, our dataset does not include instances where the same mainland species has colonised the Canaries multiple times). As our dataset only contains a single island representative per species, recolonisations are difficult to identify (see Brewer et al. (2025) for more details on sampling strategy). Given the close distance of the Canary Islands to the mainland, it is expected that recolonisations of mainland species are frequent, although population-level phylogenetic studies of non-endemic Canary Island species are sparse to validate this statement (García-Verdugo et al. 2019; Martín-Hernanz et al. 2023; Rincon Barrado et al. 2024; Arjona et al. 2025; Caujape-Castells et al. 2011). The underestimation of recolo-

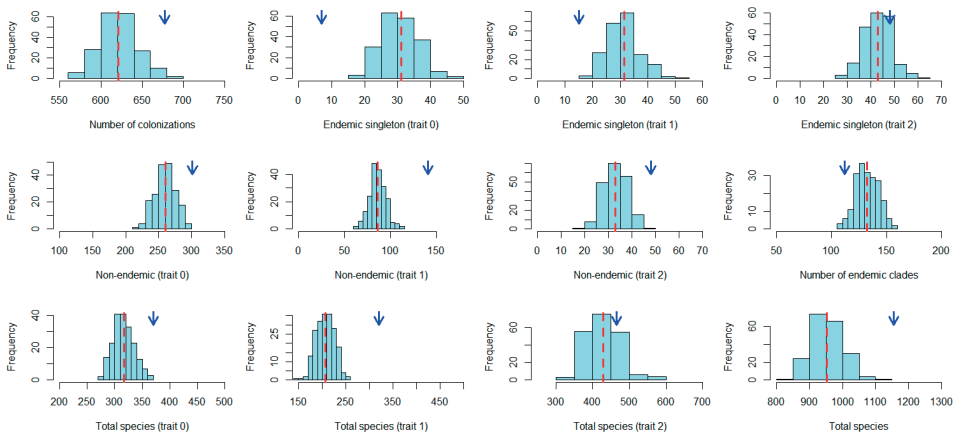


Figure 10: Goodness-of-fit plots for the three-state life span model. Histograms show the distribution of simulated values, generated using the parameters from the best fitting model (ETD3; Tables 2, S5). The dashed red line shows the median of the simulated values, and the blue arrow shows the observed value from the empirical data (see Table S3). Trait 0 = annual herbaceous, trait 1 = perennial herbaceous, trait 2 = perennial woody.

nisations for Canary Island lineages would affect colonisation and anagenesis rate estimates, which could lead to poor goodness-of-fit.

Beyond these data and sampling limitations, it is possible that neither the CTD nor the ETD model tested here provides an adequate fit to the data. The observed trait categories considered (woody vs. herbaceous, annual vs. perennial, and annual herbaceous vs. perennial herbaceous vs. woody) may be too simplistic to fully capture the diversification dynamics of the Canary Islands angiosperm flora. For example, we have already seen that annual and perennial herbs differ in their diversification behaviour, and it is plausible that woody species are similarly heterogeneous, for instance, between insular woody and non-insular woody lineages. One possible hypothesis is therefore that the higher cladogenesis associated with woody species is driven mainly by insular woody lineages, as suggested by the strong representation of insular woody lineages among the 20 largest radiations (Lens et al. 2013; Brewer et al. 2025). We note that the design of these models should be mindful not to include too many parameters due to the computational demands on likelihood maximisation. Despite poor model fit, there appears to be an effect of our examined traits on diversification, as all ETD models detect differential rates when the examined states were allowed to vary (Tables S6-S8).

To explore how missing data affects model inference and assess whether this can explain the poor goodness of fit of the models to the empirical data, we conducted a post hoc analysis, using a simulation study, with a complete and an incomplete dataset. The aim was to assess whether the poor fit observed for the

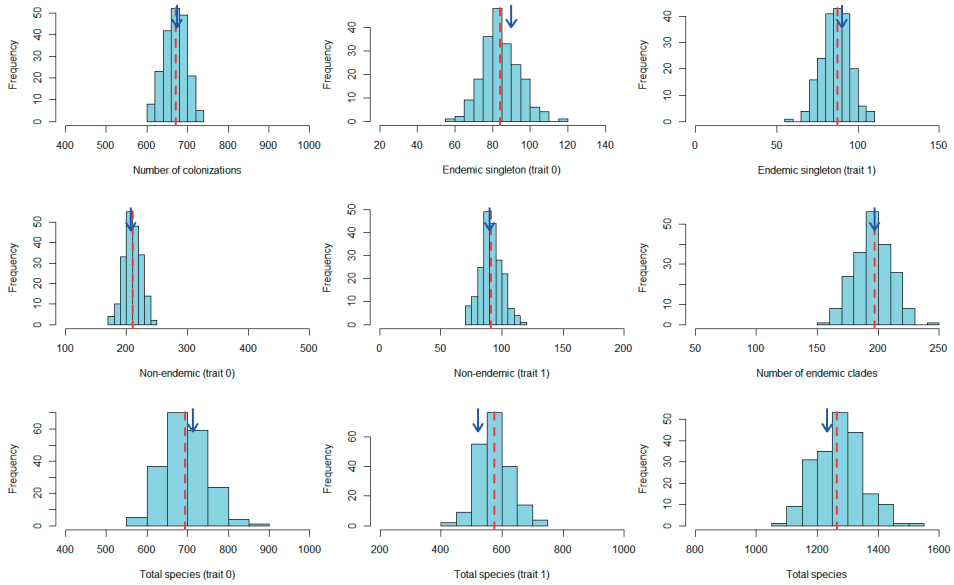


Figure 11: Goodness-of-fit plots for 200 complete simulated datasets (i.e., with no missing data). Histograms show the distribution of some summary statistics of the simulated data, generated using parameters based on the best fitting two-state growth form CTD model (CTD2). The dashed red line shows the median of the simulated values, and the blue arrow shows the generating value (see Table S3).

empirical data could partly be explained by the reduced performance of the method when datasets are incompletely sampled.

To do this, we simulated a single dataset under the best fitting two-state growth form model (CTD2), then fitted that same model to this dataset, and used the resulting parameter estimates to generate 200 replicate simulated datasets. We then compared a set of summary statistics (e.g., the number of colonisations, the number of endemic singleton lineages with trait zero or one, the number of non-endemic species with trait zero or one, the number of endemic clades, the total number of species with trait zero or one, and total species richness) computed from these 200 simulated datasets to the corresponding statistics of the original complete simulated dataset. The distributions of summary statistics computed from the 200 simulated datasets all overlap with the corresponding summary statistics of the complete simulated dataset (Figure 11). This indicates that when using a complete dataset, the maximum likelihood parameters estimated from TRAISIE generate datasets that closely resemble the original dataset across all summary statistics. We conducted the same goodness-of-fit analysis on an incompletely sampled version of the simulated dataset. We randomly removed 32% of the species from that dataset to assess the effect of missing species on inference. For the incomplete dataset, the fit

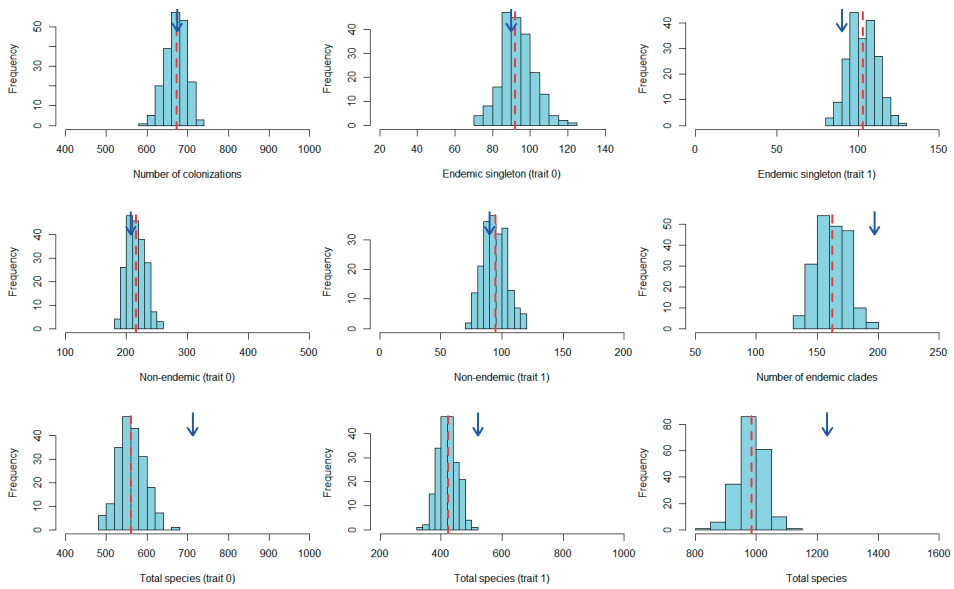


Figure 12: Goodness-of-fit plots for 200 incomplete simulated datasets (with 221 missing species across 20 clades with more than two species). Histograms show the distribution of some summary statistics of the simulated data, generated using parameters based on the best fitting two-state growth form CTD model (CTD2). The dashed red line shows the median of the simulated values, and the blue arrow shows the generating value (see Table S3).

is noticeably poorer than for the complete dataset (Figure 12). This suggests that the model is sensitive to incomplete sampling. TRAISIE uses ρ−sampling, which assumes that each species with the trait state i is randomly sampled with probability ρ_i, representing the proportion of species with trait i included in the phylogeny. However, this approach does not use information about the actual number of species missing from the tree, which may explain the observed underestimation of total species richness in the goodness-of-fit plot (Figure 12). We hypothesise that the impact of missing species could be reduced by using n-sampling (Valente et al. 2015), which explicitly incorporates the number of missing species into the model.

In conclusion, we show that with a moderate dataset size (roughly 200 species), the effect of trait evolution on diversification rates can be detected using the TRAISIE model, and that parameter estimates can be recovered with relatively high precision and recall. A more extensive simulation study is needed, however, to determine the dataset sizes required for reliable parameter inference. Our analyses of Canary Islands angiosperms indicate that lineages differing in growth form and life span follow distinct diversification trajectories: the ETD models suggest that woody and perennial lineages tend to show higher cladogenesis rates and greater reliance on in-situ diversification than

herbaceous and annual lineages, respectively. However, CTD models are a better fit for the empirical data. This suggests that the observed traits (growth form and life span) alone may not explain the diversification dynamics of the Canary angiosperms community, either because additional concealed factors are involved, or that more refined trait categories should be considered (e.g., insular woody versus non-insular woody lineages). Alternatively, the models considered here do not capture the underlying diversification processes in the Canary Island angiosperms. Models that combine features of both examined (observed) and concealed trait-dependent diversification should be explored in future work. We also observe that the goodness of fit of the preferred models to the Canary Islands data is poor; thus, these results should be interpreted with caution. Our simulation analyses suggest that this poor fit may be partly explained by missing species in the empirical dataset, as simulation analyses showed that the true generating model fits well when data are complete, but less well when data are taxonomically undersampled. Altogether, despite these limitations, our results reinforce the importance of growth form and life span in shaping the remarkable diversity of island floras, and highlight the value of integrating trait evolution into models of island biogeography.

REFERENCES

Anderson, B., J. Pannell, S. Billiard, C. Burgarella, H. de Boer, M. Dufay, A. J. Helmstetter, M. Mendez, S. P. Otto, D. Roze, et al. (2023). Opposing effects of plant traits on diversification. *Iscience,* 26(4).

Antonelli, A., R. Govaerts, E. Nic Lughadha, R. Onstein, R. J. Smith, and A. Zizka (2023). Why plant diversity and distribution matter. *New Phytologist,* 240(4), 1331–1336.

Arjona, Y., L. S. Jay-García, J. A. Reyes-Betancort, J. Morente-Lopez, M. Salas-Pascual, A. Naranjo-Cigala, G. Sicilia-Pasos, M. A. Padron-Mederos, R. Orihuela-Rivero, R. González-Montelongo, et al. (2025). Taming Darwin's conundrum: On the role of phylogenetic and functional differentiation in plant naturalization on oceanic islands. *bioRxiv,* 2025–04.

Barajas Barbosa, M. P., D. Craven, P. Weigelt, P. Denelle, R. Otto, S. D´ıaz, J. Price, J. M. Fernandez-Palacios,´ and H. Kreft (2023). Assembly of functional diversity in an oceanic island flora. *Nature,* 619(7970), 545–550.

Barba-Montoya, J., Q. Tao, and S. Kumar (2021). Assessing rapid relaxed-clock methods for phylogenomic dating. *Genome Biology and Evolution,* 13(11), evab251.

Beaulieu, J. M. and B. C. O'Meara (2016). Detecting hidden diversification shifts in models of trait-dependent speciation and extinction. *Systematic Biology,* 65(4), 583–601.

Bellvert, A., L. J. Pollock, A. Kaliontzopoulou, and M. A. Arnedo (2025). Tempo and mode of diversification of the red devil spiders (araneae: Dysderidae) of the Canary Islands. *Zoologica Scripta,* 54(3), 406–420.

Bouchenak-Khelladi, Y., R. E. Onstein, Y. Xing, O. Schwery, and H. P. Linder (2015). On the complexity of triggering evolutionary radiations. *New Phytologist,* 207(2), 313–326.

Brewer, R. F. A., O. D. Nanwou, L. van Hoek, M. J. del Arco Aguilar, Y. Arjona, W. J. Baker, N. M. van den Berg, J. Caujape-Castells, J. M. Fern´ andez-Palacios, C. Gonz´ alez-Montelongo, et al. (2025). Trait-dependent´ diversification and spatio-ecological limits drive angiosperm diversity unevenness across the Canary Islands archipelago. *bioRxiv,* 2025–09.

Brown, M. J., B. E. Walker, N. Black, R. H. Govaerts, I. Ondo, R. Turner, and E. Nic Lughadha (2023). rwcvp: a companion R package for the World Checklist of Vascular Plants. *New Phytologist,* 240(4), 1355–1365.

Carine, M. A., A. Santos-Guerra, I. R. Guma, and J. A. Reyes-Betancort (2010). *Endemism and evolution of the Macaronesian flora*. Beyond cladistics: The branching of a paradigm 3, 101.

Caujape-Castells, J., D. Bramwell, et al. (2011). *Jesters, red queens, boomerangs and surfers: a molecular´ outlook on the diversity of the Canarian endemic flora*. The biology of island floras, 284–324.

Cron, G., K. Balkwill, and E. Knox (1999). Cineraria lyratiformis (Senecioneae, Asteraceae), a new name for *Cineraria lyrata*. *South African Journal of Botany,* 65(4), 287–290.

Darwin, C. (1859). *On the origin of species by means of natural selection*. John Murray.

Dehayem, O. N., B. Haegeman, and R. Etienne (2025). An alternative approach to compute the likelihood for the daisie (dynamic assembly of island biota through speciation, immigration, and extinction) model. *bioRxiv,* 2025–09.

Denelle, P., P. Weigelt, and H. Kreft (2023). Gift—an r package to access the global inventory of floras and traits. *Methods in Ecology and Evolution,* 14(11), 2738–2748.

Díaz, S., Kattge, J., Cornelissen, J. H., Wright, I. J., Lavorel, S., Dray, S., & Zotz, G. (2022). The global spectrum of plant form and function: enhanced species-level trait dataset. *Scientific Data,* 9(1), 755.

Etienne, R. S., B. Haegeman, A. Dugo-Cota, C. Vil´ a, A. Gonzalez-Voyer, and L. Valente (2023). The phylogenetic limits to diversity-dependent diversification. *Systematic Biology,* 72(2), 433–445.

FitzJohn, R. G. (2010). Quantitative traits and diversification. *Systematic Biology,* 59(6), 619–633.

FitzJohn, R. G. (2012). Diversitree: comparative phylogenetic analyses of diversification in r. *Methods in Ecology and Evolution,* 3(6), 1084–1092.

FitzJohn, R. G., W. P. Maddison, and S. P. Otto (2009). Estimating trait-dependent speciation and extinction rates from incompletely resolved phylogenies. *Systematic Biology,* 58(6), 595–611.

Garcia-Verdugo, C., J. Caujape-Castells, M. Mairal, and P. Monroy (2019). How repeatable is microevolution on islands? Patterns of dispersal and colonization-related plant traits in a phylogeographical context. *Annals of Botany,* 123(3), 557–568.

Gillespie, R. G. (2016). Island time and the interplay between ecology and evolution in species diversification. *Evolutionary applications,* 9(1), 53–73.

Givnish, T. J. (2010). Ecology of plant speciation. *Taxon,* 59(5), 1326–1366.

Goldberg, E. E. and B. Igic (2012). Tempo and mode in plant breeding system evolution. *Evolution,* 66(12), 3701–3709.

Goldberg, E. E., L. T. Lancaster, and R. H. Ree (2011). Phylogenetic inference of reciprocal effects between geographic range evolution and diversification. *Systematic Biology,* 60(4), 451–465.

Hanz, D. M., V. Cutts, M. P. Barajas-Barbosa, A. C. Algar, C. Beierkuhnlein, J.-M. Fernandez-Palacios,´ R. Field, H. Kreft, M. J. Steinbauer, P. Weigelt, et al. (2022). Climatic and biogeographical drivers of functional diversity in the flora of the Canary Islands. *Global Ecology and Biogeography,* 31(7), 1313–1331.

Helmstetter, A. J., R. Zenil-Ferguson, H. Sauquet, S. P. Otto, M. Mendez, M. Vallejo-Marin, J. Schonenberger, C. Burgarella, B. Anderson, H. De Boer, et al. (2023). Trait-dependent diversification in angiosperms: Patterns, models and data. *Ecology Letters,* 26(4), 640–657.

Henao Diaz, L. F., L. J. Harmon, M. T. Sugawara, E. T. Miller, and M. W. Pennell (2019). Macroevolutionary diversification rates show time dependency. *Proceedings of the National Academy of Sciences,* 116(15), 7403–7408.

Herrera-Alsina, L., P. Van Els, and R. S. Etienne (2019). Detecting the dependence of diversification on multiple traits from phylogenetic trees and trait data. *Systematic Biology* 68(2), 317–328.

Illera, J. C., J. C. Rando, M. Melo, L. Valente, and M. Stervander (2024). Avian island radiations shed light on the dynamics of adaptive and nonadaptive radiation. *Cold Spring Harbor Perspectives in Biology,* 16(12), a041451.

Jaén-Molina, R., A. Marrero-Rodríguez, J. Caujape-Castells, and D. I. Ojeda (2021). Molecular phylogenetics of lotus (leguminosae) with emphasis in the tempo and patterns of colonization in the Macaronesian region. *Molecular Phylogenetics and Evolution,* 154, 106970.

