Macroevolución y microevolución en sistemas insulares: el patrón Rand Flora en el género *Canarina*

Tesis doctoral Julio 2015
Mario Mairal Pisa

Macroevolución y microevolución en sistemas insulares: el patrón Rand Flora en el género *Canarina*

Macroevolution and microevolution in island systems: the Rand Flora pattern in the genus *Canarina*

Tesis doctoral Noviembre 2015

Mario Mairal Pisa

Universidad
Rey Juan Carlos

Departamento de Biología y Geología,
Física y Química Inorgánica y Analítica

REAL JARDÍN BOTÁNICO
CSIC

Universidad
Rey Juan Carlos

DOCUMENTO PARA EL ALUMNO
EXTRACTO DEL ACTA DE DOCTOR

Curso Académico 2015/2016

TITULO DE LA TESIS REALIZADA MACROEVOLUCIÓN Y MICROEVOLUCIÓN EN SISTEMAS INSULARES: EL PATRÓN RAND FLORA EN EL GÉNERO CANARINA.

BAJO LA DIRECCIÓN DEL Dr./a. ISABEL SANMARTÍN BASTIDA, MARIA LUISA ALARCÓN Y JUAN JOSÉ ALDASORO

DPTO. DONDE SE PRESENTÓ LA TESIS BIOLOGÍA, GEOLOGÍA, FÍSICA Y QUÍMICA INORGÁNICA Y ANALÍTICA

DPTO. RESPONSABLE DEL PROGRAMA CURSADO ....

TITULO DE DOCTOR/A POR LA UNIVERSIDAD REY JUAN CARLOS CON MENCION DE DOCTORADO INTERNACIONAL

Terminada la lectura y contestadas las objeciones formuladas, el Tribunal acordó otorgar a la Tesis Doctoral, presentada por:

D./Dña. MARIO MAIRAL PISA **la calificación de:**

SOBRESALIENTE "CUM LAUDE"

En MÓSTOLES **a** 6 **de** NOVIEMBRE **de 2015**

El/La Presidente/a,	**El/La Vocal,**	**El/La Vocal,**	**El/La Vocal,**
Fdo: PILAR CATALÁN	**Fdo:** JULI CAUJAPÉ	**Fdo:** MIKE THIV	**Fdo:** GONZALO NIETO

VICERRECTORADO DE TÍTULOS PROPIOS, FORMACIÓN CONTINUA, POSTGRADO Y RELACIONES INTERNACIONALES
MCMXCVI

Extracto del Acta del grado de Doctor

El/La Secretario/a del Tribunal,

Fdo.: LUIS GIMÉNEZ

Macroevolution and microevolution in island systems: the Rand Flora pattern in the genus *Canarina*

Macroevolución y microevolución en sistemas insulares: el patrón Rand Flora en el género *Canarina*

Tesis Doctoral
Mario Mairal Pisa
Noviembre 2015

Directora
Isabel Sanmartín Bastida

Codirectores
María Luisa Alarcón Cavero
Juan José Aldasoro Martín

Universidad Rey Juan Carlos

Departamento de Biología y Geología, Física y Química Inorgánica y Analítica

Este trabajo ha sido realizado en el Real Jardín Botánico de Madrid, CSIC, bajo la dirección de la doctora Isabel Sanmartín Bastida, científica titular del Departamento de Conservación y Biodiversidad, y la codirección de los doctores María Luisa Alarcón Cavero y Juan José Aldasoro Martín, ambos del Departamento de Biodiversidad del Instituto Botánico de Barcelona, CSIC-ICUB. Además, parte de los resultados fueron obtenidos en las estancias realizadas en el Jardín Botánico 'Viera y Clavijo'- Unidad Asociada CSIC (Cabildo de Gran Canaria; Gran Canaria, España) bajo la dirección del doctor Juli Caujapé Castells, en la Université de Lausanne (Lausanne, Suiza) bajo la dirección del doctor Nadir Álvarez y en la Université de Fribourg bajo la dirección del doctor Loïc Pellissier y la doctora Myriam Heuertz. La realización de la tesis ha sido posible gracias a una ayuda predoctoral de Formación del Personal Investigador (FPI, BES-2010-037261) concedida a Mario Mairal Pisa por el Ministerio de Ciencia e Innovación (ahora Ministerio de Economía y Competitividad) y a los proyectos i) Reconstrucción del Origen de la Rand Flora Afro-Mediterránea con datación filogenética e inferencia biogeográfica (proyecto CGL2009-13322-C03-01) y ii) The assembly of African floras from lineages to populations: effects of historical climate changes and species ecology (AFFLORA) (proyecto CGL2012-40129-C02-01) concedidos a Isabel Sanmartín por la misma entidad.

Esta tesis doctoral opta a la mención de Doctorado Internacional. Para poder optar a esta mención, el documento ha sido evaluado por los doctores Thomas Couvreur del Institut de Recherche pour le Developpement y Luis Valente de la Universität Potsdam.

Real Jardín Botánico
CSIC

CSIC

institut
botànic
Centre mixt
CSIC

Cabildo de
Gran Canaria

JARDIN
BOTANICO
CANARIO
VIERA Y
CLAVIJO

Unil
UNIL | Université de Lausanne

UNI
FR

UNIVERSITÉ DE FRIBOURG
UNIVERSITÄT FREIBURG

ÍNDICE / INDEX

CHAPTER 1

CHAPTER 2

CHAPTER 3

CHAPTER 4

CHAPTER 5

CHAPTER 6

AGRADECIMIENTOS / ACKNOWLEDGEMENTS

La participación de tantas personas en esta tesis tiene un valor especial, muchísima gente me ha acompañado y ayudado en su elaboración y con vosotros quiero compartir mi ilusión y alegría en estas páginas. ¡Aunque durante el camino ha habido obstáculos, gracias a vosotros no ha habido límites!

La realización de esta tesis ha sido posible gracias al apoyo de Isabel Sanmartín, que ha sobrepasado con creces lo que se espera de un director de tesis. A Isabel le agradezco su insistencia por formarme como investigador, transmitiéndome inquietudes y guiándome de forma firme a la vez que sutil, haciendo siempre gala de una gran ética y respeto. Asimismo le agradezco su disposición para escuchar y discutir nuevas ideas, con la destreza de valorarlas o enderezarlas. Gracias también por invitarme a participar en tus proyectos, por animarme y facilitarme la asistencia a congresos, así como por toda tu ayuda con la investigación y los artículos. Viendo el camino recorrido me siento enormemente afortunado y valorado. Isabel eres un modelo de inspiración y una directora de tesis ejemplar.

Gracias al esfuerzo de mis codirectores, María Luisa Alarcón y Juan José Aldasoro, por llevar a cabo este proyecto, sobre todo en los primeros pasos y al final de esta tesis. La ayuda de Marisa ha sido fundamental, guiándome y enseñándome de una forma estupenda y con todo su empeño. Espero que vuestro esfuerzo se vea recompensado en esta tesis.

Al resto de miembros del grupo durante los primeros años: Andrea Sánchez Meseguer y Cajsa Lisa Anderson. Y en la posterior etapa: Andrea Briega, Jon Cendoya, Lisa Pokorny, Ricarda Riina y Victoria Culshaw. Lisa me ha ayudado muy generosamente con sus críticas y consejos.

A mi tutor académico José María Iriondo por toda su ayuda para facilitarme los trámites con la tesis. A Pilar Catalán, Juli Caujapé, Adrián Escudero, Gonzalo Nieto, Mike Thiv, Luis Giménez y Juan Carlos Moreno por tener la suerte de poder contar con ellos para el tribunal de la tesis.

A todo el personal del Real Jardín Botánico por acogerme y darme acceso a sus instalaciones durante todos estos años en administración, dirección, herbario, invernaderos, laboratorio y biblioteca. En el laboratorio le estoy enormemente agradecido a Fátima Durán por su paciencia y dedicación inicial cuando no sabía ni coger una pipeta; largas horas de trabajo que no habrían sido lo mismo sin la ayuda y compañía de Alberto Herrero, Emilio Cano, Gemma Andreu, Guillermo Sanjuanbenito y Yolanda Ruiz. A Charo Noya, Concha Baranda, José Luis Fernández y Marta Fernández por su ayuda en el herbario. A Gonzalo Nieto por sus consejos y la revisión de manuscritos. Mi gratitud y admiración al trabajo de *Anthos*, GBIF y *Flora iberica*. A

tantos y tantos compañeros que me han brindado su compañía, consejos o ayuda a lo largo de todos estos años en el Real Jardín Botánico: Alejandro González, Alfredo Negrillo, Ana Alaminos, Ana Otero, Andreas Hilpold, Angélica Bello, Arantxa Martínez, Beatriz Guzmán, Beatriz Rumeu, Carlos Lado, Elena Amat, Eva García, Felipe Castilla, Francisco Pando, Gabriel Páez, Helena Velayos, Inés Álvarez, Irene Villa, Isabel Liberal, Isabel Marques, Javier Diéguez, Javier Fuertes, Javier Fernández, Jesús Muñoz, Joel Calvo, José Luis Blanco, José Ruiz, Juan Carlos Zamora, Katia Cezón, Kina García, Laura Lope, María Aguilar, María José Martínez, María Luisa Navarro, María Paz Martín, Maribel Moreno, Mario Fernández Mazuecos, Mauricio Velayos, Melisa Sarmiento, Pablo Vargas, Pedro Jiménez, Rosalía Piñeiro, Santos Cirujano, Sven Buerki, Vladimir Sandoval, Yurena Arjona y Zhang Li-Fang. A mis compañeros de fiambrera: Alberto Herrero, Alejandro Quintanar, Leopoldo Medina, Marta Fernández, Patricia Barberá, Toni Buira. A Alicia Agudo, por muchas cosas. Al equipo AFROLAR, Paco Cabezas y Maximiliano Fero. ¡Un especial cariño por las agradables coincidencias con mis amigas Luz Mauro y Laura Martín!

A Toda la gente que ayudó con las expediciones y con la recolección del Material: Colectivo Azaenegue, Jano, Juli, Félix y Moi. Recorrer los entresijos de la laurisilva de los Tilos (La Palma) y La Gomera fue una gozada con los guardas del parque. Nunca olvidaré la estancia en la selva de Los Tilos. A los cabildos de Gran Canaria, Tenerife, La Gomera, La Palma y El Hierro. ¡A mi gran amiga Yulán y a Oto, Sugo, Cande y Maligna por patrocinar la *First Paleo-islands expedition*! Cruzar toda la cornisa Cantábrica en busca de *Woodwardia* y selvas terciarias fue una aventura muy divertida en compañía de mi colega Rafa Marquina.

A todos los que influyeron anteriormente en mi andadura académica para llegar hasta aquí. Aurelio Martín fue una fuente de inspiración, un gran docente que consiguió transmitirme una gran pasión por la biología de islas y la conservación de la naturaleza. La botánica no sería lo mismo sin Margarita Costa, a la que echamos mucho de menos. Ana Rosa Burgaz por los primeros pinitos en la investigación. Gracias a todos esos profesores que supieron transmitir con pasión: Eduardo de Juana, Felipe Domínguez, Gisela Oliván y Victor Monserrat. Habría sido muy aburrido sin los grandes amigos siempre dispuestos a frikar de campo, ciencia y rock: Isaacina y Mikel.

Quiero agradecer al Ministerio de Ciencia e Innovación, posteriormente Ministerio de Economía y Competitividad, por el contrato predoctoral para la Formación del Personal Investigador (FPI). Estos mismos organismos nos concedieron dos ayudas para realizar estancias breves en dos centros de investigación y desarrollo, uno español y otro extranjero.

Juli Caujapé Castells me acogió estupendamente en dos ocasiones en el Jardín Botánico ‘Viera y Clavijo’, un sitio donde siempre apetece volver y en el que el aporte científico fue perfecto. El laboratorio fue especialmente acogedor en compañía de Ruth

Jaén y la flora canaria se dejo estudiar mejor con: Águedo Marrero, David Bramwell, Julia Pérez de Paz, Olga Fernández Palacios, Óscar Saturno, Pepe Naranjo y Magui Olangua. Especialmente enriquecedora fue toda la ayuda generosa y desinteresada de Carlos García Verdugo. Con Marta Fernández y Moisés Soto pasé los mejores momentos de la estancia ¡Os echo mucho de menos! GJ accedió a mis continuas peticiones para guiarme por recónditos lugares de Gran Canaria en busca de tesoros botánicos, a veces con algún tropezón, ¡como cuando casi nos rompimos el pámpano en Güayadeque! Con la Jurria Taguante y el garrote accedimos a todos los secretos de la isla ¡No me olvido de los grandes momentos con mi familia canaria! ¡Ángel, Iván, Aldemar, Eli y Jose!

En vue du séjour à l'étranger, on a choisi l'Université de Lausanne. Durant cette période en Suisse l'aide de Nadir Álvarez a été précieuse et essentielle afin de bien accomplir les objectifs. Nadir m'a accueilli au sein de son groupe de travail et nous a traité comme l'un d'entre eux; or les contacts et le déroulement des concepts a été fortement aisé, de même Il m'a aimablement permis concilier mes tâches envers l'Université de Freiburg. Il est nécessaire de remercier à Loïc Pellissier pour son soutien à l'Université de Freiburg, il est fort lucide et m'a dédié, patiemment, des longues heures. Le séjour a été confortable près de Romain Savary, Mila Pajkovic, Nils Arrigo et Karim Ghali. Pasé unos días estupendos en casa de Myriam Heuertz y Santiago González Martínez divirtiéndome mucho con sus niñas. La ayuda, ideas y disposición de Myriam están siempre que se necesitan.

Algunas de las personas con las agradezco haber compartido charlas enriquecedoras y buenos consejos en congresos o en formato electrónico son: Alan Forrest, Alfredo Valido, Carlos Escudero, Candelaria Rodríguez, Christina Schüßler, Cristina García, David Posada, Javier Valverde, Jens Olesen, Jorge Alfredo Reyes-Betancourt, Ioanna Manolopoulou, José María Fernández-Palacios, Jorge Lobo, Lea de Nascimento, Mark Mort, Mauricio Díaz-Granados, Marcelo Rodrigues, Néstor Pérez-Méndez, Nicholas Matzke, Peter Schönfelder, Paloma Torroba, Pilar Catalán, Rafael Medina, Priscila Rodríguez, Rodrigo Megía, Silvia Santamaría, Valerie Steen, Xavier Picó. A todo el blog bel Ojo de Darwin. Con Paulina Kondraskov compartimos inquietudes. Gracias a Mike Thiv por su amistad y por las apasionantes y estimulantes conversaciones ¡y excursiones macaronésicas! Gracias a todos los editores y revisores que contribuyeron a mejorar nuestros artículos, muy especialmente a los de *Molecular Ecology*. El curso de filogenias se hizo más interesante con Miquel Arnedo y Salvador Carranza. A Arnedo, además, gracias por la preparación y el premio que nos concedieron en el II CISA.

Gracias a toda la gente comprometida con el conocimiento y defensa de la naturaleza. Qué sería del campo de Madrid sin ARBA Bajo Jarama, las salidas de Grijalbo y del Elefante Rokoso. Visitar la naturaleza macaronésica con Rubén Barone es un regalo inolvidable, gracias por compartir tu sabiduría y por tu generosidad. Aquí no pueden

faltar las hazañas naturalistas con los avezados expedicionarios del CNAE: HerpetoRichi, NHVIII, OGE y Pozosinfondo. Juntos hemos sobrevivido a cañones de $CaCO_3$, polvorones en el Sáhara, neveros, fajas, varanos, Bisbales marroquís, delitos en Suiza, a MALO y a empachos de lazitos… siempre guiados por *MAXIMUS* (y *minimus* ¿dónde está? ¿dónde está?)...¡y las que quedan! A mi grandísimo amigo, el CNAE President Javi Aoiz le agradezco que haya conseguido despegarme del trabajo de vez en cuando en emocionantes expediciones en busca de urogallos, lobos, osos, Ruaviejas y tascas madrileñas; así como por todos sus consejos editoriales.

De los mejores momentos de estos años y de mi vida han transcurrido con mis amigos; me siento muy afortunado de tener tantos y tan nobles, a muchos ya los he nombrado y sois tantos que no puedo dedicar una línea a cada uno. A mis amigos Freiji y OGE gracias por la revisión del *Canarinanomicon*. Con el CNNM hemos estado a punto de morir a causa de múltiples shocks causados por el culto a la risa: Cherman, Babía aka Astralestas, Moreno y los Vainaloquenses; Pablo Galván, más conocido como Abuelas CP y otros motes prohibidos, gracias por tu ayuda con el diseño de portada. Durante estos años hemos escuchado 20000 veces *shoot to thrill*, tirado al suelo 2 millones de barajas, celebrado 12 CMF, acabado con las reservas del Woodstock... A Isabel gracias por los zaborro-consejos. A mis amigos de Huesca de toda la vida. A Elena POSITIVE SELECTION por llenarme de alegría. A mis queridos amigos biólogos: Roci, Elvira, Pantxo, Fran, Elis, María Niño, Estefanía, Gisela, Sandra, Irenika, Hairen, Peeedro, Senda, Cuacos, Gabi, Fany. A Jus. A Miguel Ángel Granado. A Javi Ruiz. A la gente del Pardo, a Figurines Cucales and Chuan, a El Increíble PolK. A Rodri, Rubén y al Cenas. No hay emoticón para expresar lo que siento, con todo mi cariño ¡¡gracias a todos!!

A los que se han leído o se van a leer esta tesis.

El apoyo y la paciencia de Viky han sido esenciales para acabar la tesis. Gracias por tu inmensa ayuda, correcciones y críticas. Y gracias sobre todo por soportarme y darme ánimos en los momentos más difíciles de esta tesis, por tu cariño y por quererme tanto.

Reservo el final para los más importantes, mi familia. Gracias a mis padres por creer en lo que hago y aguantar la locura que supone hacer una tesis. Gracias a ellos he podido perseguir el sueño de convertirme en biólogo y seguir haciendo lo que más me gusta. ¡A Luchi, que la quiero mucho!

En Huesca a 21 de junio de 2015.

RESUMEN

I. Summary (in English)

Global change and the current biodiversity crisis have increased the need to understand the effect of climate change on patterns of biodiversity distribution and the mechanisms or factors underlying them. To understand the origin of these biodiversity patterns, scientists have resorted to macroevolutionary explanations studying the relationships between organisms and their habitat at large spatial and temporal scales, and microevolutionary explanations, studying the basis of evolutionary change within species or populations locally and over shorter time spans. One of the best-studied effects of the impact of climate change over diversity patterns is Pleistocene glaciations; however, events of global drought or aridification are more interesting under the current climatic scenario. For this, the African continent offers an ideal scenario. Africa has been considered one of the exceptions to the global Latitudinal Diversity Gradient

because a dip in biodiversity can be observed in many organisms as one move from the temperate to the tropical regions; species poverty in this continent has been sometimes attributed to extinction, mediated by climate change (gradual aridification) that also presumably generated large-scale, intracontinental disjunctions. In this sense, the so-called Rand Flora pattern, which relates angiosperm lineages that share similar disjunct distributions across the margins of the African continent and adjacent islands (Macaronesia-west Africa, the Horn of Africa-South Arabia, east Africa and Southern Africa), offers us a unique opportunity to study the effects of rapid climate change on patterns of biotic assembly. One of the most extreme representatives of this pattern is the genus *Canarina*. Belonging to the tribe Platycodoneae, a basal lineage within the Campanulaceae family, this genus comprises only three species with a widely disjunct distribution: the island-endemic *Canarina canariensis* is associated to the unique laurel habitat of the Canary Islands facing northwest Africa, while at the other end of the Sahara Desert, *Canarina abyssinica* and *Canarina eminii* are inhabitants of the East African Mountain forests. Thus, the distribution of *Canarina* species seems to be restricted to what is considered the last remnants of a subtropical vegetation that was presumably widespread throughout Africa, and which are now confined to island ecosystems on the continental margins of Africa: the Canarian species on oceanic islands, and the East African species "within-continent" islands or sky islands –geographically isolated high-altitude habitats occurring alongside different mountain ranges. The overall objective of this thesis is to study the evolutionary history of the genus *Canarina* to try to infer the macro- and microevolutionary processes that generated its current distribution. We have used an integrative and multidisciplinary approach encompassing macroevolutionary (phylogeny, biogeography, ecology) and microevolutionary (phylogeography, demography, population genetics) approaches. The phylogenetic position of the genus, its current habitat restrictions, and its extreme intracontinental disjunction, point out of *Canarina* as a potential relict, a suitable candidate for the study of patterns and processes resulting from a gradual climate change scenario. Additionally, the unique distribution of *Canarina* will allow us to study the evolutionary role played by both oceanic and continental (sky) islands in generating and maintaining genetic diversity levels.

In **Chapter 1** we use *Canarina* as a study system to try to explain the evolutionary origin of the Rand Flora disjunction and to understand the effects of extinction mediated by climate change- in the shaping of biodiversity patterns at a continental level. Two biogeographic hypotheses have been proposed in the past to explain this disjunction: i) vicariance and climate-driven extinction of an ancient widespread African macroflora, or ii) long-distance dispersal events between geographically isolated regions followed by in situ speciation. Using a combination of nuclear and chloroplast markers, we reconstructed the phylogeny of the genus, estimated lineage divergence times, and performed biogeographical and paleoclimatic reconstructions through time. Our results confirmed the position of *Canarina* within Platycodoneae and solved interspecific relationships within the genus. We show that *Canarina* originated in eastern Africa from an Asian ancestor in the Middle Miocene. The species *C. abyssinica* separated first (8 Ma), and is sister to a clade formed by disjunct species *C. eminii* and *C. canariensis*, which apparently diverged in the Late Miocene (7 Ma). We solved the lack of phylogenetically close fossils to calibrate our divergence estimates by using a "nested dating" approach, in which, a phylogenetic, higher-level partition (above or at species level), itself calibrated with external evidence (fossil or molecular), was used to constrain the molecular clock rate of additional linked data sets containing population-level data. Our temporal, spatial and climatic niche evidences indicate that the current intracontinental disjunction of almost 7000 km across the Sahara and 7 million years divergence might be explained by an event of climate-driven vicariance, in which an originally more widespread distribution across North Africa was fragmented by aridification events and later extinction of intermediate populations, leading to the currently disjunct and restricted distributions in the margins of the continent. In **Chapter 2** we reconstruct phylogeographical patterns and processes in *Canarina eminii,* an epiphyte endemic to the Eastern Afromontane forest. We use three chloroplast regions and AFLP nuclear markers to infer genetic diversity, structure, divergence times, barriers to gene flow, and phylogeographic relationships within this species. Results reveal a strong phylogeographic structure with two separate lineages at each side of the Rift Valley. Geological and climatic processes in this region of Africa, particularly the rifting of the African and Arabian plates along the Great Rift Valley, and the concomitant climatic changes, could be acting as an early barrier to gene flow, favouring the spatial and genetic isolation of populations on either side of the barrier. Our study thus stresses the importance of the Great Rift Valley in the assembly of biodiversity patterns in Eastern Africa, and supports the famous biogeographic "axiom" that geographic barriers and biota often evolve together. In addition, we detected high levels of genetic exclusivity and restricted gene flow between sky islands, which can be attributed to rapid Pleistocene climatic events that caused the expansion-contraction of the Afromontane forest during the Ice Ages. We stress the importance of sky islands as refuges and cradles of high levels of exclusive and critically endangered diversity, restricted to different patches of the Afromontane forest. In **Chapter 3**, we reconstruct phylogeographical patterns and their underlying processes in *Canarina canariensis*, an endemic of the Canary Islands associated with the main laurel forests in the archipelago. We use chloroplast regions and AFLP nuclear markers to examine spatial patterns of genetic structure, reconstruct migration events between islands and infer ancestral areas and refugia. We found a strong geographic population structure, with a first divergence around 0.8 Ma that has Tenerife as its central axis and divides Canarian populations into eastern and western clades. Genetic diversity was greatest in the "paleo-islands" of Tenerife (three geologically stable massifs: Anaga, Teno and Roque del Conde); these areas were also inferred as the ancestral location of migrant alleles towards other disturbed areas within Tenerife or the nearby islands. Our evidence suggests that barriers between oceanic islands have been less important in structuring levels of genetic diversity than geological barriers within each island. The strong structure detected as well as the presence of extinct ancestral haplotypes suggest that, following the merging of the three paleo-islands that formed Tenerife 3.5 Ma ago, the island underwent major volcanic eruptions and large landslides that apparently played a key role as genetic barriers. The topographic complexity and geological stability of the paleo-islands have favoured the preservation of genotypes and high levels of intraspecific diversity, acting both as refugia against the effects of catastrophic volcanic extinction, and as cradles and sources of new genetic diversity towards other areas. This confers these areas the characteristics of "phylogeographical hotspots", regions whose preservation should be a priority. To complement the previous two chapters, in **Chapter 4** we evaluate the need for conservation of the last fragments of subtropical vegetation remaining in Africa and adjacent islands: the laurisilva and the Afromontane forest. For this, we studied the recent demography of *C. eminii* and *C. canariensis* using microsatellite markers and ABC simulations (Approximate Bayesian Computation). Our results showed a strong decline in the population demography of these species. The downward spiral in the habitat degradation of these species, closely linked to these subtropical forests, calls for urgent action for the conservation of these relict ecosystems that harbour exceptional biodiversity. In **Chapter 5**, as an extension of Chapter 1, we expanded the research conducted within *Canarina* to a wide range of Rand Flora lineages, with the aim of establishing a temporal framework and detect the phylogenetic signature of extinction in this pattern. Using macroevolutionary and phylogenetic dating tools, we showed that, like *Canarina*, most disjunctions within Rand Flora lineages fall within the Miocene and Pliocene periods, coinciding with the onset of a major aridification trend, still ongoing, in Africa. The Rand Flora pattern seems to fit the definition of "biogeographic pseudocongruence", in which the same biogeographic pattern arises at different times (temporal incongruence) in different lineages. We also observed a gradation in the age of the disjunctions that seems to be related to the different ecological affinities of each lineage, with subhumid taxa diverging earlier and those with more xeric affinities diverging later. This hypothesis is investigated in more detail in **Chapter 6**, where we quantified and reconstructed the climatic niche of 16 Rand Flora lineages using presence data, ecological niche modelling (Species Distribution Models, SDMs), and paleoclimatic reconstructions spanning from the Mid Miocene to the Pliocene. Paleoclimate modelling revealed the existence of former "climate corridors" that connected the eastern and western sides of Africa. Furthermore, we found a correlation between the age of divergence and the climatic tolerances of each lineage. In all, our results support the hypothesis of a wider distribution range in most (many) Rand Flora lineages, which could have been extended across central North Africa. The subsequent range contraction due to the progressive aridification of the continent since the Neogene would have led to the gradual divergence of lineages at a rate that matched the climatic niche of the species. Thus, the evolutionary origins of the Rand Flora pattern would support the key role that climate change, vicariance, and extinction played in the shaping of global and regional biodiversity patterns.

II. Síntesis (en español)

El cambio global y la actual crisis de la biodiversidad han incrementado la necesidad de comprender el efecto que tienen los cambios climáticos en los patrones y mecanismos que determinan la distribución de la biodiversidad. Para entender el origen de estos patrones de diversidad, los científicos han recurrido a explicaciones macroevolutivas que estudian las relaciones entre los organismos y su hábitat a grandes escalas espaciales y temporales, y a explicaciones microevolutivas, que estudian la base del cambio evolutivo dentro de especies o en poblaciones, localmente y durante periodos de tiempo más cortos. Uno de los efectos mejor estudiados del impacto del cambio climático sobre los patrones de biodiversidad son las glaciaciones del Pleistoceno; sin embargo, el estudio de eventos de sequía o aridificación global se hace interesante dado el escenario climático actual. Para esto, el continente africano ofrece un escenario ideal. África ha sido considerado una de las excepciones al Gradiente Latitudinal de Diversidad global, ya que en muchos organismos puede observarse una disminución en la biodiversidad desde las regiones templadas a las tropicales; en ocasiones la pobreza de especies en este continente ha sido atribuida a una extinción mediada por cambio climático (aridificación gradual) que, presumiblemente, también ha generado disyunciones intracontinentales a gran escala. Respecto a esto, el patrón fitogeográfico conocido como Rand Flora, que relaciona linajes de angiospermas que comparten distribuciones disyuntas similares entre los márgenes del continente africano e islas adyacentes (Macaronesia-noroeste de África, Cuerno de África-sur de Arabia, este de África y Sur de África), ofrece una oportunidad única para estudiar los efectos del cambio climático en los patrones de biodiversidad. Uno de los representantes más extremos de este patrón es el género *Canarina*. Este género pertenece a la tribu de las Platycodoneae, un linaje basal de la Familia Campanulaceae, y está compuesto solamente por tres especies con una amplia disyunción: el endemismo insular *Canarina canariensis* está asociado a los bosques lauroides de las Islas Canarias en el noroeste de África, mientras que al otro lado del desierto del Sáhara, *Canarina abyssinica* y *Canarina eminii* crecen en las montañas del este de África. Esta distribución de las especies de *Canarina* parece estar restringida a lo que se consideran los últimos retazos de una vegetación subtropical que estuvo presumiblemente extendida a lo largo de África, pero que actualmente ha quedado confinada en ecosistemas insulares en los bordes continentales de África: la especie canaria en islas oceánicas y las especies Afromontanas en islas dentro de un continente o *sky islands* –hábitats de altura geográficamente aislados, ocupando diferentes rangos montañosos. El objetivo general de esta tesis es estudiar la historia evolutiva del género *Canarina* para tratar de inferir los patrones y procesos macro- y microevolutivos que han generado su distribución actual. Para ello hemos utilizado un enfoque integrador y multidisciplinar que abarca aproximaciones macroevolutivas (filogenia, biogeografía, ecología) y microevolutivas (filogeografía, demografía, genética poblacional). La posición filogenética del género, sus actuales restricciones de hábitat y su extrema disyunción intracontinental, señalan a *Canarina* como un potencial relicto, un candidato idóneo para el estudio de patrones y procesos resultantes de un escenario gradual de cambio climático. Adicionalmente, la singular distribución de *Canarina* nos va a permitir estudiar el papel evolutivo tanto de islas oceánicas como de islas continentales (*sky islands*), en los niveles de diversidad genética.

En el **Capítulo 1** utilizamos el género *Canarina* como caso de estudio para tratar de explicar el origen evolutivo de la disyunción Rand Flora y comprender los efectos de la extinción mediada por cambio climático en la formación de los patrones de biodiversidad dentro de un continente. Dos hipótesis biogeográficas se han propuesto en el pasado para explicar esta disyunción: i) vicarianza y extinción por cambio climático de una antigua *macroflora* africana ampliamente distribuida por el continente, o ii) eventos de dispersión a larga distancia entre regiones aisladas geográficamente y especiación *in situ*. Utilizamos una combinación de marcadores nucleares y cloroplásticos, reconstruimos la filogenia del género, estimamos los tiempos de divergencia de los linajes y realizamos reconstrucciones biogeográficas y paleoclimáticas a lo largo del tiempo. Nuestros resultados confirman la posición de *Canarina* dentro de las Platycodoneae y resuelven las relaciones interespecíficas dentro del género. Mostramos que *Canarina* se originó en el este de África a partir de un ancestro asiático en el Mioceno Medio. La especie *C. abyssinica* se separó primero (c. 8 Ma), y es hermana de un clado formado por las especies disyuntas *C. eminii* y *C. canariensis*, que divergieron posteriormente en el Mioceno Tardío (c. 7 Ma). Solucionamos la falta de fósiles filogenéticamente cercanos a nuestro grupo de estudio para calibrar nuestras estimas de divergencia, utilizando una “datación anidada” en la que una partición filogenética de nivel superior (por encima o a nivel de especies), calibrada con evidencia externa (fósil o molecular), se utilizó para informar al reloj molecular de otra partición de nivel inferior (a nivel de poblaciones). Nuestras evidencias temporales, espaciales y de nicho climático indican que la actual disyunción intracontinental de casi 7000 km a través del Sáhara y 7 millones de años de divergencia podría ser explicada por un evento de vicarianza a causa del cambio climático, en el que una distribución original ampliamente distribuida a lo largo del norte de África fue fragmentada a causa de eventos de aridificación, y más tarde, extinción de las poblaciones intermedias, dando lugar a la actual disyunción y distribución restringida en los márgenes del continente. En el **Capítulo 2** reconstruimos los patrones y procesos filogeográficos en *Canarina eminii*, un endemismo epífito del bosque Afromontano del este de África. Utilizamos regiones cloroplásticas y marcadores nucleares AFLP para inferir la diversidad genética, estructura, tiempos de divergencia, barreras de flujo génico y relaciones filogeográficas en esta especie. Los resultados revelan una fuerte estructura filogeográfica con dos linajes separados a ambos lados del Valle del Rift. Los procesos geológicos y climáticos en esta región de África, en concreto, el agrietamiento de las placas Africana y Arábiga a lo largo del Gran Valle del Rift y el consiguiente cambio en las condiciones climáticas, podrían haber actuado como una barrera incipiente para el flujo génico, lo que habría favorecido el aislamiento espacial y genético a cada lado de la barrera. Así, nuestro estudio destaca la importancia del Gran Valle del Rift en el ensamblaje de patrones de biodiversidad en el este de África, y apoya el famoso "axioma biogeográfico" que dice que las barreras geográficas y las biotas evolucionan juntas. Además, detectamos elevados niveles de exclusividad genética y restricción del flujo génico entre *sky islands*, lo que puede ser atribuido a rápidos eventos climáticos en el Pleistoceno que causaron la expansión-contracción del bosque Afromontano durante las glaciaciones. Destacamos la importancia de los diferentes rangos montañosos (altas mesetas y *sky islands*) como refugios y cunas de altos niveles de diversidad exclusiva y críticamente amenazada, restringida a cada uno de los parches del bosque Afromontano. En el **Capítulo 3** reconstruimos los patrones y procesos filogeográficos de *Canarina canariensis*, un endemismo de las Islas Canarias asociado a los principales bosques de laurisilva del archipiélago. Utilizamos regiones cloroplásticas y marcadores AFLP para examinar patrones espaciales de estructura genética, estudios demográficos, reconstruir eventos de migración entre islas e inferir áreas ancestrales y refugios de diversidad. Nuestros resultados muestran una fuerte estructura geográfica poblacional con una divergencia hace 0.8 Ma que tiene a Tenerife como eje central, dividiendo las poblaciones canarias en un clado al este y otro al oeste. Encontramos altos niveles de diversidad y exclusividad genética en las tres paleo-islas de Tenerife (Anaga, Teno y Roque del Conde, antiguos macizos geológicamente estables), las cuales a su vez identificamos como áreas ancestrales desde las que migraron haplotipos hacia otras zonas de Tenerife y otras islas del archipiélago. Nuestras evidencias sugieren que las barreras oceánicas entre islas han sido menos importantes en estructurar los niveles de diversidad que las barreras geológicas dentro de cada isla. La fuerte estructura detectada junto a la presencia de haplotipos ancestrales extintos, sugiere que, tras la unión de las tres paleo-islas que formaron inicialmente Tenerife hace 3.5 Ma, la isla experimentó importantes erupciones volcánicas y grandes deslizamientos que probablemente jugaron un papel clave en la formación de barreras genéticas dentro de Tenerife. La complejidad topográfica y la estabilidad geológica de las paleo-islas habrían favorecido la preservación de genotipos y de altos niveles de diversidad intraespecífica. Estas áreas actuarían como refugio frente a fenómenos volcánicos catastróficos, a la vez que como cunas y fuentes de nueva diversidad genética hacia otras áreas. Esto otorga a éstas áreas la característica de *hotspots filogeográficos*, regiones cuya conservación debería ser prioritaria. Como complemento a los dos capítulos anteriores, en el **Capítulo 4** evaluamos la necesidad de conservación de los últimos fragmentos de vegetación subtropical en África e islas adyacentes, es decir, la laurisilva y el bosque Afromontano. Para esto, estudiamos la demografía reciente de *C. eminii* y *C. canariensis*. Utilizamos microsatélites como marcadores y simulaciones ABC (*Approximate Bayesian Computation*) que muestran un fuerte declive en la demografía poblacional de estas especies. La dramática espiral de retroceso a la que se ven sometidas estas poblaciones, estrechamente ligadas a estos bosques subtropicales, reclama acciones de conservación urgentes para la conservación de estos ecosistemas relictos que albergan una biodiversidad única. En el **Capítulo 5**, como extensión del Capítulo 1, ampliamos la investigación llevada a cabo con *Canarina* a un amplio abanico de linajes con patrón Rand Flora, con el objetivo de establecer un marco temporal y detectar señales de extinción en las filogenias de este patrón. Utilizamos herramientas macroevolutivas y de datación filogenética, y demostramos que al igual que en *Canarina*, en los linajes Rand Flora la mayoría de disyunciones se dan entre edades que abarcan desde el Mioceno Medio hasta el Plioceno, y coinciden con el inicio de un proceso de aridificación general,

todavía en curso en África. El patrón Rand flora encajaría por tanto con la definición de *pseudocongruencia biogeográfica*, en el que el mismo patrón de distribución surge en diferentes momentos (incongruencia temporal) en diferentes linajes. También observamos una gradación en la edad de la disyunción que parece estar relacionada con las diferentes afinidades ecológicas de cada linaje; es decir, los linajes subtropicales mostraron divergencias más antiguas, mientras que los linajes adaptados a condiciones más xéricas mostraron disyunciones más recientes. Esta hipótesis se profundiza en el **Capítulo 6**, donde se cuantifica y se reconstruye el nicho climático de 16 linajes Rand Flora utilizando datos de presencia, modelización de nicho ecológico (*Species Distribution Models*, SDMs) y reconstrucciones paleoclimáticas que abarcan desde el Mioceno medio hasta el Plioceno. Las modelizaciones paleoclimáticas revelaron la existencia de antiguos "corredores climáticos" que conectarían los lados este y oeste de África. Además, encontramos una correlación entre la edad de divergencia y las tolerancias climáticas de cada linaje. Nuestros resultados soportan la hipótesis de un rango de distribución más amplio en la mayoría de linajes Rand flora, los cuales podrían haberse extendido a lo largo del centro del norte de África. La consiguiente contracción de rango debido a la progresiva aridificación del continente desde el Neógeno, habría dado lugar a la divergencia de linajes a un ritmo que se ajusta al nicho climático de las especies. De esta manera, los orígenes evolutivos del patrón Rand Flora apoyarían el papel clave que el cambio climático, la vicarianza y la extinción jugaron en la formación de los patrones de biodiversidad globales y regionales.

III. Antecedentes

En este apartado se introduce el marco conceptual de esta tesis. Para entender el origen de un patrón de especulación hay que tener en cuenta tanto la escala macroevolutiva como la microevolutiva. De manera que primero presentamos la biogeografía, sus hipótesis principales (dispersión vs. vicarianza) y los patrones de distribución, prestando especial atención al patrón Rand Flora. Posteriormente nos centramos en la filogeografía como puente entre la macroevolución y la microevolución, haciendo especial énfasis en la búsqueda de refugios de diversidad intraespecífica. Como la biodiversidad objeto de estudio se sitúa en islas de hábitat, se establece un marco en torno a la biología de islas. Asimismo, se detalla nuestro caso de estudio, haciendo hincapié en las islas oceánicas (Islas Canarias) y las *sky islands* (archipiélago Afromontano). Por último presentamos el sistema de estudio: el género *Canarina.*

1. Biogeografía: dispersión vs. vicarianza

La Biogeografía es la disciplina de la biología que estudia los procesos causales –históricos y geológicos– que determinan la distribución espacial de los organismos. Esta tesis se enfoca en el campo de la biogeografía histórica, que pretende explicar la distribución geográfica de los organismos en función de su historia evolutiva (Sanmartín 2012). La biodiversidad terrestre no está distribuida de forma homogénea, sino que tiende a acumularse en determinadas regiones y biomas, de manera que forma patrones de distribución espacial a distintas escalas geográficas. Uno de los patrones más ilustrativos que se puede observar actualmente es el aumento de la riqueza de especies desde los polos hacia el Ecuador, patrón conocido como Gradiente Latitudinal de Diversidad (Hawkins 2001, Mittelbach *et al.* 2007). La biogeografía histórica ha explicado tradicionalmente estos patrones como el resultado de dos procesos o mecanismos biogeográficos opuestos: dispersión y vicarianza. La dispersión

postula que los organismos tienen su origen en un área ancestral o centro de origen, desde donde colonizan otras áreas atravesando barreras geográficas, produciéndose más tarde procesos de aislamiento y especiación alopátrica. La vicarianza se produce tras la fragmentación del área de distribución de una especie por causas geológicas o climáticas. Sus consecuencias, finalmente, son similares: la formación de barreras que separan a las poblaciones y acaban generando especiación alopátrica. Durante varios siglos la explicación *dispersionista*, inspirada en los relatos bíblicos, fue la dominante, postulándose que tras el Diluvio Universal todos los animales del Arca de Noé habrían repoblado la Tierra desde su refugio en el Monte Ararat. Tales ideas permanecieron hasta la época del naturalista sueco Carl von Linneo, quien además creía que las especies permanecían inmutables y colonizaban otras áreas de acuerdo con sus preferencias ecológicas. En la misma época el Conde de Buffon (1787) publicaba su extensísima *Histoire naturelle* dando ya inicio a una biogeografía histórica. Posteriormente Lyell en su *Principles of Geology* (1830) argumentaba también el predominio de la dispersión para explicar la distribución de las especies. La entrada en escena de la Teoría de la selección natural de Darwin y Wallace (1859) rompe con la idea de la inmutabilidad de las especies, y propone que los organismos se originan en un centro de origen desde el que se dispersan a otras áreas de manera "improbable", atravesando una barrera y diferenciándose posteriormente. Todas estas hipótesis *dispersionistas* dominaron la biogeografía hasta mediados del siglo XX. Entonces, la fusión de la teoría de la Tectónica de Placas y la metodología cladista de Willi Hennig introdujeron un cambio de paradigma en la disciplina con el concepto de la vicarianza, aunque la panbiogeografía de Croizat (1958) ya se había adelantado con su famosa frase "la Vida y la Tierra evolucionan juntas". Las dos últimas décadas se han caracterizado por grandes avances en el desarrollo de métodos de inferencia biogeográfica que nos permitan reconstruir los patrones de biodiversidad espacial y temporal a escala de linajes y biotas. En

particular, el desarrollo de herramientas filogenéticas moleculares ha permitido un cambio de paradigma. La dispersión ha recobrado su importancia como proceso que puede producir patrones biogeográficos congruentes entre linajes, por ejemplo, mediada por vientos y corrientes oceánicas. La biogeografía vicariante, enfocada en patrones de distribución, ha sido reemplazada por una biogeografía basada en procesos (Ronquist 2003), como la extinción y la duplicación, y posteriormente por una biogeografía probabilística que modela la evolución biogeográfica como un proceso estocástico con variables o parámetros – procesos biogeográficos que determinan la tasa de cambio o migración entre rangos geográficos (Ronquist & Sanmartín 2011; Sanmartín 2012). Los últimos desafíos tratan de incorporar el registro fósil y las preferencias ecológicas con técnicas de modelización de nicho (Meseguer *et al.* 2015), como una fuente de evidencia de distribuciones pasadas o linajes extintos.