Johnson, M. G., L. Pokorny, S. Dodsworth, L. R. Botigue, R. S. Cowan, A. Devault, W. L. Eiserhardt, N. Epitawalage, F. Forest, J. T. Kim, et al. (2019). A universal probe set for targeted sequencing of 353 nuclear genes from any flowering plant designed using k-medoids clustering. *Systematic Biology,* 68(4), 594–606.

Jones, K. E., J. A. Reyes-Betancort, S. J. Hiscock, and M. A. Carine (2014). Allopatric diversification, multiple habitat shifts, and hybridization in the evolution of Pericallis (Asteraceae), a Macaronesian endemic genus. *American Journal of Botany,* 101(4), 637–651.

Kattge, J., G. Bonisch, S. Díaz, S. Lavorel, I. C. Prentice, P. Leadley, S. Tautenhahn, G. D. Werner, T. Aakala, M. Abedi, et al. (2020). Try plant trait database–enhanced coverage and open access. *Global Change Biology,* 26(1), 119–188.

Koutroumpa, K., B. H. Warren, S. Theodoridis, M. Coiro, M. M. Romeiras, A. Jimenez, and E. Conti (2021). Geo-climatic changes and apomixis as major drivers of diversification in the Mediterranean sea lavenders (Limonium mill.). *Frontiers in Plant Science,* 11, 612258.

Kozlov, A. M., D. Darriba, T. Flouri, B. Morel, and A. Stamatakis (2019). Raxml-ng: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics,* 35(21), 4453–4455.

Lagomarsino, L. P., F. L. Condamine, A. Antonelli, A. Mulch, and C. C. Davis (2016). The abiotic and biotic drivers of rapid diversification in a ndean bellflowers (Campanulaceae). *New Phytologist,* 210(4), 1430–1442.

Lambert, J., P. Neves, R. Bilderbeek, L. Valente, and R. Etienne (2022). The effect of mainland dynamics on data and parameter estimates in island biogeography. preprint. *Evolutionary Biology*.

Lens, F., N. Davin, E. Smets, and M. del Arco (2013). Insular woodiness on the Canary Islands: a remarkable case of convergent evolution. *International Journal of Plant Sciences,* 174(7), 992–1013.

Lerner, H. R., M. Meyer, H. F. James, M. Hofreiter, and R. C. Fleischer (2011). Multilocus resolution of phylogeny and timescale in the extant adaptive radiation of Hawaiian honeycreepers. *Current Biology,* 21(21), 1838–1844.

Losos, J. B. (2009). *Lizards in an evolutionary tree: ecology and adaptive radiation of anoles*, Volume 10. Univ of California Press.

Losos, J. B. and R. E. Ricklefs (2009). Adaptation and diversification on islands. *Nature,* 457(7231), 830–836.

Louca, S. and M. W. Pennell (2020). A general and efficient algorithm for the likelihood of diversification and discrete-trait evolutionary models. *Systematic Biology,* 69(3), 545–556.

MacArthur, R. H. and E. O. Wilson (1967). *The Theory of Island Biogeography* (Revised ed.). Princeton, NJ: Princeton University Press.

Maddison, W. P., P. E. Midford, and S. P. Otto (2007). Estimating a binary character's effect on speciation and extinction. *Systematic Biology,* 56(5), 701–710.

Martín-Hernanz, S., M. Nogales, L. Valente, M. Fernández-Mazuecos, F. Pomeda-Gutiérrez, E. Cano, P. Mar-rero, J. M. Olesen, R. Heleno, P. Vargas (2023). Time-calibrated phylogenies reveal Mediterranean and pre-Mediterranean origin of the thermophilous vegetation of the Canary Islands. *Annals of Botany,* 131(4), 667–684.

McPeek, M. A. and J. M. Brown (2007). Clade age and not diversification rate explains species richness among animal taxa. *The American Naturalist,* 169(4), E97–E106.

Mort, M. E., D. E. Soltis, P. S. Soltis, J. Francisco-Ortega, and A. Santos-Guerra (2002). Phylogenetics and evolution of the Macaronesian clade of Crassulaceae inferred from nuclear and chloroplast sequence data. *Systematic Botany,* 27(2), 271–288.

Muer, T., H. Sauerbier, and F. C. Calixto (2016). *Die Farn-und Blutenpflanzen der Kanarischen Inseln* . Margraf Publishers GmbH.

Naciri, Y., P.-E. Du Pasquier, M. Lundberg, D. Jeanmonod, and B. Oxelman (2017). A phylogenetic circumscription of Silene sect. siphonomorpha (caryophyllaceae) in the mediterranean basin. *Taxon,* 66(1), 91–108.

Neupane, S., P. O. Lewis, S. Dessein, H. Shanks, S. Paudyal, and F. Lens (2017). Evolution of woody life form on tropical mountains in the tribe Spermacoceae (Rubiaceae). *American Journal of Botany,* 104(3), 419–438.

Palazzesi, L., J. Pellicer, V. D. Barreda, B. Loeuille, J. R. Mandel, L. Pokorny, C. M. Siniscalchi, M. C. Tellería, I. J. Leitch, and O. Hidalgo (2022). Asteraceae as a model system for evolutionary studies: from fossils to genomes. *Botanical Journal of the Linnean Society,* 200(2), 143–164.

Patino, J., R. J. Whittaker, P. A. Borges, J. M. Fernandez-Palacios, C. Ah-Peng, M. B. Araujo, S. P. Avila, P. Cardoso, J. Cornuault, E. J. de Boer, et al. (2017). A roadmap for island biology: 50 fundamental questions after 50 years of the theory of island biogeography. *Journal of Biogeography,* 44(5), 963–983.

Poppenwimer, T., I. Mayrose, and N. DeMalach (2023). Revising the global biogeography of annual and perennial plants. *Nature,* 624(7990), 109–114.

Rabosky, D. L. and E. E. Goldberg (2015). Model inadequacy and mistaken inferences of trait-dependent speciation. *Systematic Biology,* 64(2), 340–355.

Rabosky, D. L. and A. H. Hurlbert (2015). Species richness at continental scales is dominated by ecological limits. *The American Naturalist,* 185(5), 572–583.

Ram´ırez-Barahona, S., H. Sauquet, and S. Magallon (2020). The delayed and geographically heterogeneous´ diversification of flowering plant families. *Nature Ecology & Evolution,* 4(9), 1232–1238.

Rincón Barrado, M., M. Pérez, T. Villaverde, C. García-Verdugo, J. Caujape-Castells, R. Riina, and I. Sanmartín (2024). Phylogenomics and phylogeographic model testing using convolutional neural networks reveal a history of recent admixture in the Canarian Kleinia neriifolia. *Molecular Ecology,* 33(21), e17537.

Salomon, L., J. M. Gorospe, T. F´ er, F. Kolar, P. Sklenar, and R. Schmickl (2025). So young, so rich: Habitat shifts combined with trait evolution promoted species radiation in Senecio in the Andes. *Journal of Biogeography,* e15161.

Sandral, G., M. V. Remizowa, and D. D. Sokoloff (2006). A taxonomic survey of Lotus section Pedrosia (Leguminosae, Loteae). *Wulfenia,* 13, 97–192.

Sauerbier, H., F. C. Calixto, and T. Muer (2023). *Flora vascular de Canarias*. Turquesa ediciones.

Smith, S. A. and J. M. Beaulieu (2024). Ad fontes: divergence-time estimation and the age of angiosperms. *New Phytologist,* 244(3), 760–766.

Smith, S. A., J. W. Brown, and J. F. Walker (2018). So many genes, so little time: a practical approach to divergence-time estimation in the genomic era. *PloS one,* 13(5), e0197433.

Soltis, D. E., M. E. Mort, M. Latvis, E. V. Mavrodiev, B. C. O'Meara, P. S. Soltis, J. G. Burleigh, and R. Rubio de Casas (2013). Phylogenetic relationships and character evolution analysis of Saxifragales using a supermatrix approach. *American Journal of Botany,* 100(5), 916–929.

Valente, L., A. B. Phillimore, and R. S. Etienne (2018). Using molecular phylogenies in island biogeography: It's about time. *Ecography,* 41(10), 1684–1686.

Valente, L. M., A. B. Phillimore, and R. S. Etienne (2015). Equilibrium and non-equilibrium dynamics simultaneously operate in the Galapagos Islands. Ecology Letters, 18(8), 844–852.

Valente, L. M., V. Savolainen, and P. Vargas (2010). Unparalleled rates of species diversification in Europe. *Proceedings of the Royal Society B: Biological Sciences,* 277(1687), 1489–1496.

Von Humboldt, A. and A. Bonpland (1807). *Essay on the Geography of Plants*. University of Chicago Press.

Wallace, A. R. (1855). Xviii.—On the law which has regulated the introduction of new species. *Annals and Magazine of Natural History,* Volume 16, pp. 184–196. Wiley Online Library.

Warren, B. H., D. Simberloff, R. E. Ricklefs, R. Aguilee, F. L. Condamine, D. Gravel, H. Morlon, N. Mouquet, J. Rosindell, J. Casquet, et al. (2015). Islands as model systems in ecology and evolution: prospects fifty years after MacArthur-Wilson. *Ecology Letters,* 18(2), 200–217.

Weigelt, P., C. Konig, and H. Kreft (2020). Gift–a global inventory of floras and traits for macroecology and¨ biogeography. *Journal of Biogeography,* 47(1), 16–43.

Whittaker, R. J., J. M. Fernandez-Palacios, T. J. Matthews, M. K. Borregaard, and K. A. Triantis (2017). Island´ biogeography: taking the long view of nature's laboratories. *Science,* 357(6354), eaam8326.

Wiens, J. J. (2017). What explains patterns of biodiversity across the tree of life? New research is revealing the causes of the dramatic variation in species numbers across branches of the tree of life. *BioEssays,* 39(3), 1600128.

Yu, D. and J. J. Wiens (2024). The causes of species richness patterns among clades. *Proceedings of the Royal Society B,* 291(2015), 20232436.

Zhang, C. and S. Mirarab (2022). Weighting by gene tree uncertainty improves accuracy of quartet-based species trees. *Molecular Biology and Evolution,* 39(12), msac215.

Zizka, A., R. E. Onstein, R. Rozzi, P. Weigelt, H. Kreft, M. J. Steinbauer, H. Bruelheide, and F. Lens (2022). The evolution of insular woodiness. *Proceedings of the National Academy of Sciences,* 119(37), e2208629119.

Zuntini, A. R., T. Carruthers, O. Maurin, P. C. Bailey, K. Leempoel, G. E. Brewer, N. Epitawalage, E. Francoso, B. Gallego-Paramo, C. McGinnie, et al. (2024). Phylogenomics and the rise of the angiosperms. *Nature,* 629(8013), 843–850.

For the supplementary material associated with this chapter, please see the Zenodo repository accompanying this thesis (DOI 10.5281/zenodo.22093870).





Chapter 5

Synthesis



Over recent decades, a wealth of literature has been published on the core themes of this thesis - trait-dependent diversification, island biogeography, and, to a lesser extent, diversity unevenness. In this chapter, I summarise the results of the preceding chapters and frame my results in the context of the literature to draw general conclusions surrounding these themes specifically for the Canary Islands, as well as discussing what these conclusions mean globally. As trait-dependent diversification is the central theme of this thesis (Chapters 2 and 4), I will begin there, followed by the environmental correlates of plant distributions (Chapter 3), then a brief note on colonisation of the Canary Islands by non-monocot lineages (Chapter 2), a discussion on diversity unevenness (Chapter 2), before ending on unfinished side projects from this thesis and future research. The section on environmental correlates of plant distributions also discusses an inference on why the Canaries harbour the most insular woody species globally. Throughout these sections, I also briefly present and discuss additional analyses I performed that were eventually not included in the core chapters of the thesis. These include: how lineage age is correlated with diversity across all angiosperms families (Brown et al. 2023; Zuntini et al. 2024), how Canarian radiations are delineated in environmental niche space (reanalysing data from Chapter 3), and a preliminary analysis of how the growth forms of Chapter 3 vary in environmental niche space across the archipelago of Hawaii (rerunning our analyses from Chapter 3 for Hawaiian non-monocot angiosperms). First, trait-dependent diversification.

Section 1 - Trait-Dependent Diversification

We find evidence of trait-dependent diversification in Canary Islands angiosperms in two chapters of this thesis: Chapter 2 and Chapter 4. In Chapter 2, we presented a phylogenomic tree of life for Canary Islands non-monocot angiosperms, containing ~65% of the native non-monocot lineages, as well as a trait dataset for these species containing their growth form (scored in five states: annual herbaceous, perennial herbaceous, ancestral woodiness, derived woodiness, and insular woodiness; see the Introduction for definitions). We then used a dynamic, stochastic island biogeography model, DAISIE (Dynamic Assembly of Island biota through Speciation, Immigration and Extinction), to estimate rates of colonisation, speciation, and extinction for the background community, and for three bipartitions of the dataset: herbaceous versus woodiness, non-insular woodiness versus insular woodiness (IW), and annuality versus perenniality (Chapter 2, Table 1), allowing each data subset to have different evolutionary rates. To briefly summarise: non-monocot herbaceous and woody lineages have different speciation rates, with higher cladogenesis in herbaceous lineages but higher total speciation (cladogenesis plus anagen-

esis) rates in woody lineages. Woody lineages also have lower extinction and colonisation rates in the herbaceous versus woody model. In the non-monocot non-IW versus IW model, IW lineages have faster speciation rates, lower extinction and colonisation rates, and a finite carrying capacity (i.e., as the prevalence of IW lineages increases, rates of colonisation, speciation, and extinction decrease). Similarly, in the non-monocot annual versus perennial model, perennial species have faster speciation, lower extinction and colonisation rates, and no detectable carrying capacity. These results are also robust when monocots are included in our modelling (Chapter 2, Table 1). However, DAISIE is not a trait-dependent diversification model and does not account for trait evolution along the branches of a phylogeny, nor does it account for differential trait states within the same lineages (i.e., a lineage with both herbaceous and woody species). To see if these conclusions are robust, despite the shortcomings of the approach in Chapter 2, we reanalysed this dataset using a novel trait-dependent diversification model developed in collaboration with this thesis, TRAISIE (TRait-dependent Assembly of Island biota through Speciation, Immigration and Extinction).

In Chapter 4, we present TRAISIE as a novel trait-dependent diversification model specifically for insular communities. TRAISIE is distinct from other state-dependent speciation and extinction models (SSE models; see Introduction), as it does not require a single clade as input (Maddison et al. 2007; Goldberg and Igic 2012; Herrera-Alsina et al. 2019), but rather a phylogeny of an island community, following the DAISIE framework (Valente et al. 2015). This has the benefit of including also singleton lineages (island lineages with just one species) for computing the effects of traits on diversification, an often overlooked aspect of SSE models, as well as considering evolutionary dynamics replicated across independent clades to answer broader questions about the effects of traits in diversification (Anderson et al. 2023). Using the phylogenomic framework from Chapter 2, and extending the trait database to cover all native Canary Islands non-monocot angiosperms, we confirm the findings from Chapter 2: woody species diversify faster, with faster cladogenesis rates and lower extinction rates, and colonise less frequently than herbaceous species (Chapter 4, Table 2); the same pattern is found for perennial species relative to annual species; and, when considering three states (annual herbaceous, perennial herbaceous, and perennial woody), perennial woody species have the fastest speciation rates, perennial herbaceous have the lowest extinction rates, and annual herbaceous have the fastest colonisation rates (Chapter 4, Table 2). This implies that these differential evolutionary rates are evidence of trait-dependent diversification. We also find that constant rate models (i.e., those without differential rates) fit the dataset worst of all models tested (Chapter 4,

Tables S6-8), indicating constant birth-death processes are unlikely to be operating in the Canaries. Also, using simulated data, we show that TRAISIE can accurately estimate parameters and perform model selection on moderately sized datasets (of ~100 lineages, 200 species; Chapter 4, Table 1, Figures S1-17). Though we present a comprehensive trait dataset for Canary Islands angiosperms and their continental ancestors, not all traits are used here due to time and data limitations.

However, there are some important limitations and details pertaining to i) modelling constraints, ii) unexpected results, iii) data constraints, and iv) angiosperm taxonomy, underlying the results in Chapter 4 that are worth discussing before going into what differential rates mean. First, we only tested a small range of possible model options, using 11 different models, with a mix of constant rates (CR; where all parameters have the same rates), examined trait-dependence (ETD; where diversification depends on the measured traits and allows the measured traits to have different rates), and examined-concealed trait dependence (ECTD; where we combine aspects of ETD models with concealed trait-dependence, where diversification depends on hidden traits that are not scored but can still affect diversification). We also consistently found that the ECTD models we used performed worse than the ETD models; this is unexpected behaviour and likely reflects our desire to design ECTD models with relatively few parameters (to reduce computational complexity and facilitate finding the global optimum). If we ran CTD models and combined the best-fitting one with the best-fitting ETD model, then these resulting ECTD models may very well fit better than our best-fitting ETD model. A better model fit will also likely result in a better goodness of fit than those found in Chapter 4, which, despite being the best performing models, did not fit the data well. Although it is possible the poor goodness of fit results from missing data, as we show goodness of fit from TRAISIE's best fitting parameters can be sensitive to missing data (Chapter 4, Figure 6).

Second, we consistently find that anagenetic speciation rates are estimated to be near zero throughout our analyses (except in the 'Non-IW vs IW' analysis in Chapter 2, Table 1; see also Chapter 4, Table 2). Given the presence of many endemic singleton lineages (such as the Laurel forest endemics *Laurus novocanariensis* (Lauraceae), *Ilex canariensis* (Aquifoliaceae), and *Ocotea foetens* (Lauraceae)), there is evidence that anagenesis may have occurred. Yet, our results favour a scenario where these species are a result of cladogenesis followed by extinction. We speculate in Chapter 4 that this stems from the lack of recolonisations by the same species in our dataset, and our inability to identify them given our sampling strategy of sequencing a single Canarian specimen per species. With deeper sampling at the species level, we may be better able

to identify these recolonisations by the same species and therefore better estimate anagenesis rates. However, anagenesis increases with isolation of the system (Whittaker et al. 2008; Valente et al. 2020); as the Canaries are in such close proximity to continental Africa (Western Sahara), perhaps near-zero anagenesis is explained through biogeographic phenomena.

Third, data limitations pose a substantial constraint for trait-dependent diversification analyses. For example, for TRAISIE, not only is a phylogeny and corresponding trait dataset for the island population required, but also a trait dataset for the closest continental ancestors of all lineages in the phylogeny (to assign to the root state in the phylogeny input for TRAISIE), and knowledge of trait distribution across the mainland pool (i.e., the continental source population that can colonise the island in a DAISIE framework). For the Canaries, we (conservatively) estimated the mainland pool to be non-monocot angiosperms of the Mediterranean Basin, totalling ~20,000 species, neglecting the limited number of long-distance dispersal events from the Americas (such as in *Persea* (Lauraceae) and *Visnea* (Pentaphyllaceae); Chapter 2, Figure 2). As trait distribution across the mainland pool affects the likelihood calculation in TRAISIE, and therefore model choice, poor or skewed data coverage for the mainland pool can affect our inference on trait-dependent diversification. Thus, we opted only to use traits with good coverage in the mainland pool (>80% coverage). Even for commonly scored traits, such as maximum plant height, data coverage in the mainland pool was only ~31%, a percentage of missing data that could affect model choice. With larger mainland pool sizes, data coverage would be even lower, making many traits unsuitable for such analyses. Subsequently, it may not be realistic to use TRAISIE to analyse many traits potentially interesting for diversification (such as flower/fruit-related traits) without serious investment in trait scoring for island species, their closest continental ancestors, and their mainland pools. This also does not consider the phylogeny size and sampling fraction required for diversification analyses to produce accurate results (recommended >300 species, > 60% coverage; Davis et al. 2013; Mynard et al. 2023).

Fourth, as in all diversification studies, the results of trait-dependent diversification analyses are reliant upon solid taxonomic foundations of the species included. Taxonomic uncertainty (i.e., unclear delimitation of species boundaries; Meyer et al. 2026) can result in 'over-splitting, the description of too many taxa (Pérez-Vargas et al. 2020), which can artificially inflate species richness and diversification rates within a clade (Frateles et al. 2025; Meyer et al. 2026); equally, 'lumping' of species (grouping closely related entities into fewer taxa; Meyer et al. 2026) together can reduce diversification rates. Over-splitting has been partially addressed in several Canarian angiosperm radiation (such as *Pericallis* (Jones et al. 2016), and *Lotus* (Pérez-Vargas et al. 2020)), and species

boundaries appear robust. Many radiations in the Canaries are classic examples of adaptive radiations (Böhle et al. 1996; Francisco-Ortega et al. 1996; Mort et al. 2007), but if there is lumping in these lineages, our inferences here may be affected.

But why do we find differential evolutionary rates between trait states? I conjecture that woodiness (representing a considerable part of perenniality) confers a beneficial adaptation in Canarian non-monocot angiosperms, buffering against seasonality and to a lesser extent, drought (see Section 3, 'Environmental Correlates of Insular Plant Distributions' below for further discussion), i.e., there is an evolutionary benefit to having this trait state within the given climate (for example, how deciduousness in woody angiosperms is beneficial in temperate climates; Zanne et al. 2014; Wiens 2023). My inference here is that these traits allow angiosperms species to persist for longer, thereby, extinction rates may be lower, and therefore increasing diversification rate. Though our trait-dependent diversification analyses in Chapters 2 and 4 suggest an advantageous nature of these trait states in the Canaries, the increased proportion of woody to herbaceous species in the Canaries relative to the mainland ancestors also confers an advantage (Synthesis, Table 1). There are still more native Canarian non-monocot herbaceous species than woody (explained by high colonisation rates; Chapter 2, Table 1; see also Lens et al. 2013), but where only ~18% of the continental ancestors to Canarian lineages are woody (Synthesis, Table 1; Chapter 4, Supplementary Data S3), ~40% of Canarian species are woody (Synthesis, Table 1; Chapter 4, Supplementary Data S2). The same pattern is found when considering annuality versus perenniality: 52% of the mainland ancestors to native Canarian non-monocot angiosperms are annual, whereas only 32% of native Canarian non-monocots are annual. Note that this conjecture does not apply to all of Macaronesia; most diversifying clades in the Azores are herbaceous, in contrast to the rest of Macaronesia, where they are mostly woody (Price et al. 2018). So, in summary, though there are no universal drivers of diversification in angiosperms (Helmstetter et al. 2023), we consistently find woodiness and perenniality as potential drivers of diversification at the community level for Canary Islands non-monocot angiosperms.

Section 2 - Environmental Correlates of Insular Plant Distributions

As we found growth form significantly shaped the Canarian non-monocot flora in Chapter 2, in Chapter 3, we looked deeper into the environmental correlates and distribution patterns of these growth forms. This had the additional benefit of providing us with environmental variables that could be used in trait-dependent diversification modelling, if we wished. In Chapter 3, we used a novel bioinformatics pipeline to reconstruct environmental niche space for

Synthesis Table 1. The number of Canarian non-monocot angiosperm species, and the number of continental ancestral species to the native Canarian non-monocot lineages, with each trait state. Data for Canarian species is from Chapter 4, Supplementary Data S2, and data for the closest continental ancestors is from Chapter 4, Supplementary Data S3.

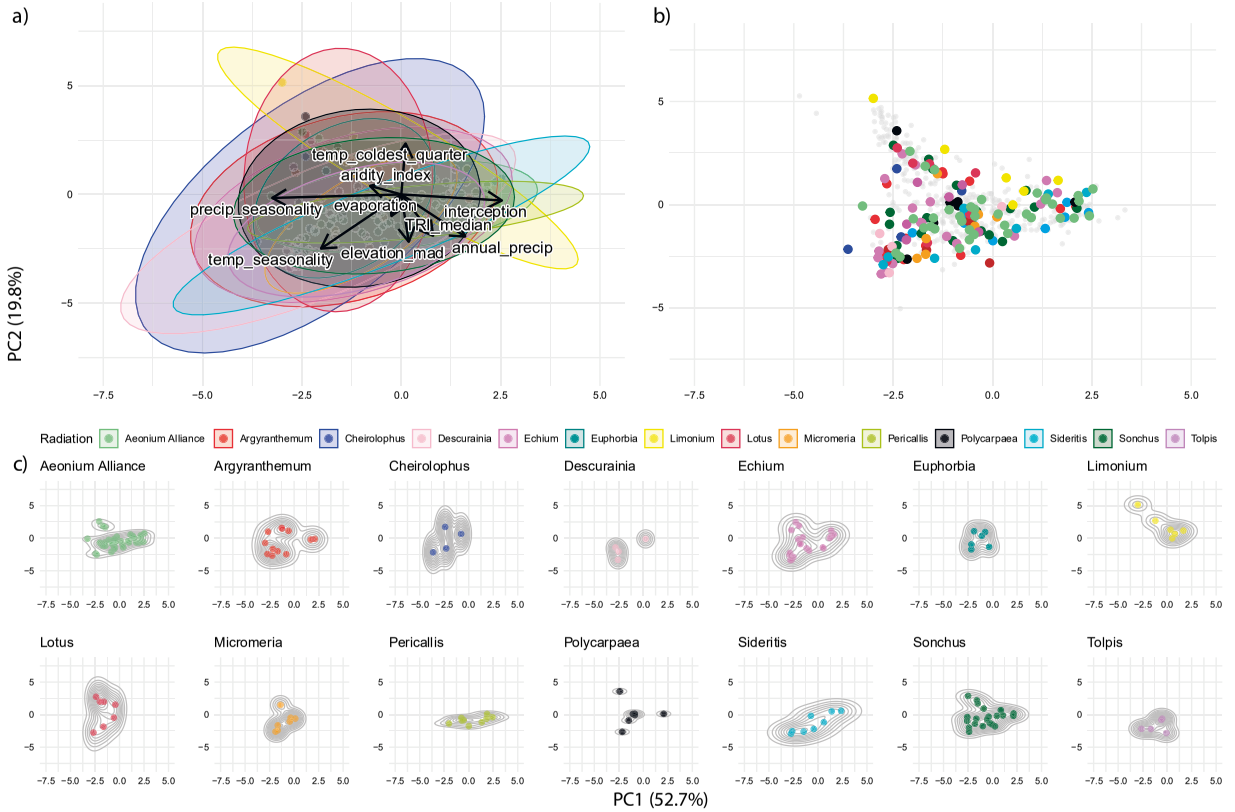
	Continental Non-Monocot Ancestors		Canarian Non-Monocot Species	
Trait State	Number of Species	Percentage of Total	Number of Species	Percentage of Total
Herbaceous	542	82%	690	60%
Woody	120	18%	465	40%
Annual	344	52%	370	32%
Perennial	318	49%	785	68%

548 Canarian non-monocot angiosperms using ten environmental variables and publicly available spatial data, whilst accounting for phylogeny (Chapter 3, Figure 2a). We used a phylo-PCA to estimate the niche of five growth forms (annual herbaceous (AH), perennial herbaceous (PH), insular woodiness (IW), derived woodiness (DW), and ancestral woodiness (AW)), given by confidence ellipses, and mapped the loadings of the PCA to reveal the variables important for shaping niche space for each growth form. We then used phylo-ANOVAs to reveal significant pairwise differences between the growth forms, for each variable (Chapter 3, Figures 3a-j). By combining the results of these analyses, we identified seasonality in both precipitation and temperature (or lack thereof), and rugosity as the most important variables associated with the occurrence of these growth forms.

We found that AH species in the Canary Islands are associated with arid regions of low precipitation and elevation, and consistently high temperatures (Chapter 3, Figures 2a, d, 3c, d, Table S3). Perennial herbaceous species show similar correlates, but with greater importance of increasing precipitation and elevation, and milder temperatures (Chapter 3, Figures 2a, S6), as do DW species, but with lower elevation and precipitation (Chapter 3, Figures 2a, 3a, e). In contrast, we found that AW species occur in areas of the Canaries with high precipitation, high topographic complexity (rugosity), and low seasonality (Chapter 3, Figures 3c, e, g). Insular woody species, unexpectedly, are found in regions of high temperature seasonality, lower temperature in the coldest quarter (almost never reaching frost), and high topographic complexity (Chapter 3, Figures 3b, c, d), with drought-related variables not recovered as significant. Our finding of rugosity being one of the most important distinguishing factors for plant distributions on the Canaries sheds new light on the biogeography of the Canarian angiosperm flora. Rugosity had previously been found to correlate with alpha and beta diversity of Gran Canarian endemic vascular plant species (J. Caujape-Castells 2025, pers. comm.), but a link with growth forms had yet to be tested.

Although our findings on the correlates of insular woody species initially appear to contrast with those of previous studies (i.e., we do not find drought-related traits correlate with the occurrence of IW species; Lens et al. 2013; Hooft van Huysduynen et al. 2021; Zizka et al. 2022), they do not contrast as substantially as they may appear. As discussed in Chapter 3, though we identify temperature seasonality as an important correlate for insular woody species occurrence (Chapter 3, Figure 3c), as well as the lowest temperature in the coldest quarter (Chapter 3, Figure 3d), these findings do not directly contradict previous findings. A growth season uninterrupted by frost (i.e., an aseasonal climate as defined by Carlquist 1974) is present in the Canaries despite the recovered preference for colder temperatures in the coldest quarter. The only habitat on the Canaries where insular woody species experience frost is the alpine desert region above 2,200 m altitude on the flanks of the Teide volcano on Tenerife. Additionally, we recover precipitation seasonality as important in shaping niche space (Chapter 3, Figure 2a), a finding in agreement with Zizka et al. (2022), who found a positive effect of precipitation seasonality on the prevalence of insular woody species on oceanic islands alone and when considering oceanic and continental islands together. Interestingly, Zizka et al. (2022) infer this positive effect, in addition to other drought-associated variables that were correlated to the occurrence of insular woody species on islands, to be evidence of drought driving the prevalence of insular woodiness - a conclusion not drawn in Chapter 3. If we had drawn the same conclusions in Chapter 3, then we would have found evidence of insular woodiness being an adaptation to drought. Thus, despite the initial contrast, Chapter 3 aligns more closely with previous studies than at first glance.

Chapter 3 constitutes a comprehensive analysis of growth forms and their environmental correlates in the Canary Islands; however, several areas related to the correlation between growth forms and environmental variables remain unexplored. First, in the discussion section of Chapter 3, we draw reference to how insular woodiness is present in most evolutionary radiations on the Canaries (Chapter 2; Carine et al. 2010; Lens et al. 2013), and how this may cause the correlation with rugosity. Yet, we did not further investigate the link between insular woodiness and radiations (defined here as lineages with at least 5 native Canarian species). Here, I further explore how radiations (including both insular woody and other types) are delineated in trait space. As our phylo-PCA contains data for 15 of the 20 most species-rich lineages in the Canaries, environmental space can be reanalysed to delimit the niches of these radiations, rather than growth forms, to see if radiations can be distinguished in environmental space (Synthesis, Figure 1). Please note, the remaining five radiations not sampled here were due to a lack of data (Chapter 3, Figure 2, Table S3), and due to



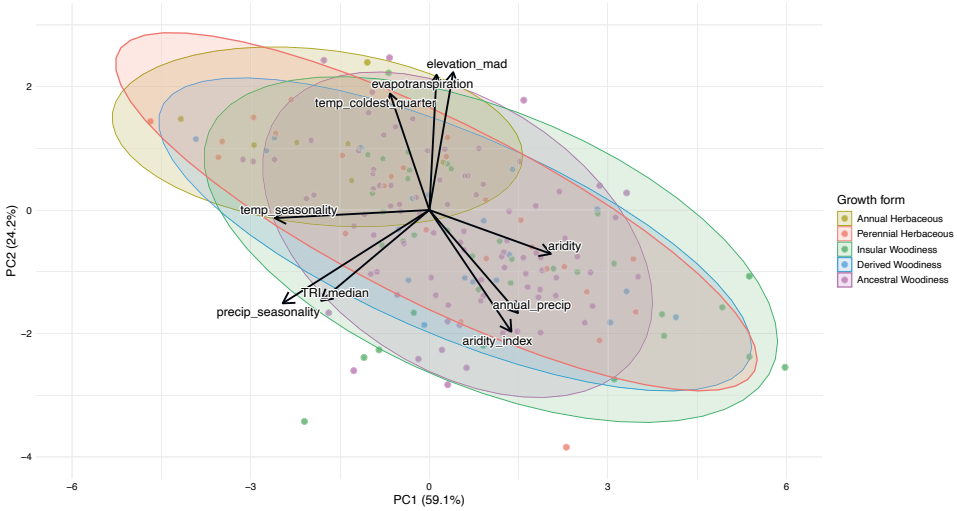
Synthesis Figure 1. Environmental niche space of the radiations of the Canary Island non-monocot angiosperm flora. Results are shown for 15 of the 20 most species-rich lineages in the Canary Islands. Data from Chapter 3, Figure 2, reorganised to a) show the environmental niche space for all 15 radiations combined, b) highlight the spread of radiations across the total environmental niche space (where grey dots are native non-monocot angiosperm species not from radiations), and c) show environmental niche space separately per radiation. Contour lines represent 2D kernel density estimation.

the small number of species in several of these radiations, phylo-ANOVAs were not run here to assess significance. Nevertheless, some clear trends emerge. The majority of radiations occupy a broad range of environmental space, suggesting that niche lability in these radiations has allowed them to colonise the diverse environmental conditions present in the Canaries. In particular, species of *Polycarpaea* occupy distinct regions of niche space, with the least overlap between species within this radiation. Across all radiations, there is substantial overlap in niche space centrally, suggesting the climatic variables included in the analysis are insufficient to separate these lineages and their niches. *Argyranthemum*, *Sonchus*, *Echium*, and *Descurainia* all overlap almost entirely, highlighting how these lineages have radiated throughout multiple environmental zones of the Canaries (also overlapping with *Euphorbia*, *Micromeria*, and *Tolpis,* which occur in much more conserved regions of niche space). As with dos Santos et al. (2022), *Aeonium* species have radiated throughout niche space, filling all regions except the most arid and highest elevations. *Sideritis, Pericallis,* and *Limonium* all diversify along a single axis, indicating their niche is expanding by diverging into similar but slightly different environments, or remains unchanged, and species are diversifying narrowly within existing niche space. Temperature seasonality shapes niche space most in *Sideritis* (given by the PCA axes in Synthesis, Figure 1), precipitation seasonality and interception in *Pericallis*, and annual precipitation in *Limonium*. Certain members of radiations also contribute to expanding niche space within a radiation, being found at the fringes, such as in certain species of *Limonium, Cheirolophus,* and *Lotus,* in agreement with Schrader et al. (2021). Taken together, these results suggest that environmental niche lability plays a major role in radiations across the Canary Islands (Steinbauer et al. 2016), with many radiations found to occupy a broad range of niche space, whilst others radiated in a more conserved niche space.

Second, and outside of the scope of the chapter, a pipeline to automate the analyses therein is presented, but no comparisons to other systems are made. Using existing datasets, here I would like to demonstrate that the pipeline we developed can be easily run for other insular communities. Here, I used several published datasets to reconstruct the same environmental niche space and the violin plots as produced in Chapter 3, but instead of the Canary Islands, I used data on the second insular woodiness hotspot (in terms of insular woody species and number of transitions), which is markedly different from the Canarian archipelago in isolation and climatic variables - the Hawaiian archipelago (National Atlas of the United States 2013; Fortini et al. 2022; Zizka et al. 2022, 2025; Zomer and Trabucco 2022; Brown et al. 2023; Poppenwimer et al. 2023; Roeble et al. 2024). First, I generated a list of native non-monocot spe-

cies of the Hawaiian archipelago (Brown et al. 2023; Roeble et al. 2024), then cross-referenced this species list with growth form datasets to assign annuality vs perenniality using Poppenwimer et al. (2023), with herbaceous vs woody data from Brown et al. (2023) and Poppenwimer et al. (2023). For woody species, I cross-referenced the species list with the insular woody species list of Zizka et al. (2022) and the derived woodiness dataset of Zizka et al. (2025). Any gaps were then checked manually (such as woody species not assigned as either insular or derived woody, to see if they belonged to a radiation to assign the same growth form as the rest of their radiation), and ancestral woodiness was assigned to woody species not belonging to a radiation of insular/derived woody species. I then checked all assigned growth forms using published literature (i.e., taxonomic revisions, protologues, journal articles with data on growth form or pictures of species). Though this approach is coarse, the published datasets were sufficiently complete to require few manual checks, and thus I considered them to be robust enough to provide a comparison with the Canary Island dataset. I then ran the pipeline from Chapter 3 to generate a phylo-PCA for the Hawaiian non-monocot community and the accompanying violin plots (Synthesis, Figures 2-4), using the bioclim variables of Fortini et al. (2022), with a resolution of ~250m, and the 'sliced' digital elevation model (DEM) of the National Atlas of the United States (2013). In the sliced DEM, pixels in the raster file correspond to discrete elevational bands (from 1-30, corresponding to the 0-4,207m present on Hawaii). Elevation and rugosity will subsequently be affected by this difference.