1.1. Patrones de diversidad: las disyunciones biogeográficas

El interés por entender la distribución global de la biodiversidad ha sido objeto de estudio durante siglos por parte de naturalistas y científicos. A día de hoy, este interés ha aumentado a causa de la necesidad de entender como la biodiversidad puede cambiar debido al actual escenario de cambio climático. Entre los patrones que más han llamado la atención de los biogeógrafos se encuentran las disyunciones –linajes supuestamente emparentados que habitan áreas sin continuidad geográfica, en ocasiones separadas miles de kilómetros (Treviranus 1803; de Candolle 1820; Hooker 1853, 1860; Raven 1972). Asimismo, la distribución de grupos australes como las aves ratites (avestruz, kiwi, casuario) o las especies del género de hayas *Nothofagus* en Sudamérica, Australia o Nueva Zelanda han sido la base de un intenso debate entre *dispersionistas* y *vicariancistas* (Croizat 1958; Darlington 1965; Heads 1999; Sanmartín & Ronquist 2004). En el Hemisferio Norte, una de las disyunciones más conocidas es la distribución *Madre-*

Tethyense de Axelrod**:** una antigua flora que se distribuía por los márgenes del Mar de Tetis y que habría quedado relegada a las regiones actuales del sur de Eurasia y oeste de Norteamérica (Figura 1a). Esto se explicaría por procesos geológicos y climáticos (vicarianza) entre Eurasia y Norteamérica a finales del Cenozoico (Kadereit & Baldwin 2012; Vargas *et al*. 2014) o alternativamente por eventos más recientes de dispersión transoceánica (Pole 1994; Givnish & Renner 2004). Otro ejemplo de disyunción intercontinental es, la llamada disyunción "Asa-Gray" o "boreotropical" (Figura 1c): una antigua flora boreotropical que migraría por puentes transcontinentales a través del Atlántico y del Estrecho de Bering, hoy reducida a linajes de plantas emparentados (más de 120 géneros de plantas; Zhengyi 1983) con una distribución disyunta entre el este de Asia y el este de Norteamérica (Figura 1b; Tiffney 1985; Wen & Ickert-Bond 2009). Evidencias paleontológicas (registro fósil), geológicas y filogenéticas apoyan la hipótesis de una flora "boreotropical" formada por elementos tropicales de hoja perenne y otros caducifolios. Esta flora se habría extendido a lo largo del Hemisferio Norte a principios del Cenozoico mediante los puentes de Beringia y del Atlántico Norte (el *North Atlantic Land Bridge*, Tiffney 1985; Wen 1999). El enfriamiento global del clima y los cambios en la configuración de las placas habrían interrumpido estas conexiones y aislado fragmentos de este bosque relicto en los extremos de su distribución. Otro ejemplo clásico son las disyunciones “anfitropicales”, que se dan entre las zonas templadas del norte y el sur de Sudamérica (Figura 1d), explicadas principalmente por dispersiones a larga distancia, con abundantes linajes que migraron de Norteamérica a Sudamérica (Wen & Ickert-Bond 2009).

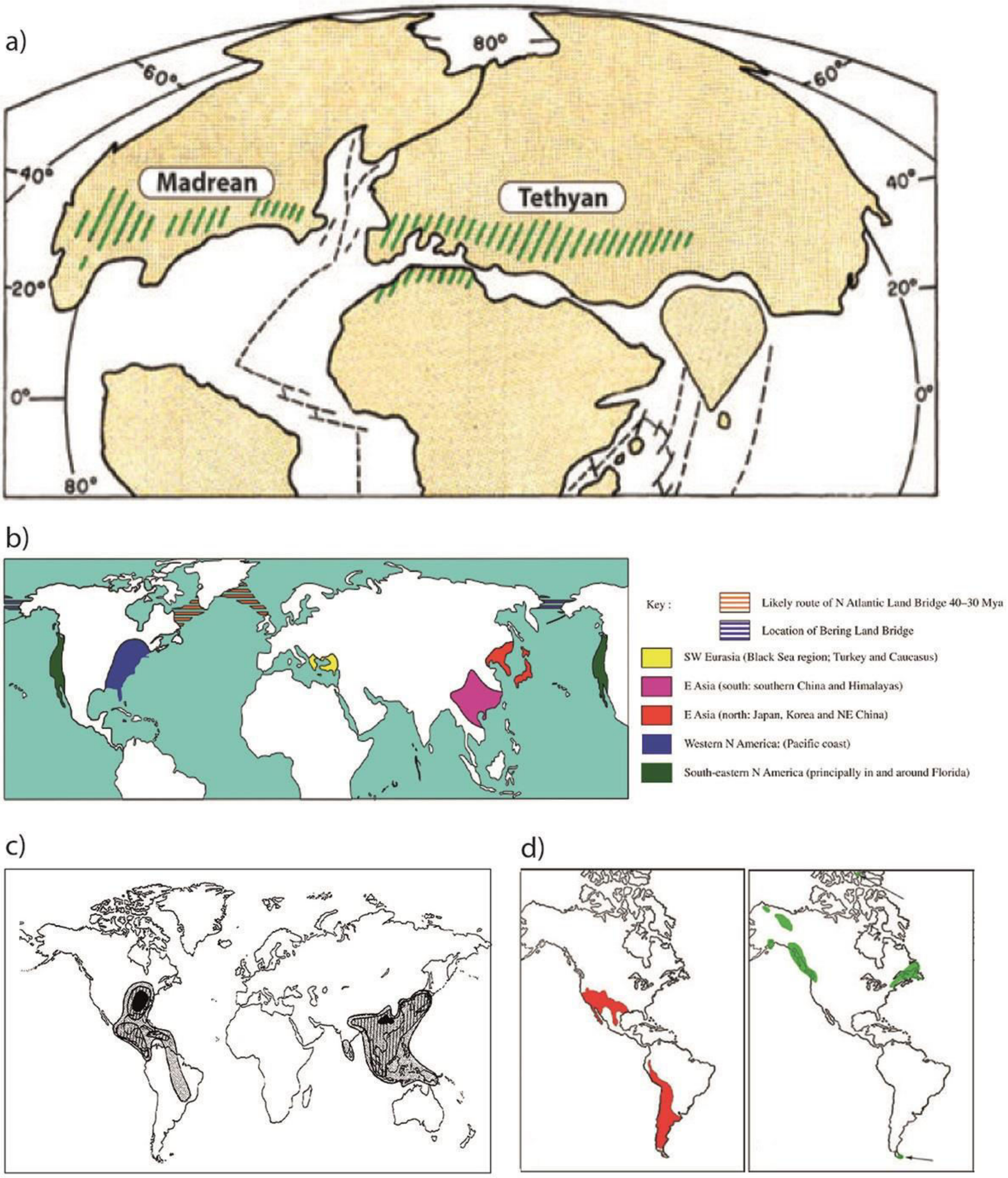


Figura 1 Disyunciones biogeográficas intercontinentales. a) Hipótesis Madreano-Tetiana de Axelrod (1975) que hipotetiza un cinturón casi continuo de vegetación perenne de hoja esclerófila, extendiéndose desde el oeste de América del Norte a Asia Central a principios del Terciario (tomado de Wen & Ickert-Bond 2009). b) Principales áreas de flora del terciario y corredores de vegetación entre Eurasia y Norteamérica (tomado de Milne 2006). El puente de vegetación en el Atlántico Norte (en naranja) es el "puente Groenlandia - Feroe" (Xiang *et al.* 2005), la única conexión que podría haber persistido después de hace 40Ma (ver Milne & Abbott 2002). Antes de este puente, una segunda conexión más al norte existió entre Groenlandia y Escandinavia (Tiffney 1985). El puente de tierra en el estrecho de Bering (en azul) existió durante la mayoría del periodo Terciario, moviéndose progresivamente hacia el sur (Milne & Abbott, 2002). c) Disyunción Asa-Gray mostrando la distribución de las magnoliáceas (sombreado): *Liriodendron* (sólido) y *Magnolia* (líneas verticales) (tomado de Zhengyi 1983). d) Disyunciones anfitropicales entre Norteamérica y Sudamérica. En rojo, la disyunción desértica del género *Ephedra*. En verde, la disyunción bipolar del musgo *Arctoa fulvella* (tomado de Wen & Ickert-Bond 2009).

En comparación con las disyunciones intercontinentales o transoceánicas, las distribuciones disyuntas entre áreas dentro de un mismo continente han sido relativamente menos estudiadas. Sin embargo, este tipo de disyunciones son especialmente interesantes en un contexto de cambio global, ya que, tradicionalmente se han explicado por vicarianza debida a cambios climáticos que habrían provocado la extinción de una biota en parte de su área de distribución, y han quedado reducidas a refugios en áreas climáticamente más estables o "islas continentales" (Raven & Axelrod 1978; Crisp & Cook 2007). Este proceso recientemente formalizado en biogeografía se ha denominado "vicarianza climática" (Wiens 2004). En estos casos, se podría testar la vicarianza observando si ha ocurrido una diversificación similar en múltiples linajes que ocupan áreas disyuntas; es decir, el origen de una barrera debería interrumpir simultáneamente el flujo génico entre múltiples especies. Crisp & Cook (2007) proporcionaron evidencia de múltiples divergencias sincrónicas en el sur de Australia para varios linajes a ambos lados de la llanura de Nullarbor, hace 14-13 Ma (Figura 2). Esto lo consideraron como una evidencia de que estas divergencias pueden atribuirse a un evento de vicarianza climática en vez de múltiples eventos independientes de dispersión desde un centro de origen.

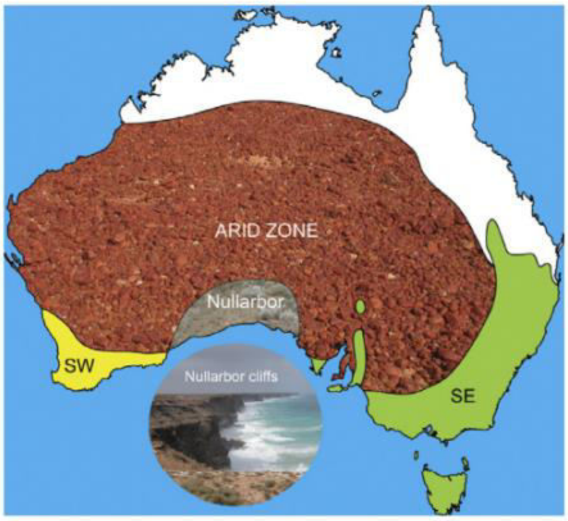


Figura. 2. Ejemplo de disyunción intracontinental. Se muestra la localización de los biomas templados al sur de Australia y las barreras que los separan. Los biomas templados del sureste y del suroeste están separados por el desierto árido de Nullarbor y la zona árida del centro de Australia (adaptado de Crisp & Cook 2007).

En esta tesis examinamos un presumible ejemplo de vicarianza climática. El patrón fitogeográfico conocido como la Rand Flora se muestra como un caso interesante ya que, presenta disyunciones similares entre los márgenes del continente africano en linajes de plantas no emparentados. La aridificación de África, con la aparición gradual de desiertos como el del Sáhara (Senut *et al*. 2009), el desierto cálido más grande del mundo, podrían ser un escenario ideal para testar disyunciones intracontinentales.

1.2 Origen evolutivo de una enigmática distribución geográfica: La "Rand Flora" Africana.

(Este apartado forma parte de un artículo aceptado en *Investigación y ciencia* (edición española de la revista americana *Scientific American)*: ver Apartado IX *List of papers*).

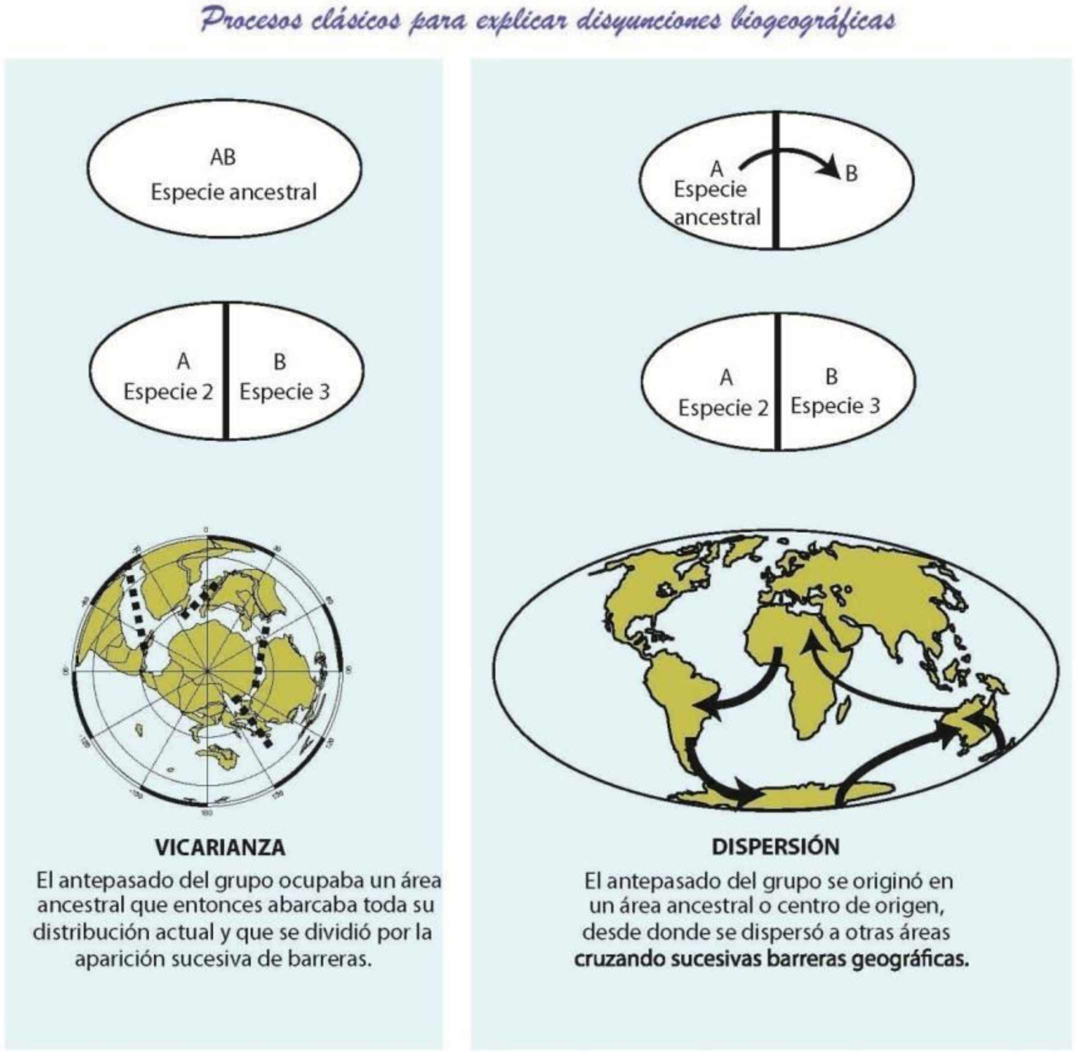
La Rand Flora es un patrón biogeográfico que relaciona especies de plantas distribuidas en los márgenes del continente africano. Se recogen aquí las diferentes hipótesis que explican su origen evolutivo, incluyendo la extinción debida a cambio climático.

Introducción

La presente crisis de la biodiversidad (se estima que alrededor de una décima parte de las especies del planeta desaparecerá en los próximos 100 años) y la acuciante necesidad de profundizar en las vías para su conservación han hecho crecer el interés por entender los mecanismos evolutivos que determinan los patrones de distribución geográfica de la diversidad biológica: ¿Por qué algunas regiones del planeta albergan mayor número de especies que otras? ¿Cuándo, dónde y cómo se formó la actual diversidad biológica de un continente?

Un tipo de distribución que ha llamado especialmente la atención de botánicos y zoólogos desde los tiempos de Charles Darwin (1859), Joseph D. Hooker (1860) y Alfred R. Wallace (1876), es el que se conoce como *disyunción geográfica*, en el que organismos emparentados habitan localidades distintas entre las que no existe continuidad geográfica.

Por ejemplo, la disyunción “gondwanica” de los mamíferos marsupiales, que habitan continentes del hemisferio sur actualmente aislados por cuencas oceánicas, como es el caso de los didélfidos en América y los uómbats y los canguros en Australia y en Nueva Guinea. Otro ejemplo es la llamada disyunción “*anfi*-Atlántica” presente en numerosas plantas leñosas del Hemisferio Norte, como los nogales o las magnolias, con especies en los Apalaches norteamericanos y cuyos parientes más próximos están en Japón o en China.



El antepasado del grupo ocupaba un área ancestral que entonces abarcaba toda su distribución actual y que se dividió por la aparición sucesiva de barreras.

El antepasado del grupo se originó en un área ancestral o centro de origen, desde donde se dispersó a otras áreas **cruzando sucesivas barreras geográficas.**

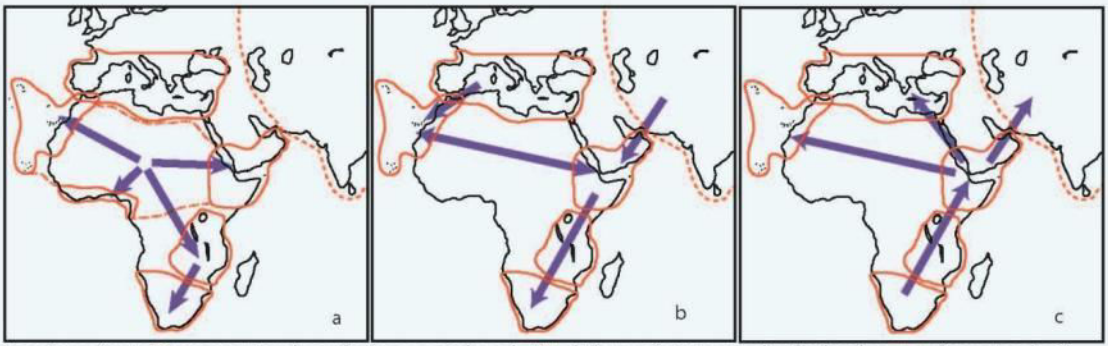
Figura 3. Ejemplos de linajes de plantas africanas con distribución Rand Flora. Algunos linajes presentan la disyunción biogeográfica entre el noroeste y noreste de África, mientras que otros presentan disyunciones aún más amplias con el este y sur de África

Se han propuesto dos procesos o mecanismos para explicar este tipo de distribución disyunta, dispersión y vicarianza (ver Figura 3). Durante mucho tiempo, la vicarianza fue la explicación favorita entre biogeógrafos, apoyada por el descubrimiento de la composición geológica de la Tierra y el subsecuente desarrollo de la teoría de la tectónica de placas. Sin embargo, el uso actual de secuencias de ADN para la reconstrucción de relaciones evolutivas y la posibilidad de datar las disyunciones geográficas utilizando el "reloj molecular" (modelo de evolución por el cual las mutaciones no selectivas se acumulan de forma constante en el ADN a lo largo del tiempo) han demostrado que muchas disyunciones son demasiado recientes para explicarse en términos de vicarianza y tectónica de placas y deben atribuirse en cambio a eventos de dispersión a larga distancia.

Frente al interés despertado por el origen de estas distribuciones transoceánicas, otro tipo de disyunción geográfica ha recibido menos atención en la literatura científica. Se trata de las llamadas *disyunciones geográficas continentales*, en las cuales especies emparentadas ocupan regiones geográficas disyuntas aisladas *dentro* de un mismo continente. En este caso, las barreras no son tanto geográficas, como en el caso de continentes separados por cuencas oceánicas, si no ambientales: las especies disyuntas están separadas por regiones intermedias de clima hostil en las que estas especies no podrían vivir y que sólo cruzan muy ocasionalmente. Un ejemplo son las regiones de montaña separadas por hábitats de llanura o los fragmentos de selva aislados por sabanas. Estas disyunciones, sin embargo, son importantes para entender el papel de otro proceso que junto con la dispersión y la vicarianza, resulta esencial para entender la distribución actual de la biodiversidad, como es la extinción asociada a cambio climático. Esto se debe a que, en muchos casos, estas disyunciones intracontinentales se explican como el resultado de cambios dramáticos en el clima, que habrían llevado a la extinción o extirpación de una flora o fauna ancestral en una parte de su área de distribución, dejando especies relictas en refugios o “islas continentales” de ambiente más favorable. Uno de los ejemplos más conocidos de este tipo de distribución es la llamada “Rand Flora” africana.