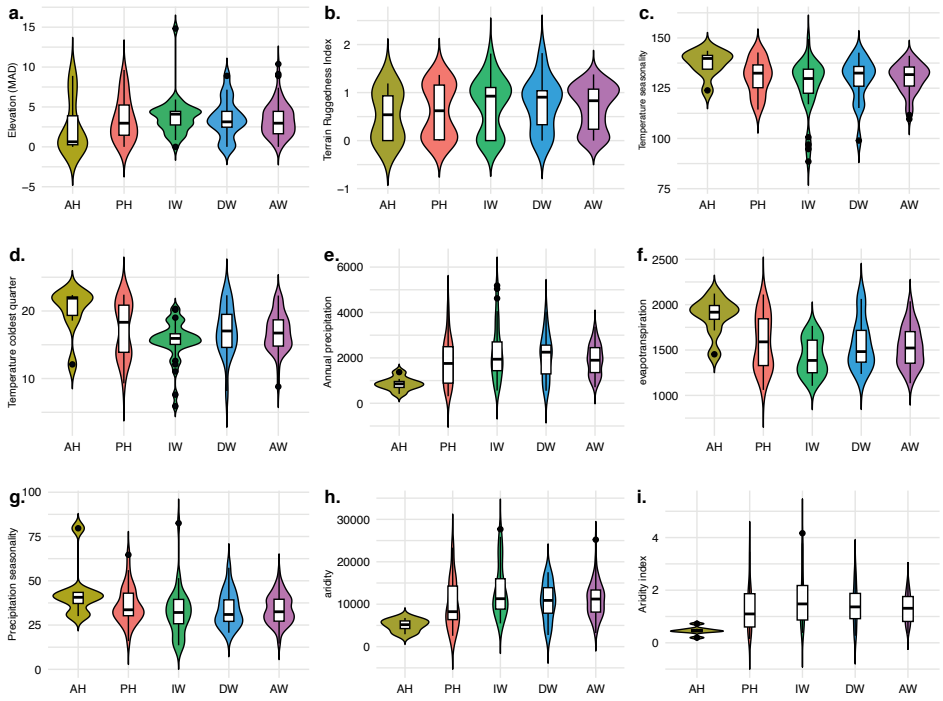
Using a coordinate uncertainty of 500m and 13 minimum occurrences (i.e., the balanced sampling approach from Chapter 3), 181 of 949 Hawaiian non-monocot angiosperm species remained after filtering (~20%) and were plotted in the phylo-PCA (Synthesis, Figure 2). Climatic seasonality (in both temperature and precipitation) and elevation (median absolute deviation) are recovered as important in shaping the environmental niche space in Hawaii (Synthesis, Figures 2, 3). Rugosity, contrary to the Canaries, contributes only marginally to niche space; however, this may be due to the scale of the DEM used. For how each variable contributes to the PC axis, see Synthesis, Figure 4. The recovered environmental variables for Hawaii show clear differences in the occurrences of insular woody species between the archipelagos (Synthesis, Figure 3). Where Canarian insular woody species were strongly associated with areas of high median temperature seasonality, Hawaiian insular woody species occur in areas with the lowest mean temperature seasonality, but at the same time, have the widest range. This suggests that, although on average temperature seasonality is low in areas where Hawaiian insular woody species occur, there are still some species that occur in areas of high seasonal-



Synthesis Figure 2. Environmental niche space of the growth forms of the Hawaiian archipelago non-monocot angiosperm flora. Phylogenetic PCA of the different environmental variables labelled per growth form for 181 of the 949 Hawaiian non-monocot angiosperm species. Each point represents a species, shaded ellipses represent 95% confidence intervals around the points for each growth form, and the arrows represent the loadings of the phylo-PCA to highlight the contribution of each variable. Colours are as in Chapter 2, Figure 2.

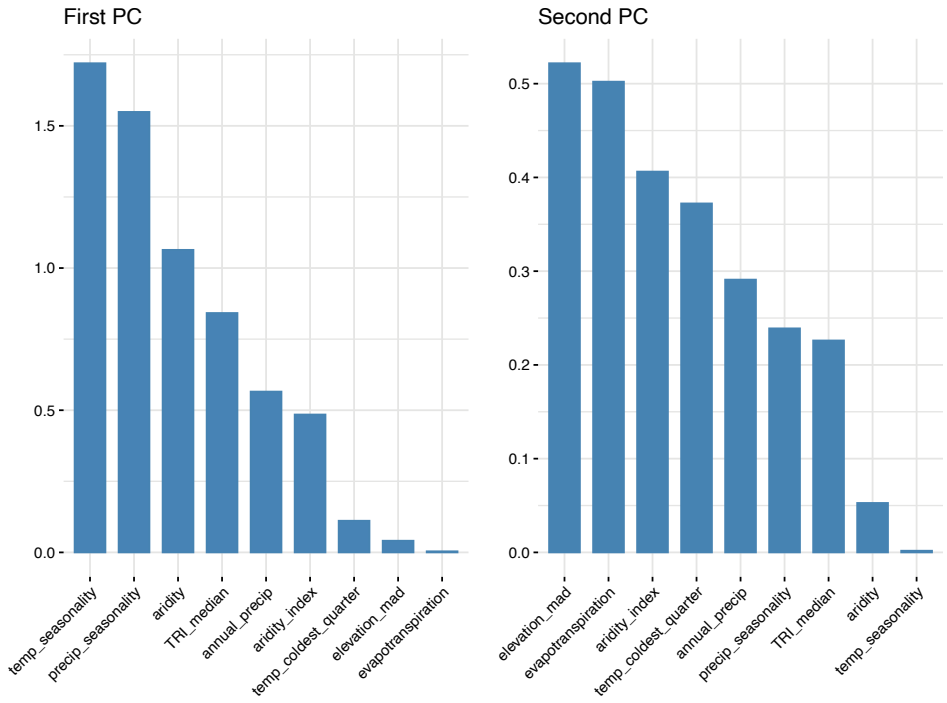
ity. Hawaiian insular woody species also occur in areas of low precipitation seasonality, suggesting that aseasonal climatic conditions are more correlated with the distribution of insular woodiness in Hawaii than in the Canaries. Also, in contrast to the Canaries, aridity is found to be strongly correlated with the occurrence of Hawaiian insular woody species, supporting the hypothesis suggesting that drought is an important driver of insular woodiness. In addition to the differences, there are also some consistencies with the Canaries. Across both archipelagos, we see a large overlap in niche space between the growth forms. Additionally, insular woody species occur in regions with high rugosity and low temperature in the coldest quarter, again without frost temperatures prevailing except at the highest mountain peaks. Interestingly, Hawaiian insular woody species occur in regions with the widest range of annual precipitation, and across both archipelagos, the second-highest median, indicating a wide precipitation tolerance. However, as only 20% of the total Hawaiian species were retained after filtering (e.g., ~16% of IW species were retained), these conclusions should be considered preliminary and may change with the inclusion of more data.

By comparing the results of Hawaii to those of Chapter 3, we find that environmental niche space is primarily driven by similar factors: an interplay of climatic seasonality and island topography, though differences in how topogra-

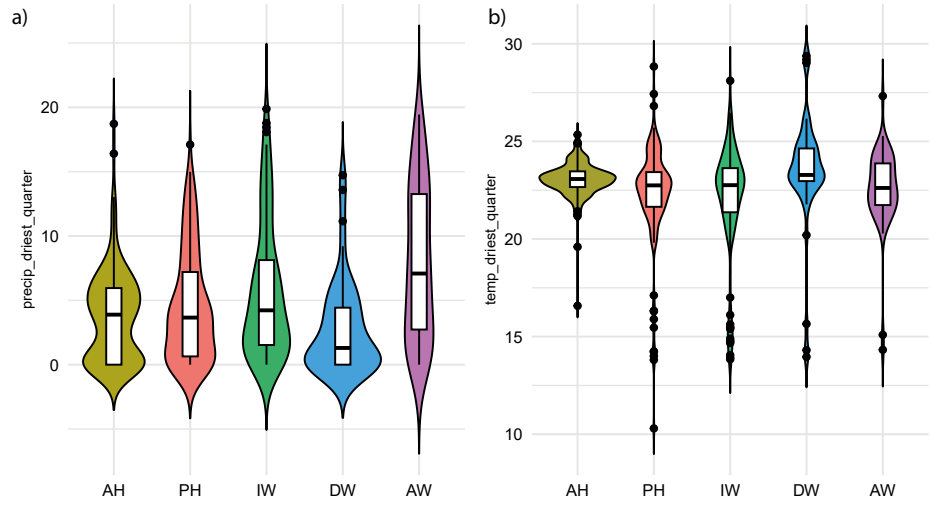


Synthesis Figure 3. Violin plots of the distribution of different environmental variables across each growth form for the Hawaiian non-monocot angiosperm flora. Boxplots show the distribution of different environmental variables across each growth form: a) Elevation (MAD), b) Terrain ruggedness index, c) Temperature seasonality (°C*100), d) Temperature in the coldest quarter (°C), e) Annual precipitation (mm), f) evapotranspiration (mm day-1), g) Precipitation seasonality (%), h) aridity (from Global Aridity Index raster, calculated as (mean annual precipitation / Mean Annual Reference Evapo-Transpiration) * 10,000), and i) Adjusted Thornthwaite aridity index. The violin plots show the distribution of each variable per growth form and the median value for each environmental variable for each growth form. AH = annual herbaceousness, PH = perennial herbaceousness, IW = insular woodiness, DW = derived woodiness, AW = ancestral woodiness.

phy affects environmental space are recovered. Interestingly, the environmental correlates between the growth forms are similar: low elevation, low precipitation, and high temperatures in annual herbaceous species; low seasonality and wet conditions in ancestral woody species; and high rugosity, low temperature in the coldest quarter (without frost), and low precipitation seasonality in insular woody species. However, each of the archipelagos shows some idiosyncrasies, likely reflecting the distinct biogeographic properties of the systems. In light of the findings from this thesis, insular woodiness appears to be an adaptation to a variable climate, allowing IW species to persist in a wide range of climates, likely a response to the climatic gradients and seasonality across the rugose areas in which they inhabit (Graham et al. 2025). Drought-related variables - high evaporation (in the form of interception, Chapter 3) and



Synthesis Figure 4. Squared cosine contribution to principal component axes, showing the quality of representation in the Hawaiian non-monocot angiosperm phylo-PCA.



Synthesis Figure 5. Violin plots of a) precipitation in the driest quarter, and b) temperature in the driest quarter for Canarian non-monocot angiosperms. AH = annual herbaceousness, PH = perennial herbaceousness, IW = insular woodiness, DW = derived woodiness, AW = ancestral woodiness.

low precipitation seasonality (Chapter 3, Figures 3f, g), and a low, non-freezing temperature in the coldest quarter are also contributing factors. This is the climatological benefit woodiness (thereby perenniality) confers against climatic variability I allude to in the trait-dependent diversification section above.

Finally, the results of Chapter 3 allow me to speculate on why the Canary Islands harbour the highest insular woody species diversity globally (Zizka et al. 2022). Poppenwimer et al. (2023) find that the prevalence of annual herbaceous species is driven by low precipitation and high temperatures in the driest quarter, likely explaining why annuals dominate in Mediterranean climates. By plotting global BioClim variables for temperature and precipitation in the lowest quarter (Fick and Hijmans 2017), the Canaries are both cooler and wetter than the southern Mediterranean Basin. I conjecture that this difference in climatic conditions (increased precipitation and lower temperatures in the driest quarter in the Canaries relative to the Mediterranean Basin, from which most colonisations occur) may have contributed to the evolution of IW, and perhaps perenniality, in the Canaries, potentially linked to a change in adaptive strategy away from the optimal strategy for survival in the hotter and drier Mediterranean Basin (see Heo et al. 2023; Zizka et al. 2025). This is supported by insular woody species occurring in cooler and wetter regions in the driest quarter than annual herbaceous species on the Canaries (Synthesis, Figure 5, adapted from Chapter 3, Table S3), and insular climates that experience less temperature and drought extremes, allowing these insular species to grow longer, thereby increasing their chance to transition from short-living herbaceous species into longer-lived insular woody species. Therefore, I propose that the coalescence of multiple factors could explain why the Canary Islands harbour the most insular woody species globally: i) the fact that the Mediterranean Basin has some of the highest proportions of annuals globally (Poppenwimer et al. 2023), resulting in a large source population; ii) the close proximity of the Canaries to the continent, which may have facilitated frequent colonisations of herbaceous species in the Canary Islands (see also Chapter 3, Table 1); iii) a shift in adaptive strategy away from survival in the hotter, drier Mediterranean Basin to increased precipitation and lower temperatures in the driest quarter in the Canaries; and iv) the year-round frost-free climate insular species inhabit, combined could have facilitated the evolution, survival and diversification of insular woody species (Chapter 3, Figure 3d).

Section 3 - Colonisation of the Canary Islands by non-monocot lineages

In Chapter 2, we inferred colonisation times of the Canary Islands for 362 Canarian angiosperm lineages using phylogenomic data, representing 751 species (or 65% of the non-monocot flora) and addressed the biogeographic history

of the extant flora using these updated colonisation times (Chapter 2, Figures 2, 3). We estimate the mean stem age for all 362 lineages to be 6.6 ± 0.5 million years (My) and the mean crown age to be 4.4 ± 0.4 My. These estimates are older than previously published values of 5.5 ± 1.3 My and 3.0 ± 1.0 My, respectively, for the mean stem and crown ages (based on 73 lineages, where our estimates for these 73 lineages are 7.8 ± 1.3 and 4.2 ± 1.2 My, respectively; Supplementary Data S13 and references cited therein). Although we consistently recovered older colonisation times than existing estimates, the inferred biogeographic history of the Canaries does not change substantially with these older estimates. All lineages, except for *Astydamia latifolia*, had stem ages that fall within the estimated age of currently emerged islands (23 My; Campeny et al. 2024). The mean lineage age we estimated overlaps with the late Miocene aridification and intensification of drought in the Western Mediterranean, Northern Africa, and Macaronesia, as previously reported (García-Maroto et al. 2009; Jiménez-Moreno et al. 2010; García-Verdugo et al. 2019a; Hooft van Huysduynen et al. 2021). Finally, our results agree with previous studies that the Canarian laurisilva is not a relic from the tertiary period (Schußler 2020), and the thermophilous forest did not originate after the onset of the Mediterranean climate 2.8 My ago (Martin-Hernanz et al. 2023), reflected in continuous colonisation times of the major taxonomic components of these communities since the late Miocene (Chapter 2, Figure 3).

We also identify differential mean stem and crown ages for insular woody, ancestral and derived woody lineages combined, and herbaceous lineages (Chapter 2, Figure 3). Insular woody lineages were the oldest, with a mean stem and crown ages of 7.7 ± 1.6 and 4.6 ± 1.3 My, respectively (Chapter 2, Figure 3a); then ancestral plus derived woody lineages, at 7.6 ± 1.0 and 5.4 ± 0.9 My (Chapter 2, Figure 3c); and herbaceous lineages were the youngest at 5.9 ± 0.6 and 3.9 ± 0.5 My (Chapter 2, Figure 3b). We hypothesise that these differences in ages reflect differential colonisation abilities of the growth forms. The younger ages of herbaceous lineages are likely explained by the propensity of herbaceous species for long-distance dispersal, combined with a shift towards open vegetation in the Western Mediterranean during the Pliocene (Jiménez-Moreno et al. 2010, 2013), which favoured the presence of herbaceous species. This hypothesis is further supported by the DAISIE analysis (Chapter 2, Table 1), and the TRAISIE analysis (Chapter 4, Table 2), where herbaceous lineages have a colonisation rate an order of magnitude higher than woody species (where woody species here combine insular woody, derived woody, and ancestrally woody lineages into one category). Differences in diversity under the best fitting models were due to this differential colonisation rate, as well as a reduced extinction rate in woody lineages, rather than differences in speciation rates using DAISIE, with

a higher diversification rate found in woody lineages using TRAISIE. These findings contribute to the field of functional island biogeography (Schrader et al. 2021; Schrader 2025), by providing empirical evidence that functional traits, namely growth form and life span, affect community composition (Chapter 2, Figure 3), dispersal rates (H2 in Schrader et al. 2021), and extinction rates (H12 in Schrader et al. 2021). Incorporating functional traits into the field of island biogeography can therefore improve our biogeographic inferences.

Section 4 - Diversity Unevenness

As I presented in the introduction of this thesis, there are three main (non-mutually exclusive) hypotheses to explain why some branches on the tree of life have more species than others. These are: clade age, variable diversification rates between clades (see Synthesis, Section 1), and ecological limits to diversification (see Chapter 1, Figure 1). In Chapter 2, we explored how these three hypotheses manifest to give rise to the Canarian non-monocot angiosperm flora. Briefly, we found no evidence of clade age affecting diversity, with no significant relationship between clade age (colonisation time) and species richness (Chapter 2, Figures 3 and 4). However, there was strong evidence of trait-dependent diversification, with woody species (at the time of colonisation), insular woody species, and perennial species exhibiting faster speciation and lower extinction rates than the background community (Chapter 2, Table 1; see also Chapter 4). Furthermore, we provided evidence of ecological limits to diversity, with insular woody species have a detectable carrying capacity (meaning that as the number of species increases, rates of speciation and colonisation decrease), and the flora having reached macroevolutionary equilibrium (Chapter 2, Table 1, Figure S2).

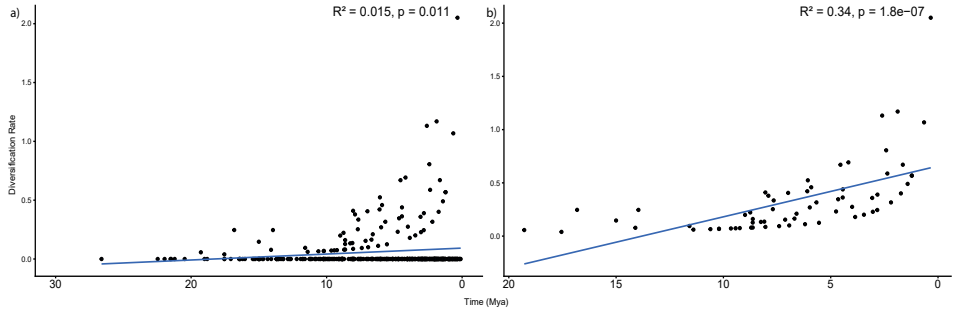
There is little consensus in the literature on which of these hypotheses best explains global diversity unevenness across clades (see Yu and Wiens 2024 for a summary). But from recent work, some general patterns emerge. Some of these patterns are notably interrelated, and thus it is difficult to tease apart different contributions. Below, I break down these patterns for the Canarian non-monocot angiosperm flora, comparing them with other published studies. Not all references in this section pertain to angiosperms, but I include them to allow a broader discussion on a global scale across the tree of life.

First, age has generally been found to affect species richness (Dimitrov et al. 2023; Yu and Wiens 2024; Quintero et al. 2025; Stephens et al. 2025). This contrasts with our findings (Chapter 2), where age does not affect species richness in the Canaries, robust to multiple ways of treating the data. However, if we include the *Aeonium* Alliance stem age in our analyses (which was removed as no close continental relatives were successfully sequenced, resulting in an

overestimated clade age), there is a significant relationship between clade age and diversity using a phylogenetic generalised least squares regression model. Though we sequenced the majority of the non-monocot lineages native to the Canaries, for three radiations (*Aeonium, Micromeria,* and the *Bencomia* Alliance), colonisation time could not be accurately estimated as no close continental relatives were successfully sequenced. Therefore, with a more accurate estimate of the ages of these clades, it is possible that we would have found a relationship between clade age and diversity. As the *Aeonium* Alliance is the most species-rich lineage on the Canaries, and indeed in Macaronesia, perhaps its removal is disproportionately affecting our conclusion. However, the *Aeonium* alliance was included in the crown age dataset (Chapter 2, Figure 4b), and in that case, there was still no relationship between lineage age and diversity.

Second, variable diversification rates are ubiquitous across the tree of life, and recent work at the global scale has found that older clades tend to have more species and reduced diversification rates, while younger lineages tend to have fewer species but higher diversification rates (Scholl and Wiens 2016; Nge and Skeels 2025; Quintero et al. 2025; Wiens and Moens 2025). To assess whether there is a relationship between clade age and diversification rates, I calculated diversification rates for all Canarian clades in Chapter 2, Figure 4 (using Magallon and Sanderson 2001), and used a linear regression to assess significance (Synthesis, Figure 6). When considering all clades (Synthesis, Figure 6a), and only clades that have diversified into at least two extant species (Synthesis, Figure 6b), I find a significant relationship between clade age and diversification rate for both scenarios, with younger clades exhibiting higher diversification rates. This supports a growing body of evidence of a relationship between evolutionary rates and clade age, with diversification rates often being found to be lower for older clades (Scholl and Wiens 2016; Nge and Skeels 2025; Quintero et al. 2025; Wiens and Moens 2025), termed the 'Zarathustran hypothesis' by Quintero et al. (2025). Therefore, although there is no relationship between clade age and diversity (species richness) in the Canaries, there is a relationship between clade age and diversification rates (Synthesis, Figure 6). However, it is worth noting that this correlation may be driven by rate scaling issues, caused by covariation between diversification rates and time (due to both variables being time-dependent), sampling bias, or survivorship bias (Harmon et al. 2021 and Louca et al. 2022).

Third, the relationship between clade age, diversification rates and species richness has been found to generally become stronger with age. This has been shown in studies that binned lineages into groups of 10 My, and then examined the correlation between diversification rates and species richness for each age category (Dimitrov et al. 2023; Stephens et al. 2025). The studies found that



Synthesis Figure 6. The relationship between lineage diversification rates and stem age estimates for Canarian non-monocots for a) all lineages, and b) only including lineages that have diversified in situ (two or more extant species), where each point represents a native Canarian non-monocot angiosperm lineage from Chapter 2.

older clades increasingly have more species and slower diversification rates. Lineage diversity in angiosperm genera and Squamates has been particularly linked to the effect of deep time, with the relationship between diversification rates and speciation rates only found to be significant after 20 My and 40 My, respectively (Dimitrov et al. 2023; Stephens et al. 2025). When I repeat this analysis for the Canary Island non-monocot angiosperm lineages from Chapter 2, Figure 4 (binning into groups of 5 mya due to a younger system age), all age categories show a significant correlation between species richness and diversification rate (Synthesis, Table 2), and the oldest bins do not have the most species, indicating that older lineages contribute less to extant diversity in the Canaries. The reasons for the disparity between my results from the Canary Islands with those of Stephens et al. (2025) for Squamates are complex, but could be due to sampling at both a low geographic scale (using only a single archipelago vs global), continental vs island systems, an excessive effect of extinction in the Canary Islands which disproportionately removed species from older clades, or insufficient sampling density (as this analysis includes 53% of non-monocot angiosperm lineages).

To compare Canarian trends with those of angiosperms, and shed some light on whether sampling issues affect these conclusions, I reran Chapter 2, Figure 4; Synthesis, Figure 6; and Synthesis, Table 2 using angiosperm-wide data from Brown et al. (2023) and Zuntini et al. (2024). Briefly, I used family age estimates from the 'young tree' of Zuntini et al. (2024), and these families with age estimates were cross-referenced with the angiosperm species richness data of Brown et al. (2023). This dataset spanned 64 orders and 352 angiosperm families, representing 350,357 species and over 150 My of angiosperm evolution. A significant negative effect of age on clade diversity is found at both the order and family levels (Synthesis, Figure 7), for both stem and crown age estimates.

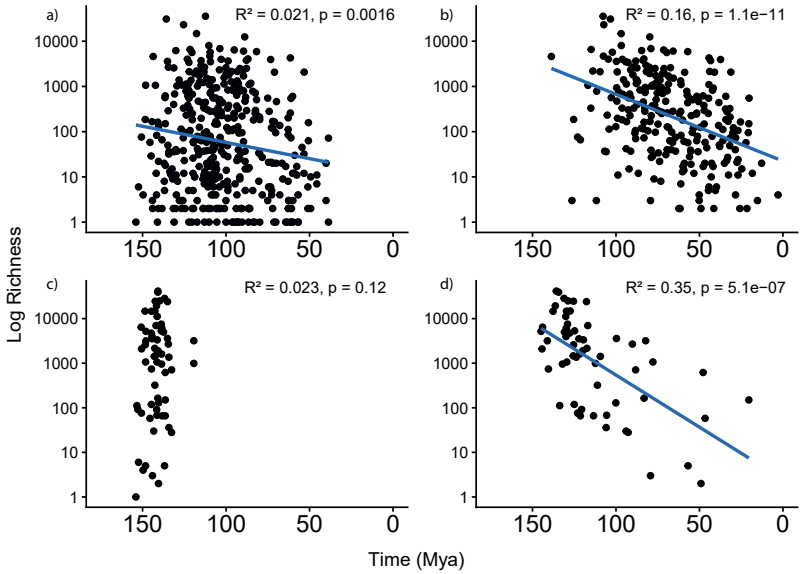
Synthesis Table 2. Relationship between species richness and diversification rate per clade age category for Canary Island non-monocots. Each native Canarian non-monocot angiosperm lineage was binned into an age category based on stem age. For each age category, the correlation between diversification rate (per lineage) and lineage species richness was computed, following Stephens et al. (2025). Adj. R^2 - adjusted R-square, p - *p*-value of the correlation.

Age (My)	Total Species Richness in category	Number of lineages in category	Adj. R^2	p
< 5	284	177	0.352	< 2.2e-16
5 -9	283	109	0.821	< 2.2e-16
10 -14	86	48	0.751	1E-15
15 -20	94	21	0.792	3.9E-08
20+	7	7	NA	NA

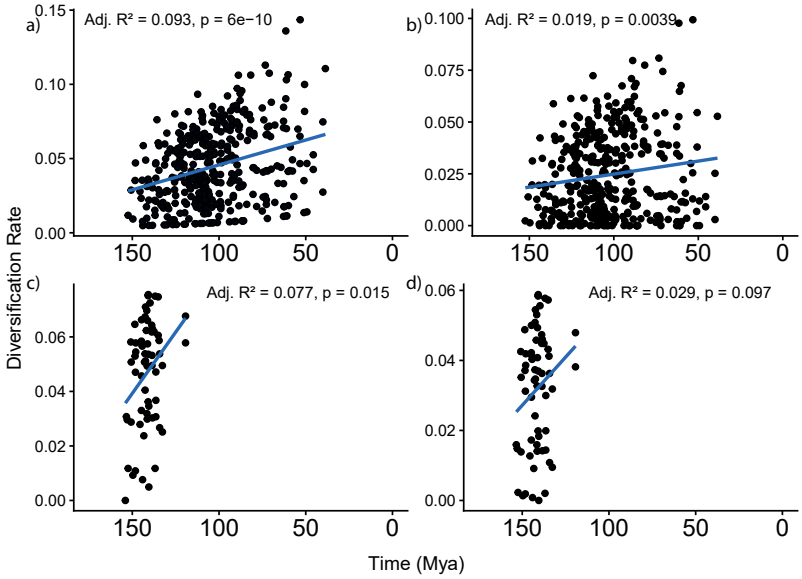
At the family level, across both the stem and crown datasets, older clades generally have more species than younger clades (Synthesis, Figure 7a, b). At the order level, for the stem dataset, there is no relationship between clade age and species diversity (Synthesis, Figure 7c). Using the crown dataset, the pattern is consistent with the family-level dataset (Synthesis, Figure 7d). Canarian trends (of no relationship between clade age and diversity) are therefore not representative of angiosperm-wide trends, and sampling biases may be affecting our inference, though there are also biological explanations for this difference.

I ran an analysis on the same dataset, but now using clade age versus diversification rate. The results show that younger clades (across both orders and families) diversify faster than older clades throughout angiosperm families and orders (Synthesis, Figure 8), consistent with what we found in the Canary Islands. Additionally, across all binned age categories (per ten million years), there is a significant relationship between diversity and diversification rates (Synthesis, Table 3), in agreement with Canarian trends that deep time is not a driver of diversity in angiosperms. The oldest age categories also did not have the most species. However, it is worth noting that Dimitrov et al. (2023) only found significance after 20 My in angiosperm genera; the youngest category in the dataset here is 40 My, as no families are estimated to be younger than that age in Zuntini et al. (2024). Therefore, the lineages analysed here may be too old to assess if there is only significance after a certain time period; equally, perhaps, the Canaries are too young a system to find a significant effect of age on diversity. I explore this line of thinking further below, in the 'Future Research' section.

Fourth, most diversity globally is contained within radiations that depart from the general trends outlined above (high speciation initially, which declines with age; Wiens and Moens 2025; Quintero et al. 2025). Radiations have high diversification rates, but these rates do not decline and appear unlinked with



Synthesis Figure 7. The relationship between clade age and species richness across angiosperms, for a) angiosperm families using stem age, b) angiosperm families using crown age, c) angiosperm orders using stem age, and d) angiosperm orders using crown age. Clade age estimates are from Zuntini et al. (2024), and species richness estimates are from Brown et al. (2023). The y-axis has a logarithmic scale, and the x-axis is fixed from 160 Mya to the present.



Synthesis Figure 8. Relationship between clade age and diversification rate across angiosperms, for a) angiosperm families and stem age, b) angiosperm families and crown age, c) angiosperm orders and stem age, and d) angiosperm orders and crown age. Clade age estimates are from Zuntini et al. (2024), and diversification rates were estimated using Magallon and Sanderson (2001). The y-axis varies across plots a - d.

time (Wiens and Moens 2025; Quintero et al. 2025); subsequently, these clades contribute overwhelmingly to total diversity. This pattern is also observed in the Canaries, with the 20 most species-rich non-monocot angiosperm lineages harbouring 343 of ~1,160 species, or 30% of diversity being contained within 3% of lineages (Chapter 2, Data S1). However, interestingly, the most species-rich clades are not the fastest diversifying (Synthesis, Table 4). Only five of the ten fastest diversifying clades in the Canaries are the most species-rich lineages; the rest are young clades, in agreement with global studies (Synthesis, Table 4; Wiens and Moens 2025). This is also true across angiosperm families (Synthesis, Table 5), and to a lesser extent, orders, which appear more time-scaled than other clades (Synthesis, Table 6).

Finally, there are several interesting topics regarding the drivers of diversity unevenness that I did not fully explore in this thesis. These include: (1) diversification rates and species richness are often unlinked for some taxa (Tietje et al. 2022; Stephens et al. 2025), with diversification rates of currently extant clades poorly reflecting modern diversity patterns (though we find diversification rates and species richness correlated across all age categories in Canarian angiosperms and angiosperm families globally; Synthesis, Tables 2, 3); (2) no universal drivers of diversification found throughout angiosperms (Helmstetter et al. 2023); and (3) the role of diversity dependence on diversification rates (Etienne and Haegeman 2012; Etienne et al. 2012; Rabosky and Hurlbert 2015; Pie et al. 2023; Quintero et al. 2025). The data made available in this thesis can help address these questions, whilst contributing to these debates by providing evidence for a link between diversification rates, age, and diversity (Synthesis, Tables 2-6). It also provides new insight into the drivers of trait-dependent diversification within a geographically restricted angiosperm assemblage (Chapter 4; see Synthesis, Section 1 above), and evidence of diversity-dependence operating within the Canary Islands (Chapter 2, Table 1). This thesis has therefore shown that diversity unevenness in the Canaries does not have a "black and white" explanation; rather, all three of the hypotheses to explain diversity unevenness operate within the non-monocot angiosperm flora of the archipelago. The interplay of these hypotheses, to the best of my knowledge at the time of writing, can be explained by the Zarathustran hypothesis of Quintero et al. (2025) to generate the flora we see today: a decrease in speciation rates and increase in extinction rates with lineage age, with few lineages withstanding the loss of diversification rate that generate a strong imbalance in diversity (i.e., radiating).

Synthesis Table 3. Relationship between diversity and species richness per age category for all angiosperm families. Each angiosperm family was binned into an age category based on stem age from Zuntini et al. (2024), the species richness was summed across the total number of families per age category, and the diversification rate (per family) was correlated with family species richness, following Stephens et al. (2025).

Age (My)	Number of Species	Adj. R^2	df	p
< 40	72	NA	0	NA
40-49	41	0.985	2	0.005
50-59	2455	0.563	10	0.003
60-69	6373	0.377	18	0.002
70-79	7623	0.475	18	4.7E-04
80-89	35731	0.354	30	2.0E-04
90-99	31184	0.438	43	4.4E-07
100-109	27958	0.435	57	8.2E-09
110-119	90050	0.286	56	9.2E-06
120-129	58500	0.253	38	0.001
130-139	49020	0.316	26	0.001
140-149	7577	0.416	14	0.004
>150	82	NA	0	NA

Synthesis Table 4. Fastest diversifying Canarian plant lineages with the number of species, stem colonisation time, and diversification rate estimated under no extinction and high extinction scenarios, using the Magallón and Sanderson method (2001). Using this method, an estimate of extinction rate (μ), from 0 (no extinction) to 1 (high extinction), is required. Diversification was estimated here under no extinction (μ = 0) and high extinction (μ = 0.9) scenarios. *Lavandula spp*. Includes *L. buchii* and *L. minutolii. Senecio spp. includes S. bollei and S. ilsae.*

Canarian Clades				
Assuming no extinction	**Number of species**	**Stem age**	**Diversification rate under no extinction**	**Diversification rate under high extinction**
Lavandula pp.	2	0.338	2.051	0.000
Silene	9	1.879	1.170	0.284
Lotus	19	2.602	1.132	0.376
Dicheranthus	2	0.649	1.068	0.000
Descurainia	7	2.415	0.806	0.172
Cheirolophus	18	4.174	0.692	0.225
Globularia	3	1.639	0.670	0.072
Micromeria	21	4.544	0.670	0.230
Erysimum	4	2.359	0.588	0.086
Senecio pp.	2	1.220	0.568	0.000

Synthesis Table 4. Continued.

Canarian Clades				
Assuming high extinction	**Number of species**	**Stem age**	**Diversification rate under no extinction**	**Diversification rate under high extinction**
Lotus	19	2.602	1.132	0.376
Silene	9	1.879	1.170	0.284
Micromeria	21	4.544	0.670	0.230
Cheirolophus	18	4.174	0.692	0.225
Echium	24	6.065	0.524	0.188
Descurainia	7	2.415	0.806	0.172
Sideritis	27	8.044	0.410	0.153
Helianthemum	15	5.906	0.459	0.139
Limonium	17	6.984	0.406	0.129
Argyranthemum	20	7.902	0.379	0.128

Synthesis Table 5. Fastest diversifying angiosperm families with the number of species (Brown et al. 2023), stem family age (Zuntini et al. 2024), and diversification rate estimate under no extinction and high extinction scenarios, using Magallon and Sanderson (2001).

Angiosperm Families				
Assuming no extinction	**Number of species**	**Stem age**	**Diversification rate under no extinction**	**Diversification rate under high extinction**
Cactaceae	2075	53.224	0.144	0.099
Brassicaceae	4274	61.490	0.136	0.098
Piperaceae	3938	73.308	0.113	0.081
Penaeaceae	72	38.669	0.111	0.053
Zingiberaceae	2066	71.014	0.107	0.074
Marantaceae	605	60.151	0.106	0.068
Cyperaceae	5895	81.729	0.106	0.077
Poaceae	12474	88.833	0.106	0.080
Dipterocarpaceae	545	61.143	0.103	0.065
Portulacaceae	158	50.672	0.100	0.055
Assuming high extinction				
Cactaceae	2075	53.224	0.144	0.099
Brassicaceae	4274	61.490	0.136	0.098
Piperaceae	3938	73.308	0.113	0.081
Poaceae	12474	88.833	0.106	0.080
Cyperaceae	5895	81.729	0.106	0.077
Zingiberaceae	2066	71.014	0.107	0.074
Asteraceae	35622	112.317	0.093	0.072
Melastomataceae	6036	89.533	0.097	0.071
Apiaceae	4066	86.868	0.096	0.069
Marantaceae	605	60.151	0.106	0.068

Synthesis Table 6. Fastest diversifying angiosperm orders with the number of species (Brown et al. 2023), stem family age (Zuntini et al. 2024), and diversification rate under no extinction and high extinction scenarios, using Magallon and Sanderson (2001).

Angiosperm Orders				
Assuming no extinction	**Number of species**	**Stem age**	**Diversification rate under no extinction**	**Diversification rate under high extinction**
Asparagales	41434	140.781	0.076	0.059
Asterales	39205	140.843	0.075	0.058
Lamiales	28461	136.848	0.075	0.058
Gentianales	24029	135.070	0.075	0.057
Poales	24991	139.741	0.072	0.056
Fabales	24519	142.387	0.071	0.054
Malpighiales	19335	141.655	0.070	0.053
Zingiberales	3188	119.294	0.068	0.048
Myrtales	14826	142.607	0.067	0.051
Ericales	14597	144.669	0.066	0.050
Assuming high extinction				
Asparagales	41434	140.781	0.076	0.059
Asterales	39205	140.843	0.075	0.058
Lamiales	28461	136.848	0.075	0.058
Gentianales	24029	135.070	0.075	0.057
Poales	24991	139.741	0.072	0.056
Fabales	24519	142.387	0.071	0.054
Malpighiales	19335	141.655	0.070	0.053
Myrtales	14826	142.607	0.067	0.051
Ericales	14597	144.669	0.066	0.050
Rosales	11184	141.218	0.066	0.049

Section 5 - Unfinished Projects

In the final sections of my synthesis, I would like to outline the projects I never had time to fully explore, as well as the future research avenues this thesis introduces. These unexplored projects vary in their degree of completeness and were not completed due to time or methodological constraints. For these, I will give a brief introduction on the subject, then provide the status of the project.