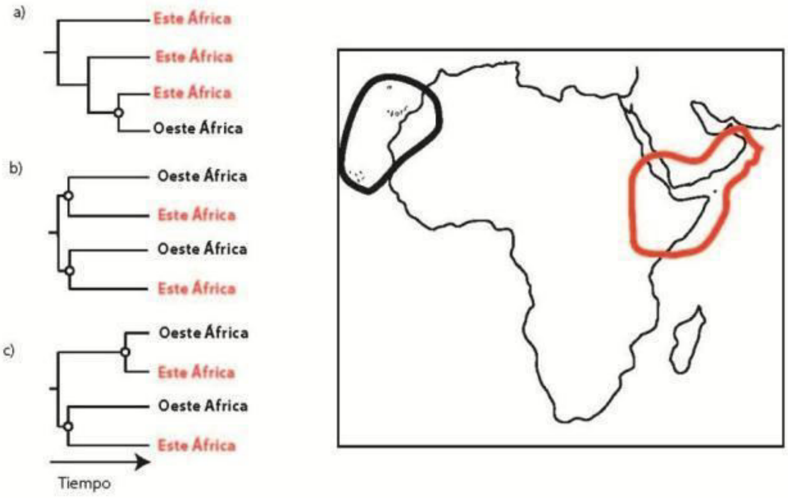
La Rand Flora

La existencia de especies de plantas emparentadas a ambos lados del continente africano, con distribuciones disyuntas entre Macaronesia, noroeste de África, el este de África (incluyendo el Cuerno de África), Sudáfrica y sur de Arabia (ver áreas rojas en Figura 4), ha intrigado a los botánicos desde hace décadas. Este patrón fitogeográfico se conoce como “Rand Flora” (del alemán para *margen* o *anillo*) o “Flora de Margen”, porque las especies se distribuyen en los márgenes del continente africano formando un anillo o reborde, mientras que el centro estaría *hueco*. Muchos de estos linajes de plantas presentan adaptaciones a climas subtropicales o a climas semi-áridos, por lo que la cuenca tropical del Congo en el centro de África o los grandes desiertos del norte como el Sáhara o el desierto arábigo representarían áreas de clima hostil más o menos infranqueables. Uno de los ejemplos más citados de este enigmático patrón florístico es el “drago” canario (*Dracaena draco*), que presumiblemente presenta sus parientes más cercanos en Arabia y el Noreste de África, pero se ha propuesto este tipo de disyunción en muchas otras familias de plantas (Figura 5).



Teorías y origen evolutivo de la Rand Flora

Las áreas (en rojo) representan los refugios asociados a la Rand Flora, mientras que las flechas hacen referencia a las 2 teorías que la explican: fragmentación (vicarianza) de una flora antes continua por aridificación de África dejando refugios en los márgenes continentales (a), o dispersión entre áreas geográficamente aisladas hacia el sur (b) o hacia el norte (c).



Patrones filogenéticos asociados a la "Rand Flora"

Se conoce como "**pseudocongruencia biogeográfica**" a la presencia del mismo patrón de distribución en dos o más grupos de **organismos sin que hayan compartido necesariamente la misma historia biogeográfica: el mismo tipo de disyunción podría haberse** formado en edades geológicas distintas (b y c).

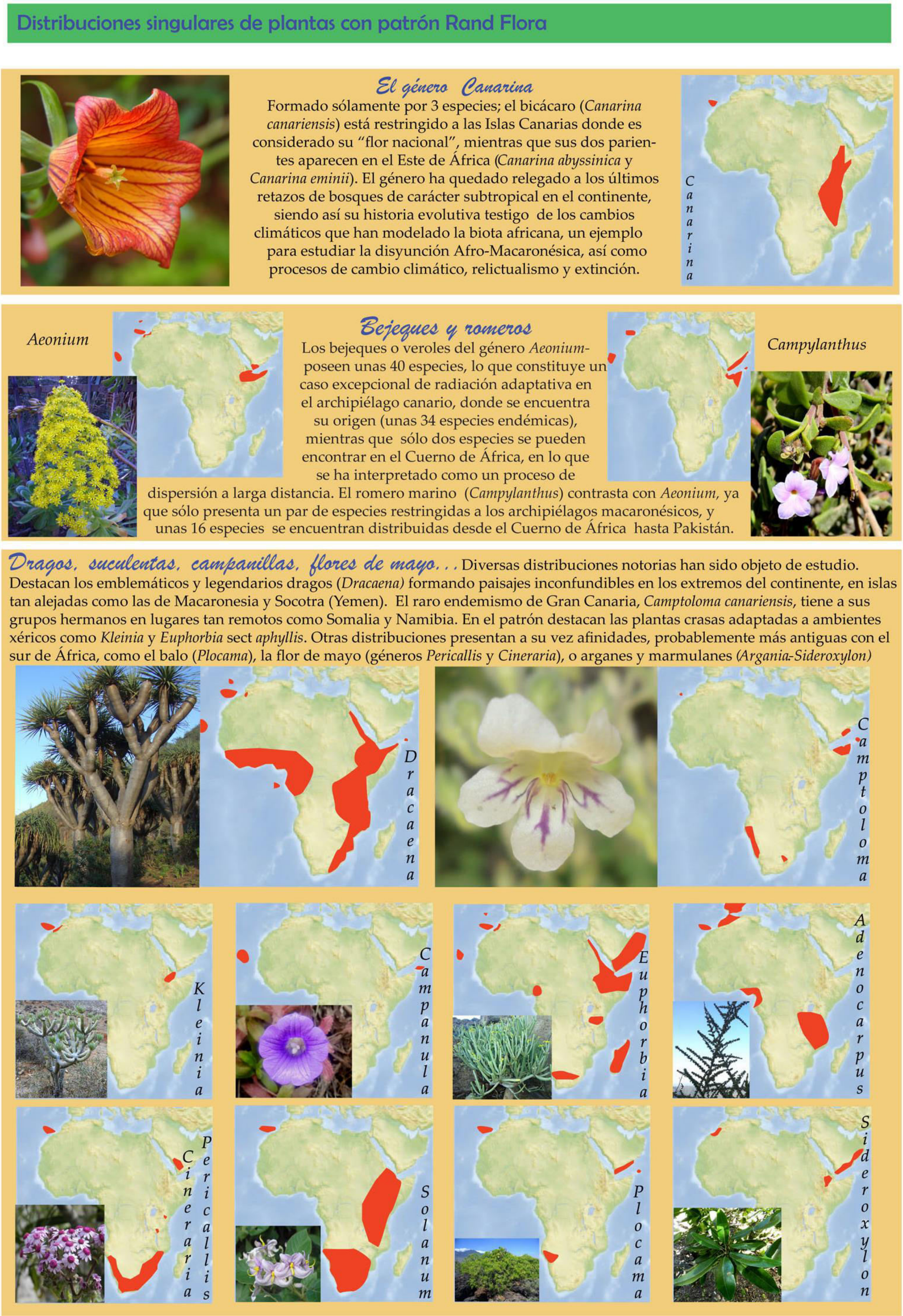
Figura 4. Superior: Hipótesis esgrimidas para explicar la existencia de especies de plantas emparentadas distribuidas en los márgenes continentales de África, en lo que se conoce como disyunción biogeográfica *Rand Flora*. **Inferior**: Patrones filogenéticos Rand Flora que se esperarían como resultado de los procesos de dispersión (a) o vicarianza (b-c). En este último caso, la dimensión temporal es importante porque debido al carácter cíclico de muchas barreras, un mismo patrón de disyunción biogeográfica (aquí entre el NO y NE de África) podría haberse formado en épocas geológicas distintas (*pseudocongruencia biogeográfica*).

Dos hipótesis alternativas se han propuesto para explicar este tipo de distribución geográfica:

1) Vicarianza y refugios climáticos (Figura 4 inferior a): las especies actuales son *relictos* de una antigua flora *macrocontinental* que se habría extendido durante el Cenozoico en África, Macaronesia y partes del Mediterráneo y que se habría extinguido en parte de su área de distribución como resultado de un proceso de *aridificación gradual* –que aún afecta al sur del Mediterráneo– dejando floras *relictas* en los márgenes continentales del este y oeste de África y Macaronesia. Esta hipótesis fue durante mucho tiempo la preferida por biogeógrafos. Algunos autores hablan de una flora semi-árida que ocupaba todo el norte de África y que habría dejado una serie de especies vicariantes en Macaronesia, Norte de África, y Arabia. Otros hablan de una flora subtropical más antigua –relacionada con la flora "Madre-Tethyense" que ocupaba las costas del antiguo Mar del Tetis y que habría sido diezmada por sucesivos eventos de aridificación, excepto en refugios climáticos en la región del Cabo , Macaronesia y áreas Afromontanas del este y oeste de África. Un ejemplo de esta teoría de vicarianza climática es la laurisilva macaronésica, que sería un remanente de una flora pre-Sahariana ancestral, que se extendió por el norte de África y Macaronesia hasta finales del Mioceno. El hecho de que las áreas donde se localizan los linajes Rand Flora presenten en general un alto grado de diversidad (especialmente en relación con las regiones limítrofes), se explicaría por su papel como *refugios* climáticos. En ellos, la estabilidad ambiental habría permitido a linajes antiguos sobrevivir y diversificarse, dispersándose a áreas adyacentes en periodos en los que las condiciones climáticas eran favorables.

2) Dispersión reciente entre regiones geográficamente aisladas (Figura 4 inferior b): según esta hipótesis, la distribución actual es el resultado de eventos de dispersión relativamente recientes entre regiones aisladas geográficamente. Numerosos estudios biogeográficos apoyan la idea de la dispersión a larga distancia como un proceso determinante en la formación de disyunciones intercontinentales. En el caso de la Rand

Flora, ciertas regiones pudieron haber actuado como *corredores* intermedios en un tipo de dispersión escalonada (*stepping stones*). Por ejemplo, se ha esgrimido que la cadena montañosa del Drakensberg en Sudáfrica y los arcos montañosos del este de África serían los corredores que habrían facilitado la dispersión hacia el norte de linajes endémicos de la flora Capense en Sudáfrica. Asimismo, sistemas montañosos aislados actualmente en medio del Sahara: por ejemplo, las montañas del Tibesti (Chad) o el macizo del Hoggar en Argelia, pudieron haber funcionado como islas continentales con hábitats favorables, permitiendo la dispersión entre el noreste y noroeste de África. La colonización de nuevas áreas geográficas sin competidores próximos habría favorecido una diversificación rápida, lo que explicaría la riqueza florística de muchas de estas regiones. Un alto porcentaje de estas migraciones habría utilizado principalmente dos rutas de dispersión distintas: una hacia el sur desde la región Mediterránea o desde Asia central, utilizando la colisión de la Placa Arábiga con Eurasia a mediados del Mioceno (16 millones de años) como puente para colonizar el este y sur de África, y otra hacia el norte, desde la región del Cabo y los montes Drakensberg al Sur de África, aprovechando el levantamiento de las montañas y altiplanos del Este de África hace 5 millones de años. Lo más probable es que, lo que se conoce actualmente como la Rand Flora, sea una combinación de procesos como dispersión, vicarianza y extinción, repetidos en el tiempo (Figura 4). Una hipótesis similar se ha argumentado para explicar la existencia de *pseudocongruencia biogeográfica* en disyunciones este-oeste de linajes tropicales africanos. Estos linajes habrían experimentado repetidos eventos de división (aridificación) y reconexión (períodos húmedos) entre el este y oeste de África como consecuencia de sucesivos cambios climáticos a partir del Mioceno.



Distribuciones singulares de plantas con patrón Rand Flora

El género Canarina

Formado sólamente por 3 especies; el bicácaro (*Canarina canariensis*) está restringido a las Islas Canarias donde es considerado su "flor nacional", mientras que sus dos parientes aparecen en el Este de África (*Canarina abyssinica* y *Canarina eminii*). El género ha quedado relegado a los últimos retazos de bosques de carácter subtropical en el continente, siendo así su historia evolutiva testigo de los cambios climáticos que han modelado la biota africana, un ejemplo para estudiar la disyunción Afro-Macaronésica, así como procesos de cambio climático, relictualismo y extinción.

Bejeques y romeros

Los bejeques o veroles del género *Aeonium*- poseen unas 40 especies, lo que constituye un caso excepcional de radiación adaptativa en el archipiélago canario, donde se encuentra su origen (unas 34 especies endémicas), mientras que sólo dos especies se pueden encontrar en el Cuerno de África, en lo que se ha interpretado como un proceso de dispersión a larga distancia. El romero marino (*Campylanthus*) contrasta con *Aeonium*, ya que sólo presenta un par de especies restringidas a los archipiélagos macaronésicos, y unas 16 especies se encuentran distribuidas desde el Cuerno de África hasta Pakistán.

Dragos, suculentas, campanillas, flores de mayo. . . Diversas distribuciones notorias han sido objeto de estudio. Destacan los emblemáticos y legendarios dragos (*Dracaena*) formando paisajes inconfundibles en los extremos del continente, en islas tan alejadas como las de Macaronesia y Socotra (Yemen). El raro endemismo de Gran Canaria, *Camptoloma canariensis*, tiene a sus grupos hermanos en lugares tan remotos como Somalia y Namibia. En el patrón destacan las plantas crasas adaptadas a ambientes xéricos como *Kleinia* y *Euphorbia* sect *aphyllis*. Otras distribuciones presentan a su vez afinidades, probablemente más antiguas con el sur de África, como el balo (*Plocama*), la flor de mayo (géneros *Pericallis* y *Cineraria*), o arganes y marmulanes (*Argania-Sideroxylon*)

Figura 5. Ejemplos de linajes de plantas africanas con distribución Rand Flora. Algunos linajes presentan la disyunción biogeográfica entre el noroeste y noreste de África, mientras que otros presentan disyunciones aún más amplias con presencia en el este o el sur de África.

Reconstruyendo el origen evolutivo de la Rand Flora

Hasta hace muy poco, los estudios biogeográficos sobre la Rand Flora se habían limitado a la comparación de patrones de distribución entre distintos linajes en busca de hipótesis que pudiesen aclarar su origen, pero sin una evaluación rigurosa de éstas. Durante la última década, tres grandes avances han hecho posible evaluar estas hipótesis por primera vez en un marco estadístico inferencial:

1) Muchas de las relaciones evolutivas esgrimidas para apoyar el patrón Rand Flora se basaban en similitudes morfológicas entre linajes. La generalización del uso de métodos moleculares en la reconstrucción filogenética demostró que, en ciertos casos, estas semejanzas eran el resultado de convergencia evolutiva (p.ej. adaptación a ambientes similares). Además, los datos moleculares han permitido por primera vez obtener estimaciones de la edad de divergencia entre los linajes disyuntos mediante el uso del "reloj molecular" –calibrado con el registro fósil o calibrado con estimas de la tasa de evolución molecular en linajes emparentados–, lo que ha permitido distinguir entre eventos antiguos de vicarianza geológica o climática o dispersión reciente.

2) Por otro lado, la acumulación de datos paleontológicos, paleogeográficos y paleoclimáticos permiten ahora la reconstrucción detallada del efecto de pasados eventos climáticos en la vegetación de África. Estas reconstrucciones indican que la aridificación del continente africano a partir del Mioceno (20 millones de años) no fue continua, sino que se produjo en *ventanas* discretas (*windows of opportunity*), coincidiendo con importantes eventos tectónicos y climáticos (ver Figura 6). Etapas de clima más húmedo, que habrían permitido a linajes de plantas extender su rango de distribución, se alternaron con etapas de aridez, en las que las condiciones climáticas hostiles habrían favorecido la vicarianza y diversificación en refugios.

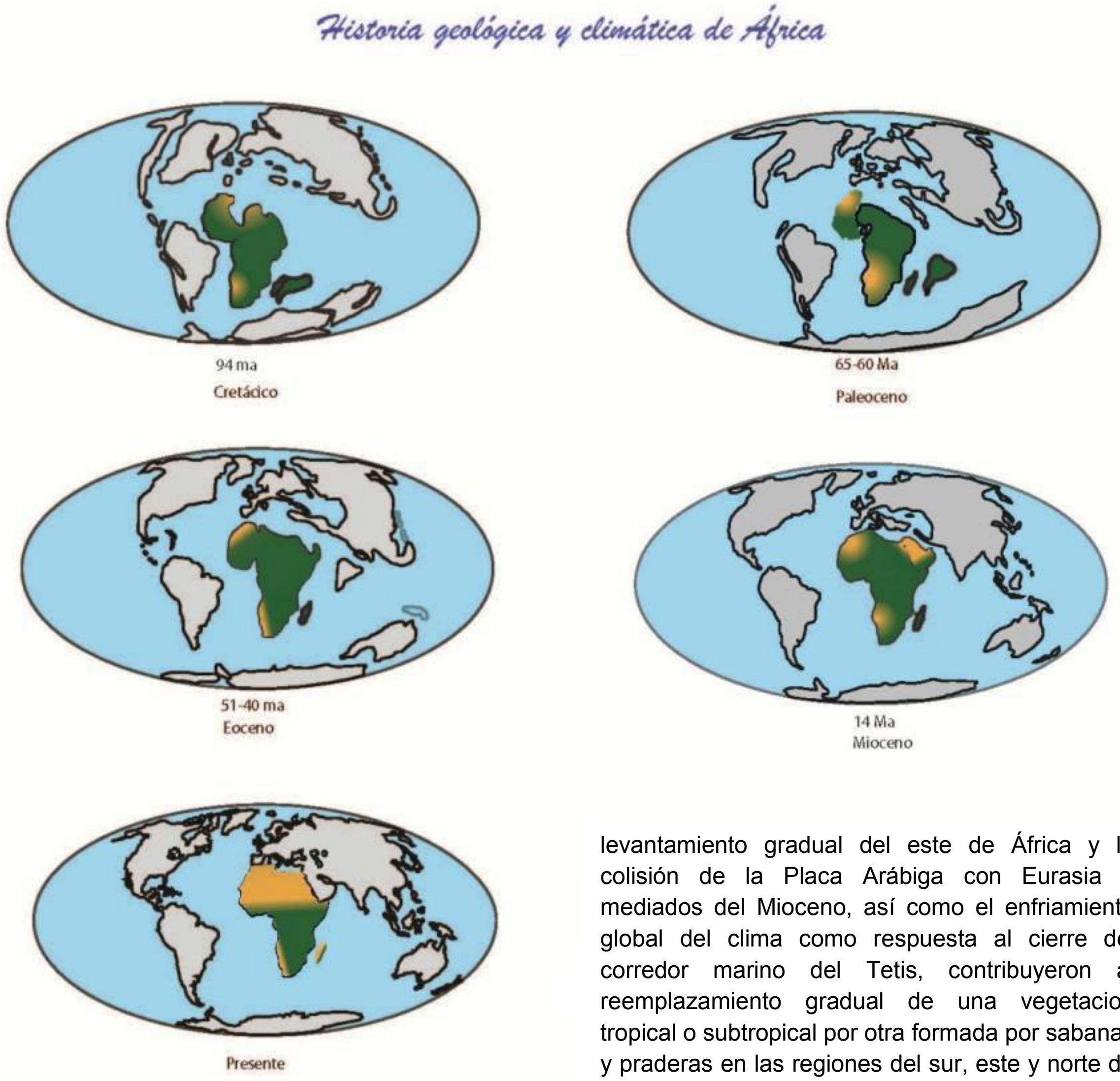


Figura 6: Breve historia geológica y climática reciente de África. Junto con Australia y la India, África es el continente que más dramáticamente ha cambiado su posición latitudinal desde el Mesozoico. A finales del Cretácico (95-80 millones de años), la placa africana que formaba parte del supercontinente meridional de Gondwana, se separó de éste y comenzó a moverse lentamente hacia el noreste hasta colisionar con la placa eurasiática a comienzos del Mioceno. La consecuencia más importante de este desplazamiento fue una tendencia global a la aridificación, desde un clima "gondwánico" templado y relativamente húmedo hacia un clima más arido a medida que se desplazaba hacia el cinturón ecuatorial. Sin embargo, no todas las regiones fueron afectadas por igual en este proceso. La aridificación avanzó desde el suroeste hacia el nordeste, de forma que se da la paradoja que la región central del Congo era a comienzos del Mioceno un desierto mientras que el norte de África mantenía condiciones semitropicales. El levantamiento gradual del este de África y la colisión de la Placa Arábiga con Eurasia a mediados del Mioceno, así como el enfriamiento global del clima como respuesta al cierre del corredor marino del Tetis, contribuyeron al reemplazamiento gradual de una vegetacion tropical o subtropical por otra formada por sabanas y praderas en las regiones del sur, este y norte de África. El proceso de aridificación se completó con la formación de los grandes desiertos del norte, Sahara y Árabe-Sinaítico a finales del Mioceno-Plioceno, contribuyendo aún más al aislamiento de elementos relictos de la antigua flora subtropical africana. Por el contrario, el patrón opuesto se observa en el centro de África, dónde el levantamiento gradual del este de África contribuyó a la subsidencia o hundimiento paulatino de la Cuenca del Congo y un cambio en el curso de los ríos, que contribuyó a la "tropicalización" de África central.