Over the last four years, several projects were started with the intention of becoming chapters in this thesis, but they never came to fruition. The first of these is a multi-lineage extension of the trait-dependent diversification model, SecSSE. As discussed in the introduction of this thesis, SecSSE (Several Examined and Concealed States-Dependent Speciation and Extinction) requires a complete tree of life for a single clade as input, restricting studies

into trait-dependent diversification across communities (Herrera-Alsina et al. 2019). Though trait-dependent diversification across communities is exactly the purpose of TRAISIE in Chapter 4, we began developing an extension of SecSSE that considered a subset of the community (rather than the entire community as assumed in TRAISIE), i.e., an SSE model that required phylogenetic trees of multiple lineages as input. We provisionally named this model 'Multi-SecSSE' (Multiple Several Examined and Concealed States-Dependent Speciation and Extinction). Unlike TRAISIE, it cannot estimate colonisation rates, but it would have the advantage that it could infer trait-dependent diversification across a subset of the community without requiring them to belong to the same clade. In collaboration with Thijs Janzen (University of Groningen), we (my supervisors, Ornela Dehayem Nanwou, and I) began developing and testing Multi-SecSSE to see how accurately it could estimate speciation, extinction, and transition rates. The model is fully coded, and we began testing the model before identifying several problems. Although under some model types (e.g., CR, CTD, ETD), Multi-SecSSE performed well, for others, Multi-SecSSE would estimate exceedingly high rates, to the point of being biologically meaningless. Also, we found that simulated data (simulated using the best-fitting (highest AIC weight) model) did not resemble the empirical data; the average tree size was too large, and the crown age distributions between empirical and simulated data were highly disparate, as were the root lengths. The cause(s) of these issues could not be identified.

The second unfinished project was a comprehensive analysis of how herbarium specimen characteristics affect sequencing success when using target capture-based methods (see Brewer et al. 2019; Kates et al. 2021). In collaboration with Kasper Hendriks (Naturalis), Lizzie Roeble (Naturalis), and the Plant and Fungal Tree of Life Project (Royal Botanical Gardens, Kew), we sought to generate the largest dataset thus far on how herbarium specimen characteristics (such as age, taxon-specific traits, climate of the country of collection, the 'greenness' of the sequenced specimen and preservation technique) affect the number of loci recovered (Brewer et al. 2019; Forrest et al. 2019; Kates et al. 2021). This dataset would have covered ~15,000 specimens, almost double the largest existing study (of 8,000 specimens; Kates et al. 2021), included traits unexplored by other studies (such as input material weight, source herbarium, extraction protocol, and the number of samples pooled before target capture), and spanned most of the evolutionary history of angiosperms (but with a bias towards Brassicaceae and Asteraceae related to family specific work of Kasper Hendriks and Lizzie Roeble, respectively). However, after discussions between all collaborators, we failed to find a novel direction for the project that would yield a publishable quality result; if we simply confirmed existing knowledge,

we agreed there would not be sufficient novelty to pursue this project. Furthermore, as standardising the datasets of the four collaborators would be a substantial undertaking, this project was abandoned at the planning stage. As many of these projects are now published and/or the datasets are open access (Hendriks et al. 2023; Zuntini et al. 2024; Brewer et al. 2025), the data are available if there is interest in pursuing this research.

Section 6 - Future Research

This then brings me to the final (sub)section of this thesis: future research. The most logical place to start, of course, is with completing the Canary Islands angiosperm tree of life. Though we have made a substantial effort here, sequencing ~50% of the native angiosperm lineages, which is unprecedented for an island plant community of this size (Chapter 2, Figure 2), substantial work remains to be done. To this end, we provide a checklist of angiosperm colonisations (Chapter 2, Table S1), as well as those we could not sequence for various reasons (Chapter 2, Table S2), to be used as a checklist for future researchers to pursue completing our sampling (though identification of the nearest continental ancestor would still be required, meaning extra sequencing efforts of the continental species are needed as well for a substantial number of island lineages). Incorporating population-level sampling at the community level constitutes another promising future goal, one that will take considerably more time to be achieved. Currently, population-level sampling is only available for select Canarian lineages (such as Asteraceae, *Kleinia, Periploca*; Garcia-Verdugo et al. 2019b; Martin-Hernanz et al. 2023; Rincón-Barrado et al. 2024; Arjona et al. 2025). Population-level sampling would also allow the possibility of tracing multiple colonisation events per species (Valtueña et al. 2016), and may allow for better inference of anagenetic speciation and transition rates in TRAISIE (Chapter 4, Tables 2). Regardless, a complete tree of life for any archipelago opens untold avenues for research, such as those given at the end of Chapter 2, and will give unparalleled insight into the evolutionary and biogeographic history of the Canaries. In combination with the trait database generated here, studying the traits best associated with radiation also becomes possible (Takayama et al. 2018).

A key question that can benefit from completing the phylogeny of Canary Island plants arises from Chapter 3: Are lignified monocot species (i.e., monocot lineages such as palms, and lineages that undergo secondary growth without developing wood, such as *Dracaena*) convergently similar in niche space to the woody non-monocot species? We found strong convergence in niche space between IW and AW species (Chapter 3, Figure 2), though DW species did not follow this trend. If secondary growth has similar environmental drivers

throughout angiosperms (i.e., whether lignification is a response primarily to drought; Lens et al. 2016; Zizka et al. 2025), I would expect these more lignified monocots to also show convergence with IW and AW species. This would also lend support to the findings of Lens et al. (2016), who found that grasses that are more resistant to embolism formation (proxy for drought tolerance) have a higher proportion of lignified tissues in their stems. Of course, including more species in the analyses of Chapter 3 may also lead to only further overlap in niche space, potentially further obscuring the already faint distinction in niche space between the woody and herbaceous Canarian species (Chapter 3, Figure 2c). Also interesting is to investigate whether obligate herbaceous monocots (obligate as they cannot evolve woodiness) occupy similar areas of environmental niche space as their non-monocot counterparts; i.e., do the annual herbaceous monocots overlap with the annual herbaceous non-monocots? These analyses would be an interesting insight into how phylogeny interacts with the environment to shape the flora.

Also related to Chapter 3, another future area of research would be to run our analyses using all global non-monocot angiosperms native to islands. The tools now exist such that this analysis would be a feasible undertaking. To start, a global checklist of all vascular plant species native to islands by Schrader et al. (2024) can be used as the species list input required to run our bioinformatics pipeline, once the non-monocot angiosperms are extracted. Next, lists of phylogenetically-informed growth forms can be cross-referenced with this list to identify IW and DW species (Zizka et al. 2022, 2025); any woody species not on these lists (which can be identified using databases such as GIFT (Weigelt et al. 2020; Denelle et al. 2023), TRY (Kattge et al. 2020; Diaz et al. 2022), rWCVP (Brown et al. 2023), or Poppenwimer et al. (2023)), by process of elimination, would be expected to be AW. Global bioclim variables at 100 m resolution are not yet available (at the time of writing), so the resolution of Chapter 3 would not be matched, but the 1 km resolution global bioclim variables of Fick and Hijmans (2017) provide a starting point. Chapter 3 then provides the bioinformatics pipeline (and preliminary figure-generating code) to redo our analyses. However, a remaining bottleneck would be the archipelago-level distribution data required by our pipeline, as Schrader et al. (2024) do not provide data on which islands within an archipelago a species is native to. Nevertheless, such an analysis would further our understanding of how environmental niche space varies between growth forms and increase our knowledge of the global environmental correlates of growth form distributions, providing detailed information on how environmental variables shape global plant distributions. This global analysis would also provide a baseline dataset for subsequent analyses, such as whether IW and DW species occur in more arid regions globally than

their herbaceous sister species, once the sister species are identified. Studies on global potential drivers of IW and DW have recently been published (Zizka et al. 2022, 2025), so addressing how prominent these drivers are per lineages is the next logical step.

Finally, in the 'Diversity Unevenness' section (Synthesis, Section 4), I began discussing how the Canaries may be too young a system, or native Canarian lineages are too young, to find a significant effect of age on diversity. Another interesting avenue of research is assessing whether there is a minimum age of a geographical system before which the time hypothesis (to explain global diversity unevenness; i.e., older lineages have more species) comes into effect. Specifically, this assumes a modification of the clade-age hypothesis that invokes an age-scaling effect, where a minimum time is required for this relationship to be established within a system. As explained above, we found no effect of time on lineage diversity (Chapter 2, Figure 4), and all age groups above are significantly correlated with species richness for Canarian angiosperms and angiosperms globally (Synthesis, Tables 2, 3), indicating that deep time is less important in driving lineage diversity than found in other systems (Dimitrov et al. 2023; Stephens et al. 2025). Stephens et al. (2025) only find a relationship between age and diversity for lineages older than 40 My, and Dimitrov et al. (2023) find most angiosperm genera are ~15-20 My old with a positive correlation between genus diversity and mean genus age (per geographic unit). In contrast, the Canarian lineages we assess here are on average 6.6 ± 0.50 - 4.4 ± 0.42 My old (stem and crown age, respectively; Chapter 2, Figure 3, Supplementary Data S12), and the currently emerged Canary Islands are only ~23 My old (Campeny et al. 2023). Therefore, if the clade-age relationship becomes stronger with age (per Dimitrov et al. 2023 and Stephens et al. 2025), the Canary Islands archipelago and Canarian lineages may simply be too young for the processes underlying the clade-age hypothesis to have affected the system. Alternatively, the unusually high turnover rates for Canary Island lineages with high colonisation and extinction rates could impose different biodiversity dynamics compared to the mechanisms that shape global diversity patterns.

Concluding Remarks

This thesis set out to investigate the processes underlying biodiversity in the Canary Island non-monocot angiosperm flora by addressing how they are distributed in time, space, and across their phylogeny. Herein, we show that Canarian non-monocot angiosperms are, on average, late-Miocene in origin, with continuous colonisations over the last ~20 My; their distribution is largely correlated with climatic seasonality and rugosity, with the impact of these factors varying per growth form; and their phylogeny is strongly shaped by

trait-dependent diversification (particularly by woody, perennial, IW species), to a lesser extent, ecological limits and time, where there is no relationship between clade age and diversity, but rather between clade age and diversification rate. We also provide the first empirical evidence of insular woodiness as a key innovation in the Canaries, with its evolution appearing to be an adaptation to a variable climate, with drought and low, non-freezing temperatures important contributing factors. These contributions add to our understanding of how species diversify and interact with their environments. But above all, I hope that this PhD thesis shows why the iconic Canary Islands flora is worthy of study and our attention.

SYNTHESIS REFERENCES

Anderson, B., Pannell, J., Billiard, S., Burgarella, C., de Boer, H., Dufay, M., & Glemin, S. (2023). Opposing effects of plant traits on diversification. *Iscience*, *26*(4).

Arjona, Y., Jay-Garcia, L. S., Reyes-Betancort, J. A., Morente-Lopez, J., Salas-Pascual, M., Naranjo-Cigala, A., & Patino, J. (2025). Taming Darwin's conundrum: On the role of phylogenetic and functional differentiation in plant naturalization on oceanic islands. *BioRxiv*, 2025-04. DOI: https://doi.org/10.1101/2025.04.09.648018.

Böhle, U. R., Hilger, H. H., & Martin, W. F. (1996). Island colonization and evolution of the insular woody habit in *Echium* L.(Boraginaceae). *Proceedings of the National Academy of Sciences, 93*(21), 11740-11745.

Brewer, G. E., Clarkson, J. J., Maurin, O., Zuntini, A. R., Barber, V., Bellot, S., & Baker, W. J. (2019). Factors affecting targeted sequencing of 353 nuclear genes from herbarium specimens spanning the diversity of angiosperms. *Frontiers in Plant Science*, *10*, 477275.

Brewer, R. F. A., Nanwou, O. D., van Hoek, L., del Arco Aguilar, M. J., Arjona, Y., Baker, W. J. & Lens, F. (2025). Trait-dependent diversification and spatio-ecological limits drive angiosperm diversity unevenness across the Canary Islands archipelago. *BioRxiv*, 2025-09.

Brown, M. J., Walker, B. E., Black, N., Govaerts, R. H., Ondo, I., Turner, R., & Nic Lughadha, E. (2023). rWCVP: a companion R package for the World Checklist of Vascular Plants. *New Phytologist*, *240*(4), 1355-1365.

Campeny, M., Menéndez, I., Quevedo, L., Yepes, J., Casillas, R., Ahijado, A., & Mangas, J. (2024). Rare Earth element distribution on the Fuerteventura Basal Complex (Canary Islands, Spain): a geochemical and mineralogical approach. *Solid Earth*, *15*(6), 639-656.

Carine, M. A., Santos-Guerra, A., Guma, I. R., & Reyes-Betancort, J. A. (2010). Endemism and evolution of the Macaronesian flora. *Beyond cladistics: The branching of a paradigm*, 101-124.

Carlquist, S. (1974). *Island Biology*. Columbia University Press, New York.

Davis, C. C. (2023). The herbarium of the future. *Trends in Ecology & Evolution, 38*(5), 412-423.

Davis, C. C., Sessa, E., Paton, A., Antonelli, A., & Teisher, J. K. (2025). Guidelines for the effective and ethical sampling of herbaria. *Nature Ecology & Evolution, 9*(2), 196-203.

Davis, M. P., Midford, P. E., & Maddison, W. (2013). Exploring power and parameter estimation of the BiSSE method for analyzing species diversification. *BMC Evolutionary Biology, 13*(1), 38.

Dehayem, O. N., Brewer, R. F., Valente, L., Lens, F., & Etienne, R. S. (2025). Impact of sampling strategy on inference of community assembly processes in phylogenetic island biogeography. *Methods in Ecology and Evolution*, *16*, 1507-1520.

Denelle, P., Weigelt, P., & Kreft, H. (2023). GIFT—An R package to access the Global Inventory of Floras and Traits. *Methods in Ecology and Evolution, 14*(11), 2738-2748.

Díaz, S., Kattge, J., Cornelissen, J. H., Wright, I. J., Lavorel, S., Dray, S., & Zotz, G. (2022). The global spectrum of plant form and function: enhanced species-level trait dataset. *Scientific Data, 9*(1), 755.

Dimitrov, D., Xu, X., Su, X., Shrestha, N., Liu, Y., Kennedy, J. D. & Wang, Z. (2023). Diversification of flowering plants in space and time. *Nature Communications, 14*(1), 7609.

Dos Santos, P., Brilhante, M. Â., Messerschmid, T. F., Serrano, H. C., Kadereit, G., Branquinho, C., & De Vos, J. M. (2022). Plant growth forms dictate adaptations to the local climate. *Frontiers in Plant Science, 13*, 1023595.

Etienne, R. S., & Haegeman, B. (2012). A conceptual and statistical framework for adaptive radiations with a key role for diversity dependence. *The American Naturalist, 180*(4), E75-E89.

Etienne, R. S., Haegeman, B., Stadler, T., Aze, T., Pearson, P. N., Purvis, A., & Phillimore, A. B. (2012). Diversity-dependence brings molecular phylogenies closer to agreement with the fossil record. *Proceedings of the Royal Society B: Biological Sciences, 279*(1732), 1300-1309.

Fick, S.E. and R.J. Hijmans (2017). WorldClim 2: new 1km spatial resolution climate surfaces for global land areas. *International Journal of Climatology, 37*(12): 4302-4315.

Forrest, L. L., Hart, M. L., Hughes, M., Wilson, H. P., Chung, K. F., Tseng, Y. H., & Kidner, C. A. (2019). The limits of HybSeq for herbarium specimens: impact of preservation techniques. *Frontiers in Ecology and Evolution*, *7*, 439.

Fortini, L. B., Kaiser, L. R., Xue, L., & Wang, Y. (2022). Bioclimatic variables dataset for baseline and future climate scenarios for climate change studies in Hawai'i. *Data in Brief, 45*, 108572.

Francisco-Ortega, J., Jansen, R. K., & Santos-Guerra, A. (1996). Chloroplast DNA evidence of colonization, adaptive radiation, and hybridization in the evolution of the Macaronesian flora. *Proceedings of the National Academy of Sciences, 93*(9), 4085-4090.

Frateles, L. E. F., Tavares, G. R. G., Nakamura, G., da Silva Jr., N. J., Terribile, L. C., & Diniz-Filho, J. A. F. (2025). The interaction between the Linnean and Darwinian shortfalls affects our understanding of the evolutionary dynamics driving diversity patterns of New World coralsnakes. *Journal of Biogeography, 52*(1), 42-54.

García-Maroto, F., Mañas-Fernández, A., Garrido-Cárdenas, J. A., Alonso, D. L., Guil-Guerrero, J. L., Guzmán, B., & Vargas, P. (2009). Δ6-Desaturase sequence evidence for explosive Pliocene radiations within the adaptive radiation of Macaronesian Echium (Boraginaceae). *Molecular Phylogenetics and Evolution, 52*(3), 563-574.

García-Verdugo, C., Caujapé-Castells, J., Illera, J. C., Mairal, M., Patiño, J., Reyes-Betancort, A., & Scholz, S. (2019a). Pleistocene extinctions as drivers of biogeographical patterns on the easternmost Canary Islands. *Journal of Biogeography, 46*(5), 845-859.

García-Verdugo, C., Caujapé-Castells, J., Mairal, M., & Monroy, P. (2019b). How repeatable is microevolution on islands? Patterns of dispersal and colonization-related plant traits in a phylogeographical context. *Annals of Botany, 123*(3), 557-568.

Goldberg, E. E., & Igić, B. (2012). Tempo and mode in plant breeding system evolution. *Evolution*, *66*(12), 3701-3709.

Graham, C. H., Araujo, M. L., Barreto, E., Dambros, C. S., Diniz-Filho, J. A. F., Zimmermann, N. E. & Coelho, M. T. P. (2025). Biodiversity patterns redefined in environmental space. *Ecology Letters, 28*(2), e70008.

Harmon, L. J., Pennell, M. W., Henao-Diaz, L. F., Rolland, J., Sipley, B. N., & Uyeda, J. C. (2021). Causes and consequences of apparent timescaling across all estimated evolutionary rates. *Annual Review of Ecology, Evolution, and Systematics, 52*(1), 587-609.

Helmstetter, A. J., Zenil-Ferguson, R., Sauquet, H., Otto, S. P., Méndez, M., Vallejo-Marin, M. & Käfer, J. (2023). Trait-dependent diversification in angiosperms: Patterns, models and data. *Ecology Letters, 26*(4), 640-657.

Hendriks, K. P., Kiefer, C., Al-Shehbaz, I. A., Bailey, C. D., van Huysduynen, A. H., Nikolov, L. A., ... & Lens, F. (2023). Global Brassicaceae phylogeny based on filtering of 1,000-gene dataset. *Current Biology, 33*(19), 4052-4068.

Heo, N., Yun, S., & Lomolino, M. V. (2023). Towards a holistic perspective on the development of island syndrome by examining its occurrence patterns in insular plants. *Frontiers of Biogeography, 15*(3).

Herrera-Alsina, L., Van Els, P., & Etienne, R. S. (2019). Detecting the dependence of diversification on multiple traits from phylogenetic trees and trait data. *Systematic Biology, 68*(2), 317-328.

Hooft van Huysduynen, A., Janssens, S., Merckx, V., Vos, R., Valente, L., Zizka, A. & Lens, F. (2021). Temporal and palaeoclimatic context of the evolution of insular

woodiness in the Canary Islands. *Ecology and Evolution, 11*(17), 12220-12231.

Jiménez-Moreno, G., Fauquette, S., & Suc, J. P. (2010). Miocene to Pliocene vegetation reconstruction and climate estimates in the Iberian Peninsula from pollen data. *Review of Palaeobotany and Palynology, 162*(3), 403-415.

Jiménez-Moreno, G., Burjachs, F., Expósito, I., Oms, O., Carrancho, Á., Villalaín, J. J., & van der Made, J. (2013). Late Pliocene vegetation and orbital-scale climate changes from the western Mediterranean area. *Global and Planetary Change, 108*, 15-28.

Jin, Y., & Qian, H. (2022). V. PhyloMaker2: An updated and enlarged R package that can generate very large phylogenies for vascular plants. *Plant Diversity, 44*(4), 335-339.

Jones, K. E., Pérez-Espona, S., Reyes-Betancort, J. A., Pattinson, D., Caujapé-Castells, J., Hiscock, S. J., & Carine, M. A. (2016). Why do different oceanic archipelagos harbour contrasting levels of species diversity? The Macaronesian endemic genus *Pericallis* (Asteraceae) provides insight into explaining the 'Azores diversity Enigma'. *BMC Evolutionary Biology, 16*(1), 202.

Kates, H. R., Doby, J. R., Siniscalchi, C. M., LaFrance, R., Soltis, D. E., Soltis, P. S. & Folk, R. A. (2021). The effects of herbarium specimen characteristics on short-read NGS sequencing success in nearly 8000 specimens: old, degraded samples have lower DNA yields but consistent sequencing success. *Frontiers in Plant Science, 12*, 669064.

Kattge, J., Bönisch, G., Díaz, S., Lavorel, S., Prentice, I. C., Leadley, P., & Cuntz, M. (2020). TRY plant trait database–enhanced coverage and open access. *Global Change Biology, 26*(1), 119-188.

Lens, F., Davin, N., Smets, E., & del Arco, M. (2013). Insular woodiness on the Canary Islands: a remarkable case of convergent evolution. *International Journal of Plant Sciences, 174*(7), 992-1013.

Lens, F., Picon-Cochard, C., Delmas, C. E., Signarbieux, C., Buttler, A., Cochard, H. & Delzon, S. (2016). Herbaceous angiosperms are not more vulnerable to drought-induced embolism than angiosperm trees. *Plant Physiology, 172*(2), 661-667.

Louca, S., Henao-Diaz, L. F., & Pennell, M. (2022). The scaling of diversification rates with age is likely explained by sampling bias. *Evolution, 76*(7), 1625-1637.

Maddison, W. P., Midford, P. E., & Otto, S. P. (2007). Estimating a binary character's effect on speciation and extinction. *Systematic Biology, 56*(5), 701-710.

Magallon, S., & Sanderson, M. J. (2001). Absolute diversification rates in angiosperm clades. *Evolution, 55*(9), 1762-1780.

Martín-Hernanz, S., Nogales, M., Valente, L., Fernández-Mazuecos, M., Pomeda-Gutiérrez, F., Cano, E. & Vargas, P. (2023). Time-calibrated phylogenies reveal Mediterranean and pre-Mediterranean origin of the thermophilous vegetation of the Canary Islands. *Annals of Botany, 131*(4), 667-684.

Meyer, L., Ladle, R. J., Trad, R. J., Frazão, A., Frateles, L., Lessa, T., & Stropp, J. (2026). Taxonomic uncertainty: causes, consequences, and metrics. *Trends in Ecology & Evolution*.

Mort, M. E., Solits, D. E., Soltis, P. S., Santos-Guerra, A., & Francisco-Ortega, J. (2007). Physiological evolution and association between physiology and growth form in *Aeonium* (Crassulaceae). *Taxon, 56*(2), 453-464.

Mynard, P., Algar, A. C., Lancaster, L. T., Bocedi, G., Fahri, F., Gubry-Rangin, C. & Herrera-Alsina, L. (2023). Impact of phylogenetic tree completeness and mis-specification of sampling fractions on trait dependent diversification models. *Systematic Biology, 72*(1), 106-119.

National Atlas of the United States. (2013). 100-Meter Resolution Color-Sliced Elevation of Hawaii. National Atlas of the United States. Available at: http://purl.stanford.edu/pc699kh4508.

Nge, F. J., & Skeels, A. (2025). Diversification patterns of the southwest Australian biodiversity hotspot reveal a novel macroevolutionary pathway to plant hyperdiversity. *New Phytologist, 248*(2), 953-967.

Pérez-Vargas, I., Portero Álvarez, A. M., Pérez de Paz, P. L., & PÉrez, J. A. (2020). Retrotransposon-based molecular markers as

a tool in delimiting species in section *Rynchol*otus, a recent radiation group of Macaronesian *Lotus*. *Systematics and Biodiversity, 19*(2), 110-120.

Pie, M. R., Divieso, R., & Caron, F. S. (2023). Clade density and the evolution of diversity-dependent diversification. *Nature Communications, 14*(1), 4576.

Poppenwimer, T., Mayrose, I., & DeMalach, N. (2023). Revising the global biogeography of annual and perennial plants. *Nature*, *624*(7990), 109-114.

Price, J. P., Otto, R., Menezes de Sequeira, M., Kueffer, C., Schaefer, H., Caujapé-Castells, J., & Fernández-Palacios, J. M. (2018). Colonization and diversification shape species–area relationships in three Macaronesian archipelagos. *Journal of Biogeography, 45*(9), 2027-2039.

Quintero, I., Andréoletti, J., Silvestro, D., & Morlon, H. (2025). Loss of macroevolutionary species fitness explains the rise and fall of clades. *Nature Ecology & Evolution*, 1-12.

Rabosky, D. L., & Hurlbert, A. H. (2015). Species richness at continental scales is dominated by ecological limits. *The American Naturalist, 185*(5), 572-583.

Rincón-Barrado, M., Villaverde, T., Pérez, M. F., Sanmartín, I., & Riina, R. (2024). The sweet tabaiba or there and back again: phylogeographical history of the Macaronesian Euphorbia balsamifera. *Annals of Botany, 133*(5-6), 883-904.

Roeble, L., van Benthem, K. J., Weigelt, P., Kreft, H., Knope, M. L., Mandel, J. R. & Valente, L. (2024). Island biogeography of the megadiverse plant family Asteraceae. *Nature Communications, 15*(1), 7276.

Scholl, J. P., & Wiens, J. J. (2016). Diversification rates and species richness across the Tree of Life. *Proceedings of the Royal Society B: Biological Sciences, 283*(1838), 20161334.

Schrader, J. (2025). Trait-based island biogeography as a tool for studying future ecological communities. *New Phytologist*.

Schrader, J., Wright, I. J., Kreft, H., & Westoby, M. (2021). A roadmap to plant functional island biogeography. *Biological Reviews, 96*(6), 2851-2870.

Schrader, J., Weigelt, P., Cai, L., Westoby, M., Fernández-Palacios, J. M., Cabezas, F. J. & Kreft, H. (2024). Islands are key for protecting the world's plant endemism. *Nature, 634*(8035), 868-874.

Schüßler, C. (2020). No Tertiary relicts? A biogeographical study on the Macaronesian laurel forest species in *Daucus* (Apiaceae), *Geranium* (Geraniaceae), *Gesnouinia* (Urticaceae), *Phyllis* (Rubiaceae), *Semele* (Asparagaceae) and *Visnea* (Pentaphylacaceae) (Doctoral dissertation, Ruperto Carola University Heidelberg, Germany).

Smith, S. A., & Brown, J. W. (2018). Constructing a broadly inclusive seed plant phylogeny. *American Journal of Botany, 105*(3), 302-314.

Steinbauer, M. J., Field, R., Grytnes, J. A., Trigas, P., Ah-Peng, C., Attorre, F. & Beierkuhnlein, C. (2016). Topography-driven isolation, speciation and a global increase of endemism with elevation. *Global Ecology and Biogeography, 25*(9), 1097-1107.

Stephens, P. R., Farrell, M. J., Davies, T. J., Gittleman, J. L., Meiri, S., Moreira, M. O., & Wiens, J. J. (2025). Global Diversity patterns are explained by diversification rates and dispersal at ancient, not shallow, timescales. *Systematic Biology*, syaf018.

Swain, H., & Chakraborty, K. (2024). Science behind herbarium and its importance in recent years. *Nordic Journal of Botany, 2024*(12), e04499.

Takayama, K., Crawford, D. J., López-Sepúlveda, P., Greimler, J., & Stuessy, T. F. (2018). Factors driving adaptive radiation in plants of oceanic islands: a case study from the Juan Fernández Archipelago. *Journal of Plant Research, 131*(3), 469-485.

Tietje, M., Antonelli, A., Baker, W. J., Govaerts, R., Smith, S. A., & Eiserhardt, W. L. (2022). Global variation in diversification rate and species richness are unlinked in plants. *Proceedings of the National Academy of Sciences, 119*(27), e2120662119.

Valente, L. M., Phillimore, A. B., & Etienne, R. S. (2015). Equilibrium and non-equilibrium dynamics simultaneously operate

in the Galápagos islands. Ecology Letters, 18(8), 844-852.

Valente, L., Phillimore, A. B., Melo, M., Warren, B. H., Clegg, S. M., Havenstein, K., & Etienne, R. S. (2020). A simple dynamic model explains the diversity of island birds worldwide. *Nature*, *579*(7797), 92-96.

Valtuena, F. J., López, J., Álvarez, J., Rodríguez-Riaño, T., & Ortega-Olivencia, A. (2016). *Scrophularia arguta*, a widespread annual plant in the Canary Islands: a single recent colonization event or a more complex phylogeographic pattern?. *Ecology and evolution, 6*(13), 4258-4273.

Weigelt, P., König, C., & Kreft, H. (2020). GIFT–A Global Inventory of Floras and Traits for macroecology and biogeography. *Journal of Biogeography, 47*(1), 16-43.

Whittaker, R. J., Triantis, K. A., & Ladle, R. J. (2008). A general dynamic theory of oceanic island biogeography. *Journal of Biogeography, 35*(6), 977-994.

Wiens, J. J. (2023). Trait-based species richness: Ecology and macroevolution. *Biological Reviews*, *98*(4), 1365-1387.

Wiens, J. J., & Moen, D. S. (2025). Rapid radiations underlie most of the known diversity of life. *Frontiers in Ecology and Evolution, 13*.

Yu, D., & Wiens, J. J. (2024). The causes of species richness patterns among clades. *Proceedings of the Royal Society B, 291*(2015), 20232436.

Zanne, A. E., Tank, D. C., Cornwell, W. K., Eastman, J. M., Smith, S. A., FitzJohn, R. G., & Beaulieu, J. M. (2014). Three keys to the radiation of angiosperms into freezing environments. *Nature*, *506*(7486), 89-92.

Zizka, A., Onstein, R. E., Rozzi, R., Weigelt, P., Kreft, H., Steinbauer, M. J. & Lens, F. (2022). The evolution of insular woodiness. *Proceedings of the National Academy of Sciences, 119*(37), e2208629119.

Zizka, A., Grenie, M., Conradi, T., Onstein, R., Bruelheide, H., & Lens, F. (2025). Global convergence in wood evolution is driven by drought on continents and frost-free temperatures on islands. *BioRxiv*, 2025-09.

Zomer, R. J., Xu, J., & Trabucco, A. (2022). Version 3 of the global aridity index and potential evapotranspiration database. *Scientific Data, 9*(1), 409.

Zuntini, A. R., Carruthers, T., Maurin, O., Bailey, P. C., Leempoel, K., Brewer, G. E., & Knapp, S. (2024). Phylogenomics and the rise of the angiosperms. *Nature*, *629*(8013), 843-850.



Summary



A fundamental characteristic of biodiversity is its uneven distribution across the tree of life. This difference in diversity between clades has been attributed to either i) time, with older clades harbouring more species as these clades have had longer to accumulate species; ii) differential diversification rates between clades due to intrinsic properties of the clades, due to, for example, morphological, environmental, and/or anatomical traits; and iii) ecological limits to diversity, where limits to resource availability or carrying capacity within a system restricts diversity. However, analysing these evolutionary patterns and processes globally is extremely complex, so smaller geographic units are often used to study these processes: islands. Islands simplify complex processes through their distinct boundaries, isolation, and comparatively simpler, but unique, biotas relative to continental systems. The Canary Islands represent one of the best-studied archipelago systems globally, with their biota studied in impressive detail, and one of the most species rich flora of any oceanic archipelago. This knowledge makes the Canary Islands an ideal arena to study the processes underlying biodiversity. In this research, I use a combination of phylogenomics, morphological and environmental trait data, and biogeographical modelling to study how the flowering plant (angiosperm) community of the Canary Islands is distributed in time, space, and across their phylogeny.

To do so, I built a dataset identifying the number of flowering plant colonisations of the Canary Islands using published literature and the number of species per lineage. Using this dataset, I generated the most complete time-calibrated phylogenomic dataset of any archipelago flora and their continental ancestors globally, using 353 nuclear genes and time-calibrated using 166 fossils, representing ~60% of the Canary Islands non-monocot angiosperm community and ~65% of the colonisation events (Chapter 2). Using a dynamic, stochastic island biogeography model, we find strong evidence of differential diversification rates between lineages comprised primarily of different traits (trait-dependent diversification). Specifically, perennial lineages and lineages that evolved woodiness after island colonisation from a herbaceous ancestor (insular woodiness) have faster rates of speciation, and lower extinction and colonisation rates than annual lineages and lineages without insular woodiness. Interestingly, woody lineages have the same speciation rates as herbaceous lineages, but lower extinction rates. We also found some evidence of ecological limits as insular woody lineages had a detectable carrying capacity (i.e., as the number of species increases, rates of speciation and colonisation decrease), and the flora was found to be at macroevolutionary equilibrium. Interestingly, there was no relationship between clade age and diversity.

As Chapter 2 found growth form to be an important aspect of the Canarian flora, we looked into which environmental variables affect the distribution of species with these different growth forms. We built a novel pipeline to extract, filter, and analyse publicly available spatial data to identify the environmental correlates of five growth forms (annual herbaceous, perennial herbaceous, ancestral woodiness, derived woodiness, and insular woodiness). We reconstructed environmental niche space for these growth forms using a phylogenetic principal component analysis with a dataset of ~600 non-monocot angiosperm species. We found considerable overlap in niche space among these growth forms, and only minor differences between a combined woody and a combined herbaceous niche. Niche space is delineated by temperature and precipitation seasonality, and topographic complexity (rugosity). Using phylogenetic ANOVAs, we find significant pairwise differences between growth forms across five environmental variables, with rugosity having the most significant differences between growth forms. From this data, insular woodiness appears to be an adaptation to a variable climate, allowing insular woody species to persist across a wide range of climatic conditions, likely in response to the topographically complex areas they primarily inhabit. Drought-related variables and a low, non-freezing temperature in the coldest quarter are also contributing factors.

I also built a trait dataset containing 24 morphological traits and environmental variables to assess how traits have affected diversification (speciation minus extinction) rates. We selected two traits with high data coverage (life span and growth form) and combined these traits with the phylogenomic data from Chapter 2 to test a novel trait-dependent diversification model specifically tailored to islands, TRAISIE (TRait-dependent Assembly of Island biota through Speciation, Immigration, and Extinction). TRAISIE is a likelihood-based framework that allows for the estimation of colonisation, speciation, and extinction rates per trait state, as well as transition rates between states, using phylogenetic and trait data for an island community. We apply TRAISIE first to simulated data to show that for datasets containing >100 lineages and >200 species, TRAISIE can accurately estimate parameters and perform model selection, with accuracy increasing with dataset size. Next, we applied TRAISIE to the Canary Islands angiosperm dataset. We found that woody species and perennial species have faster speciation rates and lower extinction rates than herbaceous species and annual species, respectively. We also tested for confounding effects of these traits using a three-state model (annual herbaceous, perennial herbaceous, perennial woody), and found that woody species have the fastest speciation rates, perennial herbaceous the lowest extinction rates, and annual herbaceous the fastest colonisation rates.

In short, this thesis finds strong evidence of trait-dependent diversification operating in the Canary Islands, providing the first empirical data supporting the hypothesis that insular woodiness promotes diversification in the Canaries; Canary plant distributions are shaped by climatic seasonality and rugosity, both of which likely contribute to the evolution of insular woodiness; and provide a comprehensive trait dataset required to test a novel trait-dependent diversification model specifically tailored to islands, TRAISIE.



Nederlandse samenvatting

Translated in Google Translate, corrected by Frederic Lens



Een fundamenteel kenmerk van biodiversiteit is de ongelijke verdeling ervan over de stamboom van het leven. Dit verschil in diversiteit tussen clades wordt toegeschreven aan: i) tijd, waarbij oudere groepen meer soorten herbergen omdat ze langer de tijd hebben gehad om soorten te accumuleren; ii) verschillende diversificatiesnelheden tussen groepen door intrinsieke eigenschappen van de evolutionaire lijnen, zoals morfologische en omgevingskenmerken; en iii) ecologische limitaties van diversiteit, waarbij beperkingen in de beschikbaarheid van hulpbronnen of de draagcapaciteit van een systeem de diversiteit beperken. Het analyseren van deze evolutionaire patronen en processen op wereldschaal is echter buitengewoon complex, waardoor kleinere geografische regio's vaak worden gebruikt om deze processen te bestuderen: eilanden. Eilanden vereenvoudigen complexe processen door hun duidelijke grenzen, hun isolatie en hun relatief eenvoudigere, maar unieke, biota ten opzichte van continentale systemen. De Canarische Eilanden zijn wereldwijd één van de best bestudeerde archipelsystemen, met een uitgebreid onderzochte biota en één van de soortenrijkste flora's van alle oceanische archipels. Deze kennis maakt de Canarische Eilanden tot een ideale omgeving om de processen achter biodiversiteit te bestuderen. In dit onderzoek gebruik ik een combinatie van fylogenomica, morfologische en omgevingskenmerken, en biogeografische modellering om te bestuderen hoe de bloemplanten (angiospermen) van de Canarische Eilanden zijn geëvolueerd.