3) Finalmente, en los últimos años se han desarrollado nuevos métodos de análisis en reconstrucción biogeográfica basados en modelos matemáticos “paramétricos”, cuya mayor ventaja es que permiten la incorporación de información temporal en la reconstrucción biogeográfica. Esto es especialmente importante para el patrón Rand Flora en el norte de África, ya que la alternancia de periodos húmedos y áridos en el desierto del Sáhara probablemente ha dado lugar a una misma disyunción biogeográfica este-oeste en épocas geológicas distintas y que por tanto no pueden ser atribuidos al mismo evento de vicarianza climática (Figura 4 inferior). Uno de los primeros estudios en hacer uso de estos avances fueron los realizados por Isabel Sanmartín *et al.* en 2010. Estos autores utilizaron datos filogenéticos y biogeográficos de diferentes linajes de plantas, que presentaban la distribución biogeográfica Rand Flora, para examinar las distintas hipótesis del origen evolutivo de este patrón en un marco estadístico (Figura 7). Para ello, consideraron las diferentes regiones Rand Flora como “islas continentales”, separadas por barreras climáticas, y adaptaron un método biogeográfico desarrollado en islas oceánicas para estimar la tasa de dispersión o intercambio biótico entre estas regiones a lo largo del tiempo. Lo interesante del método es que permitía incorporar las diferencias entre los linajes Rand Flora en la estima de la tasa de dispersión histórica entre regiones –en cuanto a su edad de origen (el orden de magnitud varía entre millones y cientos de miles de años), tasa de evolución molecular (es sabido que el “reloj molecular” varía entre linajes e incluso entre especies de un mismo linaje) o su capacidad dispersiva (p.ej. algunas especies tienen adaptaciones a la dispersión por viento, otras no). Los resultados mostraron que la tasa de dispersión más alta se daba entre Macaronesia-noroeste de África con el este de África, mientras que la tasa más baja se encontraba entre el sur de África y el resto de las regiones (con la excepción del este de África, Figura 7).

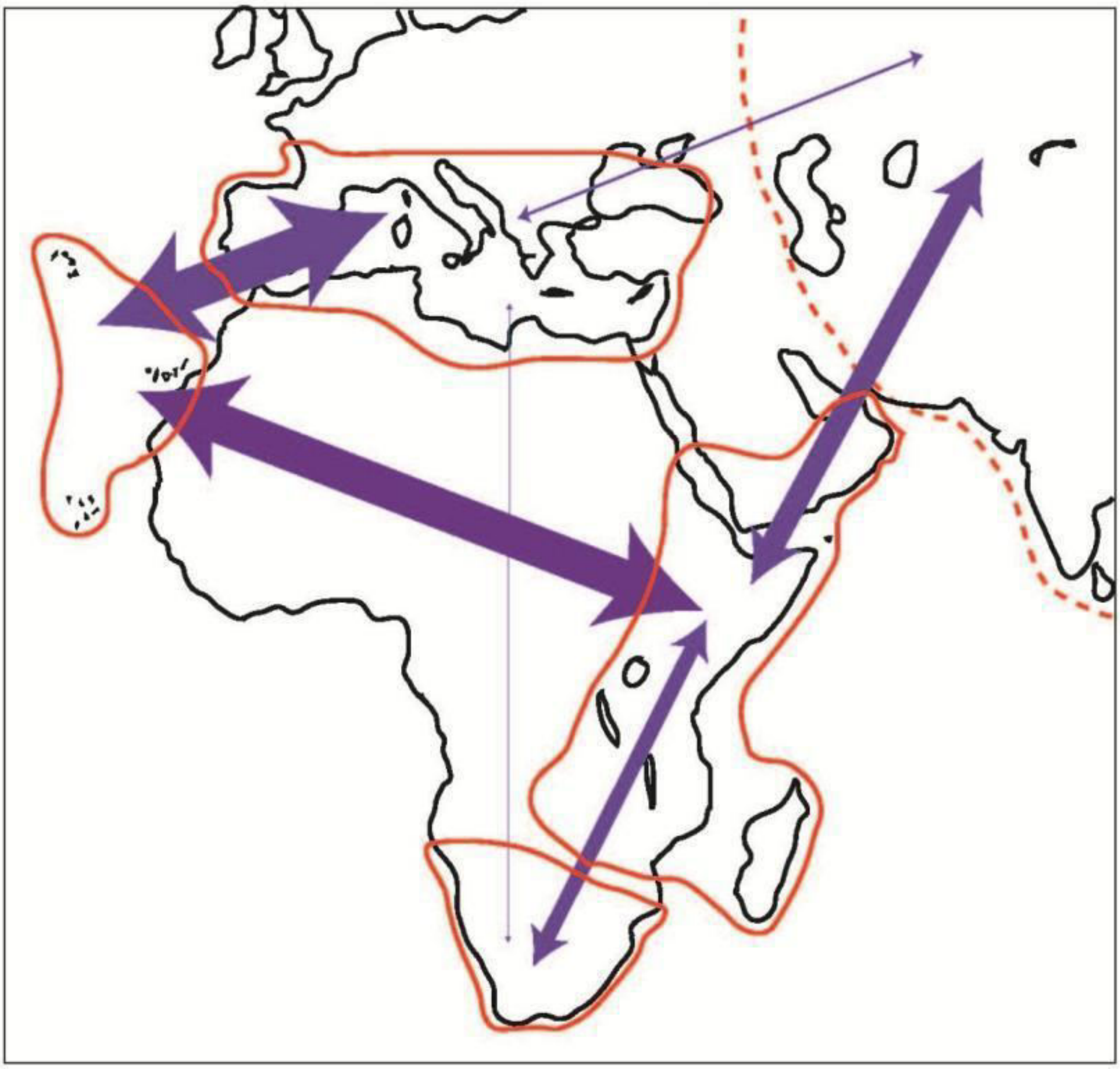


Figura 7: Origen evolutivo de la Rand Flora africana. Sanmartín y colaboradores (2010) llevaron a cabo un meta-análisis sobre datos filogenéticos y biogeográficos de 13 linajes de plantas con distribución Rand Flora. Consideraron cada una de las regiones Rand Flora (Macaronesia-noroeste de África, este de África-sur de Arabia, y sur de África) como "islas continentales" y estimaron la tasa de intercambio biótico o dispersión histórica entre regiones a lo largo del tiempo; incluyeron además dos regiones adyacentes, el Mediterráneo y Asia occidental. La tasa de migración más alta se encontró entre Macaronesia-noroeste de África y este de África, indicando una conexión biótica más reciente, mientras que la tasa más baja corresponde al sur de África, con poco intercambio biótico, excepto con el este de África. Estos resultados se corresponden bien con reconstrucciones históricas (Figura 6), que muestran cómo el proceso de aridificación avanzó desde el sur hacia el nordeste de África.

Estos resultados son interesantes puesto que concuerdan con reconstrucciones paleogeográficas (Figura 6) que señalan que el proceso de aridificación avanzó desde el suroeste en dirección noreste, de forma que la región del Norte de África mantuvo condiciones húmedas o subtropicales hasta hace muy poco, mientras que el Sur de África habría sido la primera en quedar aislada por barreras climáticas del resto de regiones Rand

Flora. Cuanto más antiguo es el intercambio biótico entre dos regiones, mayor es la probabilidad de que la extinción o eventos posteriores de dispersión hayan eliminado su rastro, lo que explicaría por qué el norte y sur de África presentaban una tasa de migración tan baja. Por otro lado, el hecho de que la tasa de intercambio biótico más alta con el sur de África se dé con el este de África concuerda con la hipótesis de que los arcos montañosos de esta región habrían sido utilizados como corredor por linajes sudafricanos para migrar al norte. En la actualidad estamos reevaluando este análisis para incluir nuevos linajes Rand Flora (Figura 5) y estimar el papel de la extinción en el origen evolutivo de este patrón biogeográfico.

Conclusión

Frente al interés despertado por regiones "hiperdiversas" como los Neotrópicos, África ha sido hasta hace muy poco el continente olvidado y existen pocos estudios sobre la evolución de su flora y fauna. El estudio del patrón Rand Flora supone una contribución en este sentido, ya que permitirá generar filogenias para numerosos grupos de plantas con su centro de distribución en África. Por otro lado, este patrón representa una oportunidad única para estudiar el papel de la extinción asociada a cambios climáticos en la evolución de la biodiversidad de un continente. En un escenario de cambio climático como el actual, entender cómo reaccionaron las especies a cambios climáticos en el pasado puede ayudarnos a predecir las consecuencias de éste en el presente o en futuros escenarios. En el caso de la Rand Flora, permitiría entender la influencia de la creciente aridificación del norte de África y sur de Europa sobre la flora endémica mediterránea.

2. La filogeografía para detectar la diversidad que "no se ve"

La filogeografía es la disciplina que estudia los patrones y procesos que determinan la distribución geográfica de los linajes genealógicos (Avise *et al*. 1987; Avise 1994, 2000). Se considera un puente entre la macroevolución y la microevolución, conectando así dos campos de estudio tradicionalmente separados: la sistemática y la genética de poblaciones (Avise *et al*. 1987). La filogeografía puede considerarse una subdisciplina de la biogeografía histórica (Avise 1994, 2000), ya que parte de la idea de que las especies poseen una estructura genética asociada con la geografía. Con esta estructura filogeográfica se pretenden inferir los patrones y procesos que determinan el origen y distribución de la biodiversidad entre poblaciones de una especie, o especies muy cercanas (Avise *et al*. 1987), y sacar así conclusiones respecto a la colonización, diversificación y extinción de linajes. Los patrones de distribución que son similares entre organismos distintos, por ejemplo, patrones convergentes entre refugios filogeográficos, podrían haber surgido en tiempos distintos y por diferentes procesos (Médail & Diadema 2009). Al comparar patrones filogeográficos entre varios linajes es posible encontrar procesos comunes y así plantear hipótesis generales sobre la distribución de la diversidad, por ejemplo, diferenciar entre hipótesis de dispersión y vicarianza. A su vez, identificando procesos demográficos y genealógicos es posible dilucidar la historia a la que se han visto sometidas las poblaciones, como: cambios climáticos, glaciaciones, vulcanismo, destrucción de islas volcánicas o efectos antropogénicos como la tala indiscriminada. Además, los métodos filogeográficos pueden dar evidencias de tamaños históricos poblacionales, patrones de migración histórica o de selección natural. Una de las ventajas de estos métodos es que permiten la búsqueda de evidencias de variación inter- e intrapoblacional, algo que es difícil de abordar con otros métodos de reconstrucción de patrones de diversidad, como por ejemplo Modelos de Distribución de

Especies (*Species Distribution Models* (SDM) o reconstrucciones fósiles. Sin embargo, una de las dificultades en filogeografía consiste en generar hipótesis demográficas alternativas. Para esto es útil acompañar estos estudios de otras evidencias (SDM, registro fósil) que permitan discriminar entre escenarios alternativos (Gavin *et al.* 2014).

En áreas que se postulan como refugios filogeográficos se espera encontrar mayor diversidad genética, mayor abundancia de alelos endémicos y ancestrales, además de una estructura genética espacial entre refugios y con respecto a aquellas áreas recientemente colonizadas (Hewitt 2000; Petit *et al*. 2003; Gavin *et al*. 2014). Probablemente, el ejemplo más estudiado es el de los refugios glaciares pleistocénicos, cuya huella puede observarse en muchos estudios filogeográficos de especies mediterráneas (Taberlet *et al*. 1998; Hewitt 2001; Petit *et al*. 2003; Gómez & Lunt 2007; Médail & Diadema 2009; Heuertz *et al*. 2010; Nieto-Feliner 2014). En estos refugios glaciares, los niveles de diversidad genética son generalmente mayores que en las áreas supuestamente colonizadas tras el retroceso glaciar. Sin embargo, la asunción de una diversidad genética reducida en áreas recolonizadas es demasiado simple. Se ha observado que la diversidad genética puede ser mayor o menor en las áreas recolonizadas, dependiendo del estimador de diversidad genética utilizado (Comps *et al*. 2001; Widmer & Lexer 2001). En sus estudios sobre el haya (*Fagus sylvatica*), Comps *et al.* (2001) encontraron una correlación negativa entre la riqueza alélica y la diversidad genética, donde áreas recolonizadas mostraban una reducida diversidad alélica pero mayor diversidad genética. En general, en loci neutrales, se espera una disminución mucho más acusada en la riqueza alélica que en la diversidad genética.

En comparación con los refugios pleistocénicos, se ha prestado poca atención al estudio de otras zonas en las que las glaciaciones han tenido un impacto menos severo. Este podría ser el caso de especies distribuidas en áreas donde ha influido especialmente la aridificación o desertización, algo especialmente interesante en el actual escenario de calentamiento global. Un ejemplo es el caso de los taxones Rand Flora estudiados en esta tesis. Los refugios a los que han quedado restringidos estas especies (sistemas insulares) probablemente han persistido a través del Pleistoceno y continuarán actuando como refugios en el futuro (Leroy & Arpe 2007). Algunos de ellos pueden haber jugado un papel importante en el posterior ensamblaje de los patrones de diversidad, actuando como "museos" (reservorios) y "cunas" de nueva diversidad génica hacia otras áreas recolonizadas, en analogía con los refugios glaciares en Europa (Hewitt, 2000). Tal es el caso de los sistemas insulares aquí estudiados, el este de África (Fjeldså & Lovett 1997; Linder 2014) y las Islas Canarias (Juan *et al*. 2000; García-Verdugo *et al.* 2015; Patiño *et al*. 2015).

2.1. Buscando refugios de biodiversidad

Los llamados "refugios filogeográficos" son áreas que dentro de una región habrían conservado el acervo genético y biológico de diferentes linajes ante cambios climáticos o geológicos, y por tanto, habrían actuado como “museos” para la conservación de la diversidad. Más tarde, algunos refugios podrían haber actuado también como “cunas” al producir nueva diversidad, y como “fuentes” cuando la diversidad migra hacia otras áreas adyacentes (Stebbins 1974). La identificación de tales refugios tiene gran importancia en la comprensión de la evolución y aparición de las distribuciones modernas de las especies (Hewitt 2004; Petit *et al*. 2005). Tradicionalmente se ha postulado que los centros de endemismo y las áreas ocupadas por disyunciones relictas son el legado de la permanencia de las especies en refugios pleistocénicos (Braun 1955; Whittaker 1961). Más recientemente la riqueza de especies y endemismos se ha explicado por estabilidad climática local (Fjeldså & Lovett 1997; Araújo *et al*. 2008; Médail &

Diadema 2009). Se piensa que el papel de los refugios climáticos ha sido crucial para salvaguardar la diversidad de los cambios climáticos del Pleistoceno. La localización precisa de estos refugios es crítica para la conservación óptima del potencial evolutivo (Petit *et al*. 2003, 2005), cuya preservación se hace muy importante tanto a nivel de especies como de poblaciones (Médail & Diadema 2009). Debido a esto, los refugios tienen un interés muy especial en políticas de conservación, ya que actúan como áreas clave para la supervivencia de la diversidad a cambios climáticos y para su persistencia a largo plazo. Una importante aportación la realizaron Myers *et al.* (2000), que definieron 25 áreas prioritarias para la conservación, denominadas *hotspots* de biodiversidad (Figura 8a) –áreas con una concentración excepcional de especies endémicas sometidas a una excepcional pérdida de hábitat. Sin embargo, esta aproximación estaba centrada a nivel de especies, quedando como asignatura pendiente el nivel intraespecífico. La variación genética intraespecífica es el nivel más fundamental de la biodiversidad; proporciona la base del cambio evolutivo y es crucial para el mantenimiento de las capacidades de las especies para adaptarse a las nuevas condiciones ambientales (Frankham *et al*. 2002; Yannic *et al*. 2013). Una importante aproximación en la búsqueda de refugios intraespecíficos la realizaron Médail & Diadema (2009) que, basándose en estudios filogeográficos, definieron 52 refugios en el Mediterráneo (Figura 8b) y acuñaron la definición de *hotspots filogeográficos* –reservorios de diversidad genética única a causa de procesos evolutivos. A pesar del conocimiento creciente sobre estos refugios en Europa, existe aún una laguna importante en otras áreas (p. ej. montañas del este de África).

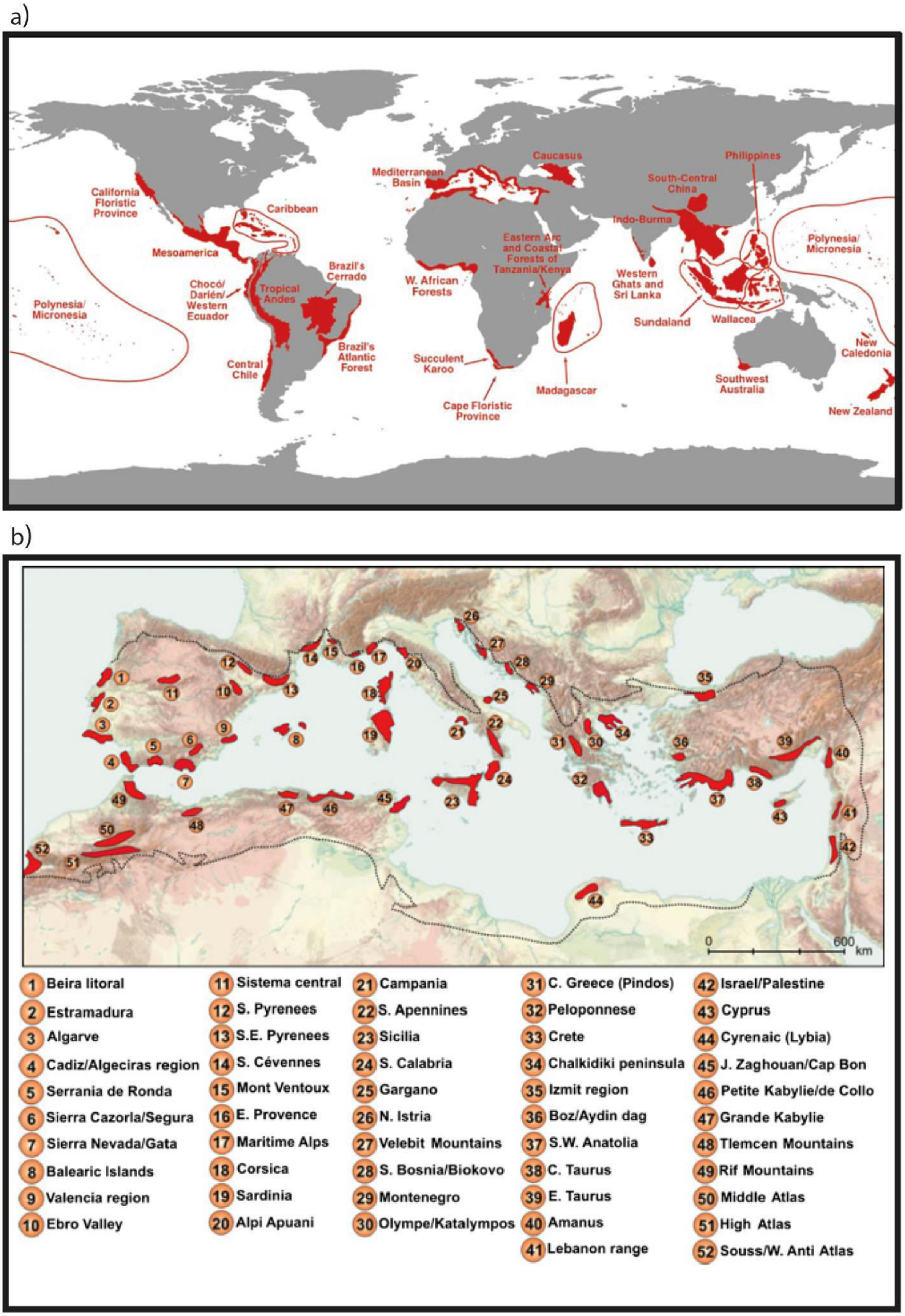


Figura 8. a) Los 25 *hotspots* biogeográficos prioritarios para la conservación de la biodiversidad según Myers y colaboradores (2000). b) Los 52 *hotspots* filogeográficos definidos para la cuenca mediterránea según Médail & Diadema (2009).

El estudio de los refugios del pasado puede ayudar a identificar los refugios actuales, así como los procesos y localizaciones geográficas donde se amortigua el impacto del cambio climático y se preserva la biodiversidad. En una revisión reciente, Gavin *et al.* (2014) resumen en tres las líneas de evidencia utilizadas para argumentar la existencia de refugios pasados: el registro fósil, los modelos de distribución de especies y la filogeografía. En esta tesis hemos utilizado estas líneas de evidencia para la búsqueda de refugios a escala regional en sistemas insulares (**Capítulos 1-4**) y a gran escala dentro de un continente (**Capítulos 5-6**). Las dificultades para obtener datos de fósiles en el norte de África ha limitado estas evidencias a su uso para calibraciones del reloj molecular (ver apartado V.3 en este resumen). Hemos hecho más énfasis en los modelos de distribución de especies, y más especialmente en las nuevas aproximaciones de la filogeografía estadística, para así identificar refugios en los sistemas insulares de Macaronesia y el este de África (ver apartado V.4 en este resumen).