Hiertoe heb ik een dataset opgebouwd die het aantal kolonisaties van bloemplanten op de Canarische Eilanden identificeert, gebruikmakend van gepubliceerde literatuur en het aantal soorten per kolonisatie. Met behulp van deze dataset heb ik de meest volledige tijdsgekalibreerde fylogenomische dataset van welke archipelflora en hun continentale voorouders dan ook ter wereld gegenereerd, met 353 nucleaire genen en tijdsgekalibreerd met behulp van 166 fossielen, wat ~60% van de niet-monocotylen angiospermengemeenschap van de Canarische Eilanden en ~65% van de kolonisatiegebeurtenissen vertegenwoordigt (Hoofdstuk 2). Met behulp van een dynamisch, stochastisch eilandbiogeografisch model vinden we sterk bewijs voor verschillende diversificatiesnelheden tussen groepen die voornamelijk bestaan uit verschillende kenmerken (kenmerkafhankelijke diversificatie). Met name meerjarige groepen en groepen die na eilandkolonisatie houtigheid hebben ontwikkeld vanuit een kruidachtige voorouder (eilandhoutigheid) hebben snellere speciatiesnelheden en lagere extinctie- en kolonisatiesnelheden dan éénjarige groepen en groepen zonder eilandhoutigheid. Interessant genoeg hebben houtige groepen lagere cladogenetische speciatiesnelheden dan kruidachtige groepen, maar hogere anagenetische speciatiesnelheden en lagere extinctiesnelheden. We vonden ook enig bewijs van ecologische limitaties, aangezien de groepen

gekenmerkt door eilandhoutigheid een detecteerbare draagcapaciteit hadden (d.w.z. naarmate het aantal soorten toeneemt, nemen de speciatie- en kolonisatiesnelheden af), en de Canarische Eilandflora bleek zich in macro-evolutionair evenwicht te bevinden. Interessant genoeg was er geen verband tussen clade-leeftijd en diversiteit.

Omdat Hoofdstuk 2 groeivorm als een belangrijk aspect van de Canarische flora had aangetoond, onderzochten we welke omgevingsvariabelen de verspreiding van soorten met deze verschillende groeivormen beïnvloeden. We ontwikkelden een nieuwe pijplijn om publiekelijk beschikbare ruimtelijke gegevens te extraheren, filteren en analyseren, teneinde de omgevingscorrelaties van vijf groeivormen te identificeren (éénjarig kruidachtig, meerjarig kruidachtig, voorouderlijke houtigheid, afgeleide houtigheid en eilandhoutigheid). We reconstrueerden de omgevingsnicheruimte voor deze groeivormen met behulp van een fylogenetische principale-componentenanalyse met een dataset van ~600 niet-monocotyle angiospermen. We vonden aanzienlijke overlap in nicheruimte tussen deze groeivormen, en slechts geringe verschillen tussen een gecombineerde houtige en een gecombineerde kruidachtige niche. De nicheruimte wordt begrensd door de seizoensvariatie in temperatuur en neerslag, en topografische complexiteit (reliëf van het landschap). Met behulp van fylogenetische ANOVA's vinden we significante paarsgewijze verschillen tussen groeivormen over vijf omgevingsvariabelen, waarbij landschapsreliëf de meest significante verschillen tussen groeivormen vertoont. Op basis van deze gegevens lijkt eilandhoutigheid een aanpassing aan een variabel klimaat te zijn, waardoor eilandhoutige soorten kunnen gedijen onder een breed scala aan klimatologische omstandigheden, waarschijnlijk als reactie op de topografisch complexe gebieden die ze voornamelijk bewonen. Droogtegerelateerde variabelen en een lage, maar geen vorsttemperatuur in het koudste kwartaal zijn ook factoren die een rol spelen.

Ik heb ook een kenmerkendataset opgebouwd met 24 morfologische kenmerken en omgevingsvariabelen om te beoordelen hoe kenmerken de diversificatiesnelheid (speciatie minus extinctie) hebben beïnvloed. We selecteerden twee kenmerken met een hoge informatiedichtheid (levensduur en groeivorm) en combineerden deze kenmerken met de fylogenomische gegevens uit Hoofdstuk 2 om een nieuw kenmerkafhankelijk diversificatiemodel te testen dat specifiek is afgestemd op eilanden: TRAISIE (TRait-dependent Assembly of Island biota through Speciation, Immigration, and Extinction). TRAISIE is een op aannemelijkheid gebaseerd model dat de schatting mogelijk maakt van kolonisatie-, speciatie- en extinctiesnelheden per kenmerkstoestand, evenals transitiesnelheden tussen toestanden, met behulp van fylogenetische en kenmerkengegevens voor een eilandgemeenschap. We passen TRAISIE eerst toe

op gesimuleerde gegevens om aan te tonen dat TRAISIE voor datasets met meer dan 200 soorten parameters nauwkeurig kan schatten en modelkeuze kan uitvoeren, waarbij de nauwkeurigheid toeneemt met de omvang van de dataset. Vervolgens hebben we TRAISIE toegepast op de bloemplantendataset van de Canarische Eilanden. We vonden dat houtige soorten en meerjarige soorten snellere speciatiesnelheden en lagere extinctiesnelheden hebben dan respectievelijk kruidachtige soorten en éénjarige soorten. We hebben ook getest op verwarrende effecten van deze kenmerken met behulp van een drie-kenmerktoestand model (éénjarig kruidachtig, meerjarig kruidachtig, meerjarig houtig), en vonden dat houtige soorten de snelste speciatiesnelheden hebben, meerjarig kruidachtige soorten de laagste extinctiesnelheden, en éénjarig kruidachtige soorten de snelste kolonisatiesnelheden.

Kortom: dit proefschrift biedt sterk bewijs voor kenmerkafhankelijke diversificatie op de Canarische Eilanden, levert de eerste empirische gegevens ter ondersteuning van de hypothese dat eilandhoutigheid diversificatie op de Canarische Eilanden bevordert; toont aan dat de plantenverspreiding op de Canarische Eilanden wordt gevormd door klimatologische seizoensvariatie en landschapsreliëf, die beide waarschijnlijk bijdragen aan de evolutie van eilandhoutigheid; en biedt een uitgebreide kenmerkendataset die nodig is om een nieuw kenmerkafhankelijk diversificatiemodel te testen dat specifiek is ontwikkeld voor eilanden: TRAISIE.



Acknowledgements



There are so many people without whom this thesis would not have been possible. It is the result of a massive collaborative effort, and without any of the people mentioned in this section, it would have looked very different. It is by no means complete, so if anyone is left off, it is not intentional; I simply cannot name everyone who has helped me along the way in these last few years. Throughout my PhD, joy and motivation came from so many places, so I will try to express my gratitude to those sources from along the way. That being said, I'd be remiss if I didn't start with my supervisors: Frederic Lens, Luís Valente, and Rampal Etienne.

Frederic, it's been a real pleasure working with you. I want to start by thanking you not just for the opportunity to do this PhD, but also for giving me so much flexibility and freedom to shape this thesis. Not everyone can say that they truly felt they had control to take their thesis in the direction they wanted, but I definitely can. I hope that this thesis also achieved what you wanted from the project. Thank you for teaching me so much about plants, phylogenetics, the Canary Islands, and, of course, why the evolution of woodiness is fascinating. I will forever be grateful for being given the chance to advance human knowledge on these topics. On a less professional note, I've also enjoyed the few times we've indulged my British fondness for a drink, either as a lab, at conferences, or at Naturalis Christmas Festivals; I'll never forget your love for making nearby people dance with you.

Luís, thank you for all your patience along the way, for helping me understand how modelling works after my hopeless attempts on my own, and for all the feedback on how to give a good presentation and write a good manuscript. You've been so generous with your time, and for that, I cannot thank you enough. As a conservationist with an interest in biogeography that I hadn't yet had the opportunity to explore, thank you for giving me the opportunity and for allowing me to teach MSc students about the topic. I can't think of a better person to learn about the world of island biogeography from. Your immense knowledge of the field is truly impressive. Though my contribution to the field does not feel like much, I hope it was what you had in mind.

Rampal, though we only worked closely together in the second half of my PhD, I feel like I've learnt so much from you, both professionally and personally. I don't know how you do it all - academia, family life, music, and hobbies; it's truly inspirational. As an early-career researcher, I hope one day to have the work-life balance you've achieved. Thank you for your patience with me as I learned the world of diversification modelling, and for teaching me why the

small details in complicated matters matter. I hope it's not been too frustrating watching me start from the very beginning and make such simple mistakes along the way. To all three of my supervisors, thank you for everything. It's been an unforgettable experience and a pleasure working with you all.

I also want to thank my fellow PhD candidate on this project, Ornela Dehayem Nanwou. Ornela, your mathematical knowledge is incredible. How you can look at the equations underlying the models used in our theses and not only understand them but actually build on them is mind-blowingly impressive. It's amazing what you've achieved in building TRAISIE. Listening to you and Rampal talk about the nuances of how TRAISIE was constructed always confused me, and I was always impressed by how seamlessly you could answer such complicated questions. I hope your foray into the world of botany was enjoyable and you enjoyed working on this project as much as I did. After 4.5 years, we did it!

Within Groningen, I was part of the Theoretical & Evolutionary Community Ecology research group. I'd like to thank Josh, Raph, Yang, Tianjian, Shu, Pedro, and Thijs for all the interesting conversations, island biogeography or not. Even though I was not in Groningen often, it was always comforting to know there was such a warm community there whenever I turned up. Particularly, I'd like to thank Thijs for the discussions on all things trait-dependent diversification modelling and for the help running simulations for me. Your running of a few thousand simulations really saved me a lot of time and eased the stress of getting my fourth chapter together. Within Naturalis, I had the pleasure of being a member of two research groups: Functional Traits and Biodiversity Hotspots (as an honorary member). Thank you to Marta, Kasper, Clara, Anai, Vicky, Ajaree, Giovanni, Renske, Frank, Liliana, Vincent, Menno, Sebastian, Shweta, Yuanshu, Saijing, Isaak, Lizzie, Daniel, Adriana, Michael, Pepijn, Maarten, and Mathijs (as well as many more students past and present) for the interesting discussions and support. I'd also like to thank the two students I supervised, Noor and Lucas, for their contributions and help with this thesis; my second and third chapters, in particular, would look very different without their input. I hope my supervision was not too chaotic. Thank you to the lab technician for this project, Laura, for her continued patience with me whilst I learnt to work in a lab. Without your incredible contribution, I would probably still be doing DNA extractions to this day. I'd also like to thank members of the Tropical Botany research group and the herbarium team for all the interesting conversations and coffees over the years: Leni, Sylvia, Hans, Angelica, Peter, Susana, Roxali, Marnel, Nicolien, and Sander.

I also would like to thank all of the collaborators, co-authors, and curators who helped throughout my PhD. In particular, I'd like to thank those who helped with building the Canary Islands non-monocot tree of life in Chapter 2, whether it was sending plant material or figuring out the bioinformatics: Marcelino, Yurena, Bill, Juli, José Maria, Cristina, Ruth, Louis, Aguedo, Sara, Olivier, Raquel, Jairo, Stephan, Pablo, Alex, Javier, Steven (Meise Botanic Gardens, BR); Suzanne (Royal Botanical Gardens Edinburgh E); Ranee (Natural History Museum, London); Roderick (Hortus Botanicus, Leiden); Leopoldo (Real Jardín Botánico de Madrid, MA); Andreas (Naturhistorisches Museum Wien, W); Henry (The Botanical Museum of the Finnish Museum of Natural History, H); Patrik and Arne (Lund University Botanical Museum, LD); Filipe (Herbarium of the University of Coimbra, COI); Reto (United Herbaria, Z & ZT); Abelard and Rafael (Universidad de Sevilla), Juan (Royal Botanical Gardens, Kew), and Roxali, Susana, and Marnel. Without all of your valuable contributions, generating the data contained in this thesis would have been so much more difficult.

People say that your colleagues really make or break a job, and I can safely say the PhD/PostDoc community at Naturalis made my PhD experience what it was. I'm so happy to have met all of you. So thank you to PhDs/PDs past and present: Aafke, Ajaree, Anai, Annas, Andres, Annabelle, Benoit, Bob, Brigida, Burooj, Caroline, Cas, Clara, Corné, Deyi, Elsa, Eva, Giovanni, Iliana, Isaak, Jada, James, Jimmy, Kasper, Kevin, Lauranne, Leon, Lizzie, Luiz, Maartje, Mainah, Margret, Mariya, Merijn, Mike, Nemi, Nicholaas, Niels, Panos, Pasha, Remon, Rita, Roderick, Roel, Sebastian, Seppe, Sjoerd, Tim, Vicky, Werner, William, and Yanell. In particular, I want to thank the Gentlemen's Club: Mike, Werner, and Vicky. Our takeaway and drinks evenings always gave me something to look forward to and made Leiden feel like home. Mike, thanks for being one of the chillest people I've met, your Hawaiian shirts always brightened up the office. Werner, thank you for the long coffee breaks, the trips into the museum and collections, and of course, the biscuits. Vicky, a big thanks for all the fun moments around the museum, from trying to find Linnaeus, leaving you in the dark in the collections, and for always bringing reason to our stupid ideas. I'd also like to thank the 'drinking buddies' group in particular: Kevin, Andres, Roderick, Giovanni, Mahout, Iliana, Panos, Annabelle, and Tim (and everyone else who joined along the way). Thank you for all the Friday evening pub visits, trips to get Turkish, and of course, swimming in the canals. These were truly some of my favourite moments in Leiden.

Outside of work, there was one place I spent more time than any in Leiden: the climbing gym. Throughout my PhD, I managed to bring so many people along

with me to share my love for climbing, so thank you to everyone who came along. Particularly, I want to thank the regulars of the Naturalis Bouldering group (in no particular order): Marina, Esther, Pierre, Harold, Jada, Andres, Kevin, Ruthglen, James, Isaak, Corné, Niels, and Vicky (as well as all the others who came along). I hope you enjoyed our climbing sessions and trips to various Dutch cities as much as I did; I'll miss our weekly climbs greatly. My apologies to anyone who got injured from climbing! I'd also like to thank the friends I made from the gym: Bram, Maxi, Steijn, Jack, and Hannah. You were some of the first and closest friends I made in Leiden. It was a pleasure getting to know you all, and I will always remember our time spent climbing, going for food, and meeting up around the city. Leiden was never the same after you (mostly) all moved away. I'd also like to thank Connor, Bart, Pim, Niels and everyone else from climbing over the years. I wish you all the best and hope we can climb together again in future.

Before ending, I'd like to thank my family. Mum and Dad, thank you for all of your support over the years, for encouraging me to apply for this PhD in the first place, and for all of your advice and help. Your support throughout all of my years at various universities has meant the world to me. From helping me move various times to our weekly Sunday evening calls, I don't know where I'd be without your support. So, thank you for everything. Rich, thank you for always being there to listen. Whether it's about work problems, social life, or life in general, I always appreciate our conversations and your advice. I'm greatly looking forward to seeing more of you all now that I'm back from the Netherlands.

Finally, Grace. Your unwavering love and support over the last 4 ½ years means everything to me. Though neither of us planned to be apart for so long, and it can't have been easy with so many missed opportunities to be in the same country, you were always my rock to lean on. Your continuous support and belief in me kept me going, even during the hardest parts of my PhD, and made me feel like we were never really apart. You mean more to me than I can put into words. I can't wait to spend the rest of my life making it up to you for all this time spent apart. Thank you for everything.



Curriculum Vitae



Ryan Brewer was born in Guildford, Surrey, United Kingdom on the 30th of June 1996. After completing his secondary education at Therfield School, Leatherhead, he moved to Canterbury, Kent, for undergraduate studies in Wildlife Conservation at the University of Kent (2015-2018). During his undergraduate studies, Ryan undertook field courses in tropical ecology in Malaysian Borneo, and for his dissertation, studied population trends in the yellow-spotted river turtle (*Podocnemis unifilis*) in the Pacaya-Samiria National Reserve, Peru.

Ryan continued his studies at University College London (UCL), where he undertook a Master of Research in Biodiversity, Evolution, and Conservation (2017-2018). During his postgraduate studies, Ryan became fascinated with the fields of phylogenetics and biogeography, as well as the various factors affecting extinction risk in biodiversity. He completed two postgraduate research projects at UCL: one on the phylogenomics and biogeography of the caecilian family Caeciliidae (Amphibia: Gymnophiona) at the Natural History Museum, London, under the supervision of Mark Wilkinson and Jeff Streicher; and another on disease risk in big-headed turtles (*Platysternon megacephalum*) confiscated from the Vietnamese live-animal trade, with the Zoological Society of London, under the supervision of Tony Sainsbury and Ben Tapley.

After his postgraduate studies, Ryan joined the Royal Botanic Gardens, Kew, as a Species Conservation Researcher on the Sampled Red List Index 2020 team (2018-2021). Here, he researched the factors affecting extinction risk in monocot angiosperms and the family Fabaceae to assess how extinction risk for 2,000 species changes through time. The IUCN Red List assessments he generated were published on the IUCN Red List and the data featured in Kew's State of the World's Plants 2020 report 'Trends in Extinction Risk in Plants and Fungi' chapter and the associated peer-reviewed publication. This was Ryan's first foray into the world of botany and introduced him to the interesting, complex world of botanical evolutionary biology.

Continuing with his newfound interest in botany, Ryan was given the opportunity to begin his PhD studies at the Naturalis Biodiversity Centre, Leiden, The Netherlands, with the University of Groningen in 2021. On the phylogenomics, diversification, and biogeography of Canary Island angiosperms, Ryan was supervised by Frederic Lens, Luís Valente, and Rampal Etienne. Alongside his research, Ryan supervised two masters students; presented at the Flora Macaronesia Conference in 2022, International Botanic Congress in 2024, Netherlands Society for Evolutionary Biology in 2025, and Munich Botanic Garden Summer Seminar Series in 2025; and taught on courses at the University of Groningen and University of Leiden. Ryan was also a member of the PhD Council of the Naturalis Biodiversity Center from 2022 to 2023.