3. Biología de islas

Los estudios de Charles Robert Darwin en las Galápagos (1859) y de Alfred Russel Wallace en el archipiélago malayo (1864), fueron pioneros en proponer la existencia de islas que nunca habían estado en contacto con el continente, característica que exigía su colonización por procesos de dispersión a larga distancia. Wallace, en su obra *Island Life* (1880), reparó en que el diferente origen de las islas debería afectar a la composición de sus biotas y realizó una primera clasificación de las islas que todavía sigue vigente, separando a las islas en oceánicas y continentales. Las islas oceánicas son islas de origen volcánico que se levantan directamente del lecho oceánico, por lo que nunca han formado parte de un continente y suelen estar separadas por grandes profundidades de agua. Estas islas poseen una ontogenia ya que nacen, crecen, se erosionan y finalmente se sumergen. Después de crecer, se produce su desmantelamiento a causa de diversos eventos catastróficos como: vulcanismo, colapsos de caldera y grandes deslizamientos de ladera. En cuanto a su biota se caracterizan por su *disarmonía* (Carlquist 1974) o disonancia –biotas generalmente más pobres con ausencia de grandes taxones (mamíferos terrestres, organismos dulciacuícolas) y menos especies por unidad de superficie. En estas islas, inicialmente desprovistas de vida, son las habilidades y mecanismos de dispersión los que determinan la probabilidad de llegada. Por otra parte Wallace separó las islas continentales en función de su conexión al continente, clasificándolas en recientes o antiguas. Las islas continentales recientes están próximas a un continente y unidas a éste por fragmentos de tierra que pueden emerger a causa de los niveles eustáticos del nivel del mar. Se caracterizan por su *armonía* –composición de la biota muy similar a la del continente cercano. Las islas continentales antiguas estuvieron en su momento unidas a un continente y, tras el fraccionamiento de la placa continental, fueron convertidas en fragmentos continentales libres. Su biota se caracteriza por una *armonía antigua*, siendo propia de la del continente del que se fragmentó pero con considerables ausencias taxonómicas.

La biogeografía de islas como ciencia alcanza su madurez hace ahora 50 años de manos de los ecólogos MacArthur & Wilson (1963, 1967), que la definen como una disciplina que trata de explicar los factores que afectan a la riqueza de especies de

Figura 9. Predicciones teóricas de la riqueza de especies en islas (modificado de MacArthur & Wilson 1967).

comunidades naturales. MacArthur & Wilson propusieron la conocida *Teoría del Equilibrio de la Biogeografía insular* (TEBI), que propone que la composición del número de especies presentes en una isla se debe a la existencia de un equilibrio entre una tasa de migración y una tasa de extinción. La tasa de inmigración es mayor en un principio y se irá reduciendo a media que las primeras especies ocupen los nichos disponibles. La tasa extinción es muy baja al principio e irá aumentando a medida que el número de especies colonizadoras aumente. Con el paso del tiempo, estos procesos se compensan llegando a un punto de equilibrio donde la tasa de inmigración es igual a la tasa de extinción (Figura 9a). Además, propone que la riqueza de especies de una isla es función de dos parámetros, tamaño y distancia al continente. Las islas de mayor tamaño contienen mayores áreas, por lo que hay oportunidades para una mayor variedad de hábitats y se reduce la probabilidad de extinción debido a fenómenos estocásticos; sin embargo, la tasa de extinción aumenta a medida que la isla disminuye (Figura 9b). Mientras que la inmigración está afectada por la distancia de una isla a la fuente de colonización. Tradicionalmente esta fuente es el continente, pero también pueden ser otras islas. En las islas que están más aisladas es menos probable que se produzcan inmigraciones que en las islas que están menos aisladas (Figura 9c).

MacArthur & Wilson (1963, 1967) postularon un equilibrio dinámico para explicar el proceso de especiación. En este equilibrio el número de especies se mantiene pero la composición de especies cambia (reemplazo o *turnover*). Décadas después Marc Lomolino (2000) introdujo el proceso de especiación y predijo que ésta aumentará cuando la inmigración y la extinción disminuyan, además de aumentar con la mayor área de la isla y con la mayor distancia al continente (Figura 10a).

Hasta el día de hoy, los biogeógrafos han seguido añadiendo complejidad a los modelos biogeográficos. Uno de los últimos modelos es el de Fernández-Palacios *et al.*

(2015), en el que han propuesto integrar el efecto de las glaciaciones. La principal diferencia con el TEBI es que, en lugar de representar solamente una 'instantánea' de un único conjunto de tasas de inmigración y extinción en islas inhabitadas y geográficamente estáticas, este modelo incorpora dos conjuntos de tasas resultantes de un período glacial y un período interglaciar, respectivamente, para la misma isla (Figura 10b).

Posteriormente a la TEBI, también surgieron los modelos para las islas continentales (Diamond 1972). Estos postulan que con las subidas del nivel del mar las islas continentales quedarían aisladas del continente, alojando un número de especies mucho mayor que el número de especies que podría sostener la nueva y reducida área. La isla quedaría sobresaturada de especies continentales en un periodo inicial, pero poco después irían extinguiéndose hasta alcanzar el punto de equilibrio. Cuanto más reducida sea el área, mayor será la tasa de extinción y viceversa. Esta relación ha sido ampliamente empleada a la hora de crear reservas, postulando que, cuanto mayor sean las áreas a preservar, mayor será la garantía de supervivencia de un mayor número de especies; lo que en biología de la conservación ha llevado al debate conocido como "sólo grande o varios pequeños" (SLOSS, por sus siglas *single large or several small*; Quamenn 1996).

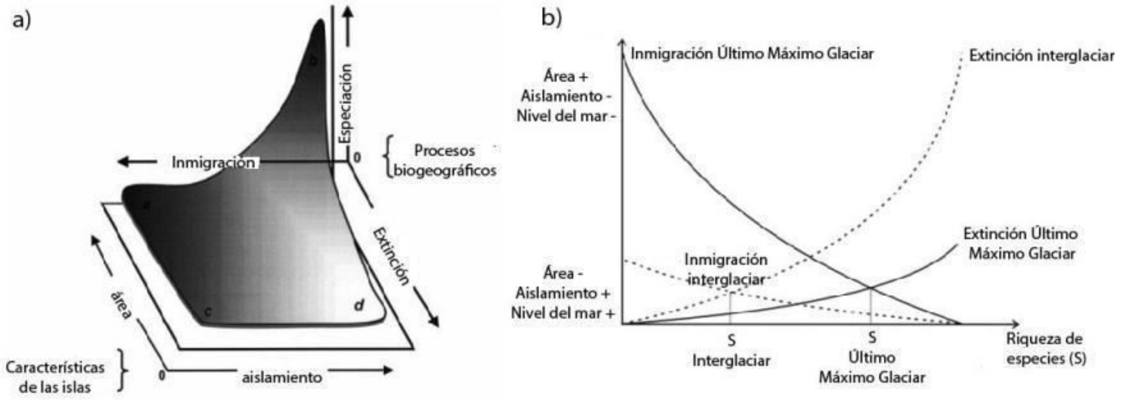


Figura 10. Algunos de los modelos que amplian la TEBI. a) Modelo de biogeografía de islas tripartito de Lomolino, ilustrando la variación en tres procesos fundamentales (inmigración, extinction y especiación) como función de las características de una isla. b) Modelo sensible a las glaciaciones propuesto por Fernández-Palacios *et al.* 2015.

A partir de las ideas del entomólogo Richard Levins (1969, 1970) se desarrolló el concepto de metapoblaciones –conjunto de poblaciones de una misma especie interconectadas por migraciones de individuos. Levins asumió un modelo de parches de hábitats discretos, conectados entre sí, de los mismos tamaños y proclives a la extinción. Se propone que, cuando la tasa de extinción es mayor que el tiempo de recolonización, una población puede desaparecer. De esta manera las metapoblaciones consideran una escala espacial y temporal que no se había abordado con la ecología poblacional y que ha tenido una gran acogida en la biología de la conservación. Considerando estas ideas metapoblacionales, Hanski & Simberloff (1997), en lugar de utilizar un continente como fuente de migración se refieren a parches de hábitat discontinuos que se asemejan a islas. De manera que los parches mayores actúan como fuentes, nutriendo a los parches menores o sumideros, que se extinguirían sin estos aportes.

Actualmente la teoría de islas se ha extendido a fragmentos terrestres continentales separados por áreas de clima hostil, como por ejemplo, montañas separadas por desiertos de arena, lagos o una matriz agrícola (Whitakker & Fernández-Palacios 2007).

La biota insular se caracteriza por diferentes peculiaridades. Su composición depende de diferentes factores propios de las islas como: edad geológica, superficie y topografía, presencia de islas puente (*stepping stones*) o puentes de tierra (*land bridges*), factores climatológicos como monzones y corriente marinas. Además, en los ambientes insulares se producen síndromes diversos como: amplitud de hábitat y de nicho, gigantismo y enanismo, pérdida de capacidad de dispersión, radiaciones adaptativas y desarrollo de leñosidad en linajes herbáceos (Whitakker & Fernández-Palacios 2007). Todo esto puede dar lugar a una biota singular que ha quedado refugiada en islas, las cuales albergan incluso comunidades enteras que se habrían extinguido en los continentes. Esto puede deberse a diferentes causas: i) que los hábitats dónde vivían estas especies se han destruido en otras zonas continentales y han quedado refugiadas en las islas, ii) que los ancestros continentales evolucionaron o iii) que los archipiélagos son de más difícil acceso para el hombre y los depredadores. A estas especies muchas veces se les considera relictas, reliquias vivas restringidas a sistemas insulares. Según Cronk (1997), los endemismos insulares pueden dividirse en paleoendemismos y neoendemismos. Los paleoendemismos son especies relictas cuyo aislamiento taxonómico es debido a la extinción en su área ancestral. Los neoendemismos son especies cuya forma ancestral se ha dispersado a una nueva área donde han sufrido una acentuada radiación. Los paleoendemismos se habrían originado en el continente, mientras que los neoendemismos se habrían originado *in situ*. En general las islas albergan una desproporcionada cantidad de especias relictas y endémicas y se hace urgente seguir desarrollando la teoría de islas, para hacer frente a las necesidades de su conservación.

En la presente tesis se hace especial énfasis en los procesos y patrones evolutivos tanto en islas oceánicas como en islas dentro de un continente.

3.1. Islas oceánicas e islas dentro de un continente: las Islas Canarias y las sky islands Afromontanas.

Los patrones y procesos evolutivos que caracterizan a islas oceánicas y *sky islands* o islas de altura, aunque no funcionan siempre en paralelo, son muy parecidos. En analogía con la TEBI, en las *sky islands* la distancia entre parches de hábitat correspondería a la distancia entre islas oceánicas, mientras que el tamaño del parche sería análogo al área de las islas oceánicas. En las *sky islands*, las llanuras (*lowlands*) actuarían de manera análoga a las masas de agua en los archipiélagos oceánicos, ejerciendo como barreras para la dispersión y el flujo génico (Figura 11). En ambos sistemas, las comunidades están estructuradas en cinturones de vegetación altitudinales, en los cuales pueden producirse drásticos desplazamientos debido al efecto de los ciclos glaciares. El descenso de tales cinturones podría conectar los ecosistemas a través de corredores y puentes de vegetación en las *sky islands (*Kebede *et al*. 2007) o mediante el descenso del nivel del mar en las islas oceánicas (Rijsdijk *et al*. 2014; Fernández-Palacios *et al*. 2015). En ambos sistemas la colonización puede ocurrir por migración; sin embargo, mientras que en islas oceánicas esto es obligatorio, en las *sky islands* la colonización también puede ocurrir desde las zonas vecinas colindantes, por ejemplo, mediante cambio de nicho. Por otro lado, ambos sistemas tienen en común su papel como generadores de diversidad (genética y específica) y por tanto son idóneos como escenario para realizar estudios evolutivos sobre diversos procesos y factores que generan la biodiversidad (McCormack *et al*. 2009; Warren *et al*. 2015).

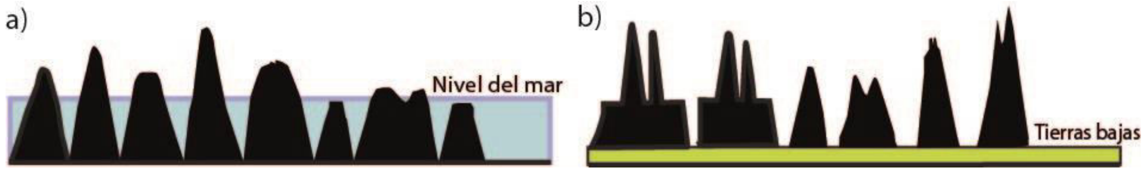


Figura 11. (a) Barreras clásicas en islas oceánicas y (b) en islas dentro de un continente.

En esta tesis nos centramos en dos sistemas insulares situados a ambos extremos del continente africano, el archipiélago Afromontano o *sky islands* en el este de África y las Islas Canarias en el oeste. En concreto, nos centramos en dos de sus hábitats: el bosque Afromontano en las *sky islands* y la laurisilva en las Islas Canarias. Estos dos sistemas insulares poseen numerosas afinidades que se detallan más abajo.

Debido a su posición latitudinal más baja, el efecto en África y Macaronesia de las glaciaciones pleistocénicas habría sido más moderado, sin llegar a producir extinciones en todo su rango de distribución y permitiendo el estudio de patrones y procesos adicionales a los clásicamente estudiados a latitudes más altas. Aunque

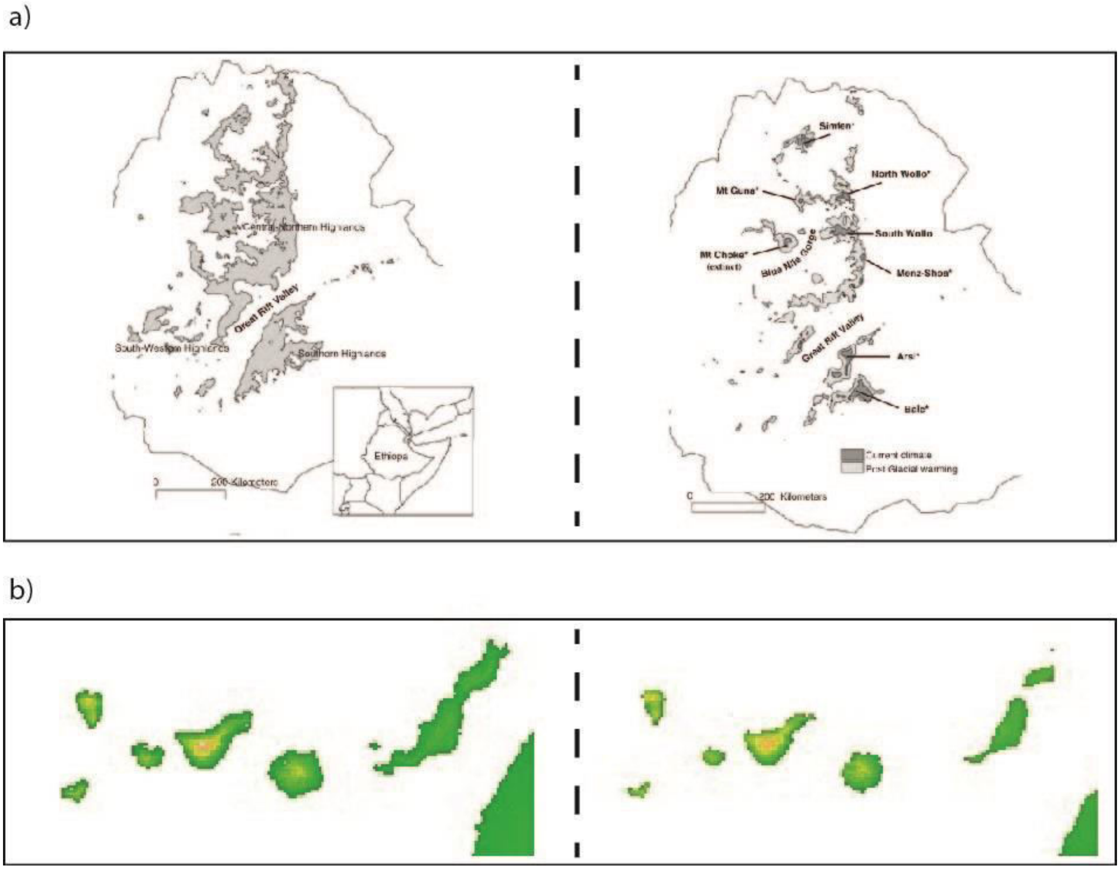


Figura 12. Efecto de las glaciaciones en el área de los sistemas insulares. a) Distribución estimada de los hábitats Afroalpinos etíopes durante el Último Máximo Glacial (izquierda), y bajo las condiciones actuales (derecha; gris oscuro) (tomado de Gottelli *et al*. 2004). b) Modelización de la superficie de las Islas Canarias en el Último Máximo Glacial (izquierda) y en la actualidad (derecha) (Mario Mairal, datos no publicados).

tradicionalmente las islas se han considerado como áreas climáticamente estables en el tiempo geológico, esta visión está cambiando con el estudio del efecto de las glaciaciones en los sistemas insulares (Gottelli *et al*. 2004; Harter *et al*. 2015; Fernández-Palacios *et al*. 2015). En el archipiélago Afromontano los cinturones de vegetación podrían haber descendido más de 1000 metros de altura y haberse formado puentes de vegetación entre *sky islands* (Coetzee 1964; Hedberg 1969; Maley 1996; Gottelli *et al*. 2004) (Figura 12a), mientras que en las Islas Canarias el descenso del nivel del mar habría provocado que se duplicara casi su superficie actual, con la consecuente emergencia de montes marinos y reconexiones entre islas (Rijsdijk *et al*. 2014; Fernández-Palacios *et al*. 2015) (Figura 12b). En definitiva, estos dos sistemas insulares proporcionan un laboratorio ideal para evaluar, en dos conjuntos de áreas aisladas, el efecto de las glaciaciones en los patrones de biodiversidad y la respuesta de los organismos a rápidas oscilaciones climáticas.

En ambos sistemas insulares los procesos geológicos han sido muy dinámicos. Ambos han surgido en los últimos 20Ma por actividad volcánica (Hughes *et al*. 1991; Wolfenden *et al*. 2005; Carracedo *et al.* 1998; Zaczek *et al.* 2015), dinamismo que puede proporcionar una referencia temporal para determinar la edad de estas islas. Estas áreas proporcionan un escenario interesante para comprender en qué medida algunos procesos geológicos, como la actividad volcánica, el levantamiento de montañas o la erosión, actúan sobre el origen, estructura y distribución de la biodiversidad. La correlación espacial y temporal entre procesos geológicos y patrones de diversidad sugiere la existencia de un nexo en el que los procesos geológicos actuarían como promotores de la especiación o la extinción. En este sentido, las dinámicas vulcano-tectónicas del Valle del Rift (**Capítulo 2**, Figura 1) y de las Islas Canarias (**Capítulo 3**; Figura 1) se presentan como un escenario ideal para testar estas hipótesis. Aunque es evidente que las glaciaciones pleistocénicas tuvieron su efecto a estas latitudes, la localización de estos hábitats subtropicales húmedos a una altitud media (Afromontano y laurisilva), junto a lo diverso de su topografía, proporcionan una alta diversidad local de nichos ecológicos y de microhábitats que podrían haber actuado como refugios a largo plazo y mitigado los efectos del clima extremo (Tzedakis 2009). Así, los organismos pueden migrar en altitud y las condiciones topográficas y ecológicas favorecen el aislamiento de las poblaciones, permitiendo una elevada divergencia incluso entre pequeñas distancias geográficas (Petit *et al*. 2003; Harter *et al*. 2015) (Figura 13). Los sistemas montañosos parecen haber jugado un papel fundamental en la formación de patrones de diversidad; y podrían haber actuado como refugio y como áreas de endemismo para diversas especies (Schönswetter *et al*. 2005). Sin embargo, aunque las montañas han sido destacadas como importantes *hotspots* de diversidad, en África permanecen menos exploradas que en otras partes de la Tierra, sobre todo en latitudes bajas. Ese desconocimiento a su vez se ve agravado por el incremento

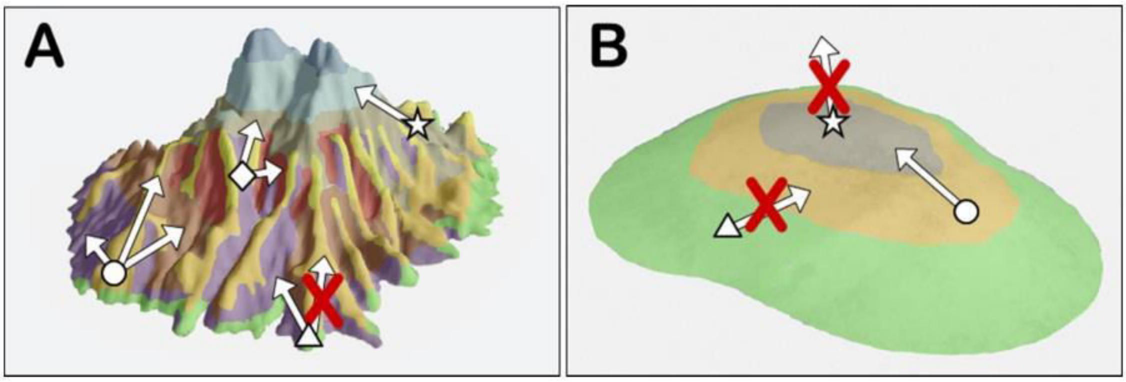


Figura 13. Efecto en las islas de la altura, topografía y diversidad de hábitats en los cambios de rangos potenciales para las plantas (tomado de Harter *et al*. 2015). Los tipos de hábitat están simbolizados por diferentes colores; las flechas indican migración hacia nuevos hábitats en respuesta a cambios climáticos. Las cruces rojas simbolizan la imposibilidad de una migración hacia otro hábitat debido a limitaciones ecológicas o espaciales (p. ej. especies competidoras predominantes, falta de especies facilitadoras, ausencia de sustrato, etc.). A) Las islas más elevadas y topográficamente estructuradas contendrán mayor diversidad de hábitats, por lo que las especies de plantas tendrán diferentes opciones para encontrar micro-refugios adecuados (ver movimientos hacia diferentes hábitats). B) En islas planas y menos estructuradas, las especies se encontrarán limitadas para encontrar micro-refugios accesibles y adecuados. La dispersión limitada, los filtros establecidos en una dirección, o la ausencia de alturas más elevadas podrían causar problemas para la supervivencia de las especies si no logran adaptarse a las nuevas condiciones climáticas en sus hábitats originales.

de la actividad humana en el fértil archipiélago Afromontano del este de África (sobreexplotación de recursos naturales, tala indiscriminada, cultivos, caza excesiva y sobrepastoreo; Hawando 1997) y por las predicciones futuras de cambio climático (Seto *et al*. 2013). En definitiva, aunque los sistemas insulares han estado sometidos a grandes cambios climáticos y procesos geológicos de tipo catastrófico (vulcanismo) en escala de millones de años, es a corto plazo el efecto de las actividades humanas sobre estos ecosistemas el que puede ser devastador para su diversidad y su potencial evolutivo.

Esta tesis se enfoca en diversos patrones y procesos que afectan a la biodiversidad en diversos sistemas insulares: las *sky islands* del bosque Afromontano (**Capítulos 1**, **3**, **4**), las islas oceánicas del archipiélago canario (**Capítulos 1**, **2**, **4**) y las "islas climáticas" para los linajes Rand Flora (**Capítulos 5** y **6**). El estudio de especies distribuidas en sistemas insulares en África y regiones adyacentes es importante para entender cuáles son los procesos responsables de la formación de patrones de diversidad genética y biótica en estas latitudes. Para ello, hemos elegido el género *Canarina*, cuya distribución abarca tanto el bosque Afromontano como las Islas Canarias.

3.2. El archipiélago de islas Afromontanas

Las comunidades Afromontanas consisten en una serie de altiplanicies (mesetas y volcanes) aisladas a modo de archipiélago entre zonas de llanura, que constituyen un centro de endemicidad en las montañas del este y oeste del continente (White 1983). Las montañas del este de África y sur de Arabia están reconocidas como uno de los puntos calientes para la conservación global de la biodiversidad (Mittermeier *et al*. 2004) habiéndose denominado: *The Eastern Afromontane Biodiversity Hotspot* (Figura 14a). En esta área, de las 10856 especies encontradas (7600 especies de plantas) al menos un tercio son endémicas (30.8% de endemicidad: Mittermeier *et al*. (2004)). Aquí habitan especies tan emblemáticas como el gorila de montaña, el babuino gelada, el lobo etíope, el café y las violetas africanas del género *Saintpaulia* (Figura 14b).

Figura 14. a) *The Eastern Afromontane Biodiversity Hotspot* (Mittermeier *et al.* (2004); b) especies emblemáticas de las montañas esteafricanas (de izquierda a derecha y de arriba a abajo: *Saintpaulia*, gorila, grulla carunculada, lobo etíope, café y gelada); c) *Sky island* de Ol Doinyo Lengai (Tanzania) circundada por la sabana. Fuente fotos: Wikimedia commons.

Topografía e historia geológica de los archipiélagos Afromontanos

En África, los archipiélagos Afromontanos se encuentran distribuidos principalmente en la zona ecuatorial y el este del continente africano (Figura 15a). Geológicamente, la estructura más destacable en el este de África es el Gran Valle del Rift –una fractura de 4830 km en dirección norte-sur (Figura 15b). En su parte norte el Valle del Rift está circundado por el macizo etíope, que se extiende por: Etiopía, Somalia y Eritrea. Esta región está formada por altas mesetas basálticas con una altitud media de 1500 metros, salpicada por más de 20 cumbres volcánicas que superan los 2500 metros de altitud y algunas que llegan a los 4000 metros. A estas extensas altiplanicies se les ha denominado "el techo de África". El macizo etíope, a su vez, está dividido en dos grandes unidades: el macizo de Abisinia al noroeste y el macizo del Harar al sudeste. Ambos macizos están separados por el Gran cañón del Valle del Rift, donde se han formado numerosos lagos en el piso de la fosa tectónica. En África Central el Valle del Rift se divide en dos grandes arcos volcánicos: el Rift Oriental (Gregory Rift) y el Rift Occidental (Albertine Rift), que se vuelven a unir al sur de Tanzania. Estos arcos están salpicados por numerosos accidentes geográficos: antiguos macizos (Rwenzori, Usambara, Kipengere, Udzungwa); elevadas mesetas (Myika, Rukiga) y varios volcanes de más de 2500 metros de origen más reciente (Meru, Kilimanjaro, Kenia, Rungwe, Usambara, Uluguru, Virunga). Tanto el macizo etíope como los arcos volcánicos se encuentran bien delimitados geográficamente. El macizo etíope está rodeado por: desiertos (en Eritrea, Sudán, Arabia, Somalia y Kenia), el Mar Rojo y Arábigo al Este y el Nilo Azul y las tierras bajas al oeste. Los arcos volcánicos están delimitados por: los numerosos lagos del Valle del Rift, la Cuenca del Congo al oeste y el océano Índico al este (Figura 15b). Así, este alto grado de aislamiento tiene un buen soporte geográfico, y se hace una analogía del archipiélago Afromontano con "islas del cielo" o *sky islands* (Hedberg 1969, 1970). Nos encontramos ante un mosaico de islas dispersas dentro de un continente; grandes mesetas y montañas aisladas de elevada altitud, donde las comunidades de plantas y animales destacan por su singularidad y endemicidad.

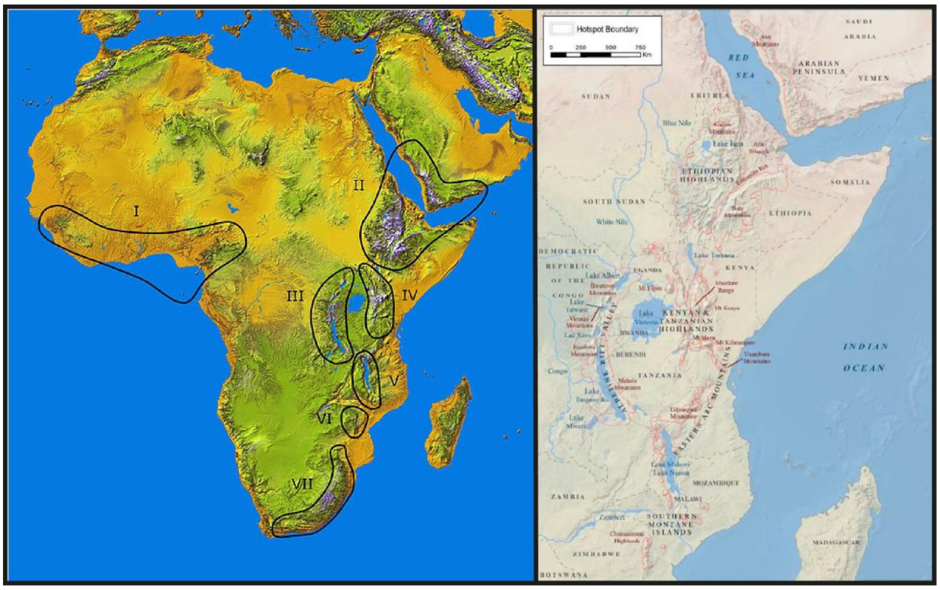


Figura 15. a) Montañas Afroalpinas; b) Mapa del este de África. Fuente imágenes: Wikimedia commons.

Estas montañas tienen orígenes diferentes desde el Mioceno hasta el Pleistoceno Tardío (Tabla 1) y han permanecido aisladas desde su origen (Hedberg 1961a). Las primeras elevaciones se produjeron en el macizo etíope; posteriormente se elevó el arco volcánico del este (Gregory Rift), mientras que el arco oeste (Albertine Rift) es el último que aparece, llegando al máximo levantamiento durante el Plio-Pleistoceno (Chorowicz 2005). El levantamiento del este de África fue una fuerza determinante en el establecimiento del clima africano en el Neógeno. Ocasionó una reorganización de la circulación atmosférica y una fuerte aridificación. Se redujeron los bosques tropicales y aumentaron los hábitats de sabana. La comparación de faunas desde el Mioceno Temprano al Mioceno Medio sugiere un gran cambio climático a finales del Mioceno; las comunidades del Mioceno Medio con dietas basadas en especies forestales (con metabolismo C3 dominante) fueron reemplazadas por nuevas familias de mamíferos ungulados consumidores de herbáceas (metabolismo C4) (Sepulchre *et al*. 2006). De esta manera, el relieve ha tenido una fuerte influencia en el origen de la diversidad que encontramos en el este de África.

Tabla 1. Principales eventos geológicos que originaron el Gran Valle del Rift y las montañas del este de África

	Millones de años (Ma)	**Eventos geológicos**	**Bibliografía**
ESTE DE ÁFRICA			
	30 Ma	Fracturas en la depresión del Afar y en la meseta etíope, debido a un punto caliente de actividad responsable de la emisión de flujos de lava	Hoffman *et al*. 1997; Mege & Korme 2004
	29 Ma	Se abre el Rift en el sistema del Mar Rojo - Golfo de Adén.	Wolfenden *et al*. 2005; Hughes *et al*. 1991
	25 Ma	Vulcanismo asociado al Rift. Aparecen los primeros lagos	Roberts *et al*. 2012
	27.5 – 23 Ma	El Rift se intensifica en la parte sur del Mar Rojo	Hughes *et al*. 1991
	20 Ma	Primeros signos de vulcanismo en el centro de Kenia, en la futura unión de los tres Rift.	Pickford, 1982; Fitch *et al*. 1985
	16 Ma	Colisión de la placa Arábiga con Eurasia	Sanmartín 2003; Oberprieler 2005; ver Allen & Amstrong 2008 para una datación más antigua
	13.8 Ma	Levantamiento de los bordes del Mar Rojo	Goudie 2005
	11 Ma	Se comienza a abrir el Rift en Etiopía	Wolfenden *et al*. 2004
	11 Ma	Ausencia de puentes de tierra posteriores al Mioceno en el Mar Rojo	Fernandes *et al*. 2006
	12 – 10 Ma	Se comienza a abrir la rama oeste del Rift, aparecen la cuenca y el Rift del Tanganica.	Lezzar *et al*. 1996; Cohen *et al*. 1993
	11 Ma	Actividad volcánica en el Rift de Kenia	Ebinger 1989a
	12 – 7 Ma	Vulcanismo en el Rift occidental; el vulcanismo comienza hace 12 Ma en el norte y hace 7 Ma en el sur	Ebinger 1989a

	8 – 6 Ma	Periodo de actividad tectónica que produjo el levantamiento masivo de la topografía del este de África y de las tierras altas de Camerún	Sepulchre *et al*. 2006; Cerling *et al*. 1997
	7 – 6 Ma	Primeros indicios de aridificación del desierto del Sáhara	Senut *et al*. 2009
	5 – 2 Ma	Levantamiento topográfico de los rifts de Tanganica y Malawi	Ebinger *et al*. 1993
Gregory Rift (rama este)			
	23 – 11 Ma	Primeros signos de vulcanismo y depresión topográfica	Hackman *et al*. 1990; Smith and Mosley 1993
	2 – 1 Ma	Episodios magmáticos	Abebe 2007
	Desde 650 Ka	Episodios magmáticos	Abebe 2007
Albertine Rift (rama oeste)			
	12 Ma	Comienza el vulcanismo en el norte de la rama, en el Rwenzori	Kampunzu *et al*. 1998
	11 – 9 Ma	Comienza el vulcanismo en los montes Virunga. Después de un periodo de quiescencia (9–3 Ma) aparecen los volcanes en el Plio-Pleistoceno.	Kampunzu *et al*. 1998
	8 Ma	Comienza el vulcanismo al sur de la región del Kivu. Se produce una gran fractura al norte del Lago Kivu	Pickford *et al*. 1993
	8.6 Ma	La actividad volcánica comienza al sur del arco volcánico, en el volcán Rungwe	Ebinger *et al*. 1989b
	12 – 9 Ma/ 8 – 7 Ma	La cuenca central comienza a formarse entre 12- 9 Ma; la cuenca norte entre 8 y 7 Ma	Cohen *et al*. 1993
	6 – 5 Ma	Se producen mayores fracturas al norte del Lago Malawi (Nyasa), y el Rift aumenta en longitud hacia el sur	Flannery & Rosendahl 1990

La biodiversidad asociada a las *sky islands*.

En las *sky islands* la altitud de las montañas determina el rango de los hábitats, de forma que podemos encontrar tres cinturones de vegetación altitudinales bien definidos (Figura 16): i) bosque Afromontano, ii) de ericáceas (subalpino) y iii) Afroalpino (Hedberg 1951). En el cinturón de bosque húmedo Afromontano abundan las formaciones de: *Hagenia abyssinica*- *Hypericum* sp., *Podocarpus falcatus* y *Juniperus procera*, bosques de bambú (*Sinarundinaria* sp.), y otros géneros como *Prunus* sp., *Olea* sp., *Ocotea* sp., *Impatiens* sp. y *Cyathea* sp. El cinturón subalpino está compuesto por varios géneros de ericáceas (*Blaeria*, *Erica*, *Vaccinium*; White 1983) y otras especies como *Hypericum revolutum* y *Adenocarpus manii*. El cinturón de mayor altura o Afroalpino está dominado por pastizales, matorrales y páramos de altura. Aquí dominan las plantas gigantes con roseta (*Lobelia* sp. y *Dendrosenecio* sp.) y los géneros *Alchemilla* y *Helichrysum*. En las zonas bajas, las barreras que separan las *sky islands* están formadas por extensas sabanas secas de transición, semidesiertos y selvas de baja altitud (*kolla*) dominadas por los géneros *Acacia*, *Boswellia*, *Commiphora* y *Terminalia*. Actualmente, muchas de las zonas bajas han sido reemplazadas por una matriz agrícola donde abundan cultivos de tef (*Eragrostis tef*), cebada (*Hordeum vulgare*) y café (*Coffea arabica*), así como plantaciones de *Cupressus lusitánica* y *Eucalyptus globulus*.

La flora que habita en las *sky islands* está compuesta por elementos de origen mixto: un componente importante está relacionado con la flora de las tierras bajas tropicales; sin embargo, también existen afinidades notables con la región de El Cabo, el Mediterráneo, y el Hemisferio Norte (Engler 1892; Hedberg 1961a). Según el cinturón de vegetación ocupado, la flora que habita las *sky islands* se ha dividido en: Afroalpina y Afromontana, mostrando ambas diferentes patrones y afinidades biogeográficas, mientras que el cinturón de ericáceas se corresponde con una zona de transición que presenta ambos tipos de floras. La flora Afromontana es mayormente de carácter forestal, presentando afinidades con las tierras bajas tropicales. Además tiene un importante componente de especies con amplia distribución a lo largo de los archipiélagos Afromontanos, las Montañas Drakensberg y la región del Cabo –patrón acuñado por White (1978) como *track Afromontano*. Adicionalmente, estas áreas comparten algunas afinidades biológicas con los sistemas montañosos de Angola y Camerún (p. ej. gorilas, suimangas, turacos y lobelias gigantes), razón por las que se les denominó como "Islas Galápagos de África" (Kingdom 1990). La flora Afroalpina ha sido considerada tradicionalmente como un elemento relicto con afinidades con la flora templada Euroasiática (*African pan-temperate element*: Engler 1904; Hedberg 1961a). Este cinturón posee una alta endemicidad de plantas vasculares (80% según Hedberg (1969); 77% según Gehrke & Linder (2014)), con una buena representación de elementos holárticos (hasta un 13%). De esta manera, los parientes más cercanos de estas plantas no se encuentran en las áreas bajas circundantes, sino que se encuentran en lugares más lejanos como las áreas templadas de ambos hemisferios, el Mediterráneo o el Himalaya (cf. Hedberg 1965; Linder 2014).

En las *sky islands* Afromontanas es destacable la presencia de diferentes especies emparentadas habitando en diferentes macizos. Por ejemplo, los artrópodos de la hojarasca tienen un alto grado de endemicidad en cada una de las montañas del Rift Oriental (Scharff, 1993). En el cinturón Afroalpino son emblemáticas las radiaciones de lobelias gigantes (*Lobelia* sp.) y senecios arborescentes (*Dendrosenecio* sp.), radiaciones que se extienden a las numerosas moscas ápteras que viven refugiadas en estas plantas.

Origen de la biodiversidad en las *sky islands*: patrones y procesos evolutivos

El levantamiento del este de África, el Gran Valle del Rift y el vulcanismo asociado, son elementos primordiales para entender los patrones de diversidad en las montañas del este de África. Se asume que el surgimiento de estas montañas sería el pistoletazo de salida para su colonización y diversificación. Por otro lado, también habría que subrayar el efecto del clima asociado a las montañas (ver revisión del clima de África en el **Capítulo 5** de esta tesis). De esta manera, se asume que el origen y la edad de diversificación de los grupos más antiguos debería ser igual o menor que la edad de emergencia de las montañas. Para explicar los patrones de diversidad en las *sky islands* se ha debatido entre la fragmentación de comunidades ampliamente extendidas o migraciones entre montañas. Lo que a priori parece destacar es la notoria brecha geográfica que supone el Gran Valle del Rift, el cual ha sido determinante para explicar los patrones de diversidad en numerosos grupos independientes: mamíferos (babuino gelada, Belay & Mori 2006; ñu, Arctander *et al.* 1999), aves (avestruz, Freitag & Robinson 1993; género *Sheppardia*, Roy *et al.* 2001; nectarínidos, Bowie *et al.* 2004), insectos (termitas, Wilfert *et al*. 2006; mosquito *Anopheles gambiae*, Lehmann *et al*. 2000) y plantas (ver tabla 2 del **Capítulo 2** de la presente tesis).

La hipótesis montane forest bridge

Actualmente, los bosques del este de África se encuentran altamente fragmentados, principalmente debido a la expansión de la agricultura. Sin embargo una extensa cobertura forestal se extendió durante los periodos glaciares-interglaciares hasta el Holoceno (Bonnefille *et al*. 1990; Kuper & Kropelin, 2006). La hipótesis *montane forest bridge* postula que, cuando los bosques de montaña estaban más extendidos, la colonización entre montañas adyacentes fue más probable, tanto durante los previos periodos interglaciares como en el presente periodo interglaciar, antes del impacto de la agricultura (Hedberg, 1969; Kebede *et al*. 2007; Figura 16a).

Figura 16. a) Cinturones de vegetación y su movimiento según la hipótesis *mountain forest bridge* (Kebede *et al*. 2007); b) cinturón de vegetación Afroalpino; c) cinturón de ericáceas (delante) y Afromontano (detrás). Fuente fotos: Wikimedia commons.

Hedberg (1969) ya sugirió que sin el impacto humano, los bosques montanos de Elgon, Charangani, Aberdare y Monte Kenia, podrían haber estado en contacto directo en tiempos históricos (Holoceno). Posteriormente numerosos estudios han sugerido que el cinturón Afroalpino y el de ericáceas descendieron entre 1000-1500 metros cubriendo áreas mucho más extensas que las actuales (Coetzee 1964; Flenley 1979; Maley 1996; Arctander 1999; Gottelli *et al*. 2004; Kebede *et al*. 2007), con la unión de parches forestales aislados a través de puentes de vegetación entre *sky islands*. Por todo esto, un bosque montano histórico parece más que probable en las mesetas etíopes, a ambos lados del valle del Rift. Sin embargo, parece más improbable que las zonas bajas del norte de Kenia –que separan las mesetas etíopes de los arcos volcánicos– hayan estado cubiertas por extensas superficies forestales. Aunque sí que se han encontrado algunas afinidades en la vegetación a lo largo de la parte oeste del valle del Rift, como entre la meseta de Abisinia y las montañas de Imatong en Sudán del Sur (Gehrke & Linder 2014). Incluso en el Valle del Omo (al norte de la cuenca del Turkana) hay evidencias de mayores superficies forestales (Bobe *et al*. 2006). Todo esto indica que en algunas áreas bajas, los valles fluviales sirvieron como refugios y centros de endemicidad para algunas especies forestales hasta el Pleistoceno Inferior. Estos valles a su vez pudieron actuar como corredores norte-sur durante las oscilaciones del Plio-Pleistoceno (Kingdom 2003; Bobe & Behrensmeyer 2004). Por lo tanto, los bosques montanos quizá sí que descendieron lo suficiente como para unirse en algunas zonas. Sin embargo, el contacto en los cinturones de ericáceas y Afroalpinos parece poco probable (Hedberg 1970). Y aunque pudieran haberse dado reconexiones a través de las altiplanicies que unen las montañas, se cree que el Valle del Rift siguió actuando como barrera este /oeste para la distribución de la biodiversidad.

Se ha propuesto un largo corredor de vegetación para explicar las afinidades bióticas entre el archipiélago Afromontano y las montañas de Camerún (Moreau 1966). Un ejemplo es la semejanza genética que existe entre estas áreas en las poblaciones del árbol *Prunus africana*, lo que indica la presencia de un posible corredor a través de África Ecuatorial (Kadu *et al.* 2011). Parece ser que una hipótesis similar se puede postular para *Hypericum revolutum* en el Pleistoceno-Holoceno, con poblaciones actuales en el este de África y en el oeste de Camerún (Cendoya & Sanmartín, datos sin publicar)

En general, las *sky islands* habrían estado sometidas a varios periodos de expansión y contracción de rango, especialmente durante las glaciaciones. Especies y poblaciones quedaron fragmentadas y aisladas durante las contracciones en diferentes macizos montañosos. Esta vicarianza provocaría diferenciación genética y especiación (Taberlet *et al.* 1998), y ocasionaría las numerosas radiaciones adaptativas que ahora observamos.

Dispersiones entre montañas

Alternativamente, otros estudios postulan el predominio de las dispersiones a larga distancia para explicar las afinidades entre los parches de vegetación montana del este de África. Hedberg (1970, 1986) ya sugirió que los eventos de migración entre montañas dependían de dispersiones a larga distancia a causa de vientos monzónicos y pájaros. Estas migraciones implicarían una mayor estocasticidad en la distribución de los taxones, por lo que no esperaríamos encontrar patrones genéticos congruentes, sino patrones aleatorios con poca estructura geográfica.

La presencia del elemento holártico en la flora Afroalpina se ha explicado por dispersiones recientes desde el norte en el Plio-Pleistoceno (Koch *et al.* 2006; Assefa *et al.* 2007). Se cree que una ruta migratoria existió en el norte del continente a través de los sistemas montañosos de la Península Arábiga. Las montañas Arábigas presentan una gran similitud florística con las montañas Simien en el noroeste de Etiopía (Gillett 1955; Puff & Nemomissa 2001) y constituirían el punto de partida para las radiaciones de la flora holártica en África. Así se ha explicado la distribución actual de numerosos elementos holárticos como: *Arabis alpina* (Koch *et al.* 2006), *Lychnis* sp. (Popp *et al.* 2008), *Carduus schimperi* y *Trifolium cryptopodium* (Wondimu *et al.* 2014). Una ruta similar se ha propuesto para los escasos taxones con origen asiático, tal es el caso de *Hypericum* (sección *Campylosporus*, Meseguer *et al*. 2013) y *Canarina* (**Capítulo 1** de esta tesis).

Las colonizaciones por dispersión desde el Hemisferio Norte, junto a la radiación *in situ*, han tenido un importante papel en la composición de las floras montanas de África (Gehrke & Linder 2009). Estas colonizaciones no son fenómenos excepcionales sino que habrían ocurrido de manera recurrente. Por ejemplo, se han detectado 2 colonizaciones independientes para *Arabis alpina* (Koch *et al.* 2006), 4-6 para el género *Ranunculus* y 13 para el género *Carex* (Gehrke y Linder 2009). Sin embargo, otros grupos habrían colonizado las montañas mediante una única dispersión y posterior radiación adaptativa (p. ej. *Dendrosenecio*, Pelser *et al*. 2007; *Lychnis*, Popp *et al*. 2008). Sin duda, la región holártica ha sido la más importante fuente de linajes para las montañas del este y sur de África (Gehrke & Linder 2009). La presencia del cinturón de ericáceas a ambos lados del Valle del Rift también se ha explicado por dispersión (Gizaw *et al*. 2013). Además, en el cinturón Afromontano se han detectado algunos fenómenos raros de dispersión en la especie forestal *Hagenia abyssinica* (Ayele *et al*. 2009). Incluso se ha detectado flujo génico a ambos lados del Rift en grupos sin adaptaciones evidentes para la dispersión a larga distancia (Wondimu *et al*. 2014).

La dispersión también ha podido producirse de manera gradual, mediante migraciones entre parches de hábitat a modo de *stepping-stones*. Este modelo predice patrones en los que la diferenciación genética depende de la distancia geográfica. Adicionalmente, el descenso altitudinal de los cinturones de vegetación habría aumentado el área disponible y la producción de diásporas, y se habrían reducido las distancias entre parches de vegetación. De esta manera, se ha detectado que una gran cantidad de grupos Afromontanos se habrían originado en la región del Cabo, siguiendo una migración unidireccional hacia las montañas Drakensberg, y desde allí, a los archipiélagos

Afromontanos del este de África (Galley *et al.* 2007), donde se habrían producido especiaciones *in situ* y radiaciones adaptativas. Estas migraciones desde El Cabo han ocurrido en los últimos 17 millones de años, coincidiendo con el surgimiento de las regiones montañosas del este de África (Galley *et al.* 2007). Para explicar las relaciones entre los Montes de Camerún y los archipiélagos Afromontanos, se ha especulado con migraciones utilizando las montañas Marra (Sudán) a modo de *stepping stones* (cf. Wickens 1976). Para las relaciones del este de África con el Mediterráneo y la Macaronesia se han propuesto sucesivas expansiones y contracciones de rango a lo largo del norte de África (Bramwell 1972; Sanmartín *et al.* 2010; Désamoré *et al.* 2011); en este caso, el este de África actuaría como área fuente y los macizos montañosos en medio del Sáhara (p. ej. Hoggar y Tibesti) podrían actuar a modo de *stepping stones* (*Erica arborea,* Désamoré *et al.* 2011).

Conservación

Los bosques montanos tropicales son uno de los ecosistemas más ricos en especies de la Tierra (Bussmann 2001). El estado de los archipiélagos Afromontanos es crítico, sobre todo en Etiopía. Estos ecosistemas han sido muy reducidos y fragmentados, su protección es escasa y actualmente están representados solamente en algunos Parques Nacionales (Simien, Bale Mts., Awash, Omo y Nechisar). Un informe de la FAO (2001) ya alertaba sobre la alarmante destrucción de estos hábitats de montaña en el último siglo, los cuales siguen sufriendo un rápido proceso de deforestación y degradación del suelo (Kloos & Legesse 2010).

A esto hay que sumar lo poco que se conoce sobre su biodiversidad. Algunos estudios han revelado que existe una fuerte diferenciación en la diversidad genética de especies entre montañas muy cercanas. Incluso en grupos bien conocidos como las aves, nuevos estudios genéticos revelan la existencia de especies crípticas y han encontrado evidencias para separar 6 especies montanas en 18 (Kahindo *et al.* 2007). El conocimiento de los límites de cada especie y la existencia de especies crípticas son fundamentales a la hora de determinar planes de conservación.

La endemicidad, *relictualismo*, exclusividad filogenética y alta biodiversidad, son sólo algunas de las razones por las que salvar estas comunidades de montaña debería ser prioritario para la comunidad conservacionista.

Generalidades de la biodiversidad Afromontana

La diversidad en África se ha interpretado como el resultado de la suma de diferentes floras. La flora tropical de las tierras bajas albergaría las especies más antiguas desde finales del Cretácico, las floras áridas y Austro-templadas se añadirían en el Paleoceno, mientras que las floras Afromontanas y Afroalpinas serían de origen más reciente, coincidiendo con el levantamiento y vulcanismo que comenzó en el Mioceno (Linder 2014).

En general, los bosques Afromontanos han actuado como áreas sumidero. Se cree que gran parte de la biodiversidad se ha generado *in situ* en el propio continente, a partir de la flora de las tierras bajas tropicales. La existencia de estrechas bandas de vegetación horizontales en las montañas habría facilitado el cambio a nichos de diferente clima con posteriores oportunidades para la adaptación. Adicionalmente, existen dos componentes que llegaron a través de rutas recurrentes de migración: el componente Afromontano que migró desde la región de El Cabo y el componente holártico que migró (en su mayoría) desde el Hemisferio Norte, aunque también desde Asia y Norteamérica. Las migraciones desde El Cabo se habrían producido desde el Mioceno, mientras que las migraciones desde el holártico son más recientes. En consecuencia, los clados más antiguos de la flora Afromontana datan del Mioceno Temprano, mientras que los clados Afroalpinos son los más jóvenes de África y datan del Mioceno Superior/Plioceno (Linder 2014). Tras la colonización, las especies habrían quedado aisladas en diferentes rangos montañosos y se habrían producido los característicos fenómenos de radiaciones adaptativas en islas. A este escenario habría que añadir dos elementos que se han relevado de gran importancia para la distribución de la biodiversidad en las montañas africanas: el escaso flujo génico a ambos lados del Gran Valle del Rift y las contracciones-expansiones de rango de los cinturones de vegetación.

Sin embargo, queda mucho trabajo por hacer con la flora Afromontana. Por ejemplo, las dataciones son escasas y podrían existir elementos con edades anteriores al Mioceno, las relaciones entre las selvas Afromontanas y las de Camerún apenas se han estudiado, tampoco el papel de las selvas Afromontanas como áreas refugio y fuente; por otro lado, las *sky islands* de tamaño y edad similar, geográficamente cercanas, tienden a tener fuertes similitudes en la composición de especies (Gehrke & Linder 2014), pero esta similitud interespecífica puede ser engañosa a nivel intraespecífico, ya que la diferenciación genética entre montañas puede ser muy alta (Ehrich *et al*. 2007) y una gran parte de la biodiversidad parece residir en especies crípticas que todavía no se han estudiado (Kahindo *et al*. 2007).

3.3. Las Islas Canarias

Las Islas Canarias comprenden siete islas volcánicas (junto a algunos islotes y roques) situadas solamente a 96 km de la costa noroeste de África (**Capítulo 3**, Figura 1). Estas islas forman la región biogeográfica de la Macaronesia junto a otros cuatro archipiélagos de origen volcánico: Azores, Madeira, Salvajes y Cabo Verde, además de un enclave continental africano conocido como *Argan sector* (Engler 1879; Takhtajan 1969; Sunding 1979; para otra definición de Macaronesia ver Vanderpoorten *et al*. (2007)). Las Islas

Canarias poseen un elevado porcentaje de diversidad exclusiva y relativamente amenazada, lo que ha llevado a incluirlas dentro del *hotspot* mediterráneo (Myers *et al.* 2000). En Canarias se pueden encontrar más de 12.500 especies terrestres de las cuales el 30.2% son endémicas, destacando el número de endemismos en grupos como: gasterópodos (247 especies; Ibáñez 1994), arañas (300 especies; Wunderlich 1991) y coleópteros (1160 especies; Carrillo & Oromí 2000). Muchos de estos grupos han sufrido una enorme diversificación, especialmente espectacular en invertebrados (los curculiónidos del género *Laparocerus* poseen 102 especies endémicas y los caracoles del género *Hemicycla* 76 especies) y en plantas (*Aeonium* (39 especies), *Argyranthemum* (22), *Cheirolophus* (17), *Echium* (16), *Limonium* (16), *Lotus* (17), *Pericallis* (13), *Sideritis* (24), *Sonchus* (15), etc.).

La flora vascular canaria se caracteriza por un elevado número de especies endémicas (alrededor de 600; un 40% de endemicidad según Santos-Guerra (1999)). Alrededor del 70% de estos endemismos poseen un carácter leñoso, en ocasiones derivado de la adaptación de especies continentales herbáceas a medios insulares (Carlquist 1970), tal es el caso de las alianzas de *Bencomia* y *Dendrosunchus*, así como de los géneros *Argyranthemum* y *Echium*.

Breve historia geológica de las Islas Canarias

Las Islas Canarias se encuentran sobre la placa Africana, encima de un punto caliente volcánico de, al menos, 68 Ma. Aunque su origen todavía es motivo de controversia, las últimas investigaciones apuestan por un origen a causa de una pluma mantélica (Zaczek *et al.* 2015). Las islas volcánicas siguen un ciclo en el que nacen, crecen, se desmantelan, se erosionan y se vuelven a sumergir. Todos estos estados se pueden observar en la provincia volcánica de Canarias, la cual está formada por un archipiélago de islas emergidas y otras sumergidas. Estas islas se han ido desplazando de oeste a este debido a la expansión de la dorsal Mesoatlántica, por lo que las islas situadas al este son más antiguas (muy erosionadas o ya sumergidas), mientras que las islas van disminuyendo en edad en sentido oeste. Existen al menos cinco grandes islas ya sumergidas, conocidas como las Paleo-Canarias, cuyo origen se remonta hasta hace unos 60 Ma en el monte submarino Lars (Fernández-Palacios *et al*. 2011). En las islas emergidas la edad es menor y abarca desde 21 Ma en Fuerteventura al este, hasta 1.1 Ma en el Hierro al oeste (Tabla 2; Figura 1 en **Capítulo 3**). Conocer la edad de las islas es primordial para comprender su estado evolutivo, ya que las islas más antiguas son las que han sufrido mayores procesos erosivos y destructivos, característicos de islas oceánicas. En un principio, las islas atraviesan una fase constructiva en la que crece el edificio volcánico, para posteriormente entrar en una fase destructiva en la que las islas se desmantelan debido a fenómenos erosivos y catastróficos (p. ej. hundimientos y deslizamientos gigantes o megadeslizamientos), que reducen la superficie insular. Un ejemplo ilustrativo es la isla de Fuerteventura, que alcanzó una altura de 3300 metros en el Mioceno y actualmente ha quedado reducida a 810 metros (Stillman 1999). En Canarias los deslizamientos gigantes o grandes avalanchas de derrubios han tenido un papel muy importante en el modelado de la topografía. En los últimos dos millones de años han ocurrido hasta 12 deslizamientos gigantes de más de 20 km (Figura 17; Tabla 2). Uno de los más impresionantes es el deslizamiento de El Julán en El Hierro, cuyos restos se localizan hasta la llanura abisal de Madeira, 600 Km al oeste (Figura 17b).

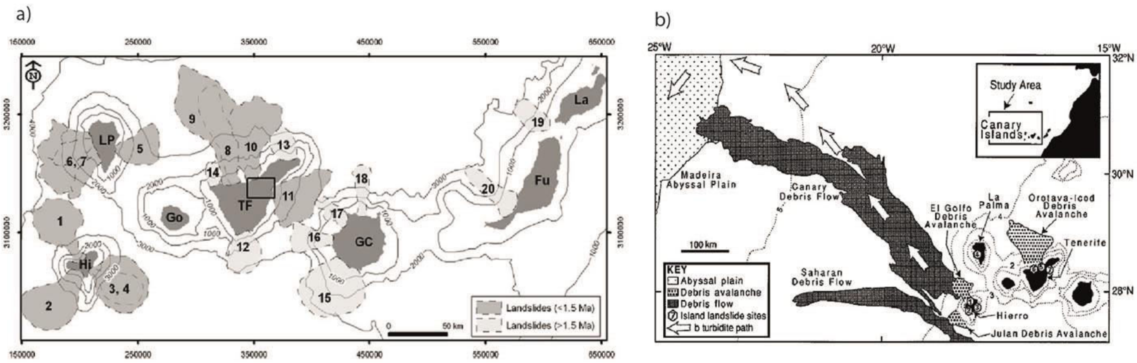


Figura 17. a) Principales deslizamientos gigantes de las Islas Canarias mostrando los depósitos dejados por las avalanchas de derrubios en los flancos submarinos (adaptado de Seisdedos *et al*. 2012). b) Flujo de derrubios del deslizamiento de El Julán (Masson 1996).

Conocer la evolución geológica de los archipiélagos volcánicos es clave a la hora de comprender la historia de colonización de los organismos. Las diferentes reconstrucciones de la Macaronesia muestran el papel que podrían haber tenido los islotes sumergidos a modo de islas puente (*stepping-stones*). Además, aunque el archipiélago canario nunca ha estado conectado con el continente, el brusco descenso del nivel del mar durante el Cuaternario pudo promover la emersión de estos montes submarinos (p. ej. Amanay) y la unión de algunas islas (p. ej. Lanzarote, Fuerteventura y el archipiélago de Chinijo estaban unidos en una única isla llamada Mahan, a tan solo 60 km del continente africano). En definitiva, la edad antigua de las Canarias, sumado a su proximidad al continente y la presencia de *stepping-stones*, es clave para comprender el papel que tienen estas áreas como refugios de diversidad. También parece destacable el posible papel de las zonas geológicamente estables, donde existe un gran número de linajes endémicos, por ejemplo, en las paleo-islas de Tenerife (Figura 18).

Tabla 2. Principales eventos geológicos que han intervenido en el origen de las Islas Canarias

Cronología	Eventos	Bibliografía
60–25 Ma (Paleógeno)	Emersión y posterior inmersión de las islas de Paleo-Macaronesia	Geldmacher *et al.* 2001, 2005
25–5 Ma (Mioceno)	Emergencia de las islas Canarias más antiguas: Fuerteventura (21 Ma), Ajaches (15), Amanay (15, ya sumergida), Gran Canaria (15), Famara (10), La Gomera (12), Adeje (11.5), Teno (8) y Anaga (6)	Fernández-Palacios *et al.* 2011
25–5 Ma (Mioceno)	Antes de producirse la actual erosión dominante y los mega-deslizamientos, las islas más viejas eran mucho más altas y grandes e incluían hábitats que hoy se han perdido. La elevación de Fuerteventura alcanzó 3300 m en el Mioceno.	Stillman 1999; Acosta *et al.* 2005
<2Ma	12 deslizamientos gigantes de más de 20km. Se estima una media de un mega-deslizamiento cada 125-170 Ka en el Pleistoceno	Krastel *et al.* 2001
<0.9 Ma	Los bajos niveles del mar predominan durante los periodos glaciares. Las especies o poblaciones que habían estado aisladas se reconectan cuando desciende el nivel del mar pudiendo ocurrir competición interespecífica e hibridación, o respuestas evolutivas que facilitan la coexistencia.	Ruddiman 2003
125 ka (último interglaciar)	La línea de costa de Fuerteventura se encontraba 5 metros por encima del nivel del mar actual. En esta época, las Islas Canarias eran ligeramente más pequeñas y estaban más aisladas que en la actualidad	Meco *et al.* 2006
120 ka-presente	Durante la mayor parte del último ciclo del nivel del mar todas las islas han sido más grandes y han estado mejor conectadas que hoy.	Rijsdijk *et al.* 2014
76–16 ka	Reducción de las áreas de hábitat y de la capacidad de carga de la mayoría de especies terrestres	Rijsdijk *et al.* 2014
LGM	En el Último Máximo Glaciar, la isla de La Gomera (373 km2) tuvo cerca de dos veces su superficie actual, mientras que Gran Canaria (1574 km2) fue 1.6 veces más grande de lo que es hoy	Rijsdijk *et al.* 2014
8 ka	Periodo húmedo importante	Meco *et al.* 2006
5.5 ka	Se establece el actual nivel del mar	Meco *et al.* 2003
<5000 años	Todas las islas (excepto La Gomera) han estado activas durante los últimos 5000 años	Schmincke 1998
<2000 años	Durante el último periodo de elevación del nivel del mar, y sólo en 2 milenios, muchas islas han reducido su superficie abruptamente más del 50% o se han dividido en varias islas.	Rijsdijk *et al.* 2014
Gran Canaria		
14.5–7.3 Ma	Emergencia de Gran Canaria	Schneider *et al.* 2004
5.5–3 Ma	Colapsos tectónicos debido a un fuerte episodio de actividad volcánica en el volcán Roque Nublo.	Pérez-Torrado *et al.* 1995; Marrero & Francisco-Ortega 2001; Emerson 2003
600.000-51.700 años	Formación de la depresión de Tirajana	Lomoschitz *et al.* 2002
Holoceno	24 erupciones Holocenas en Gran Canaria	Rodríguez-González *et al.* 2012
Tenerife		
11.6–6.4 Ma	Emergencia de Roque del Conde	Ancochea *et al.* 1990, 1999
8.0–3.2 Ma	Emergencia de Anaga	Ancochea *et al.* 1990, 1999
7.4–4.5 Ma	Emergencia de Teno	Ancochea *et al.* 1990, 1999
6 Ma	Avalancha de derrubios de Teno (dirección norte)	Cantagrel *et al.* 1999.
4.7–4.1 Ma	deslizamiento de Anaga (dirección norte)	Walter *et al.* 2005
3.5 Ma	Fusión de las tres proto-islas en la isla actual	Carracedo *et al.* 2007
3.3–1.9 Ma	La isla atraviesa un periodo de quiescencia volcánica y erosión.	Ancochea *et al.* 1990
2.6–2.3Ma	Avalancha de derrubios de Tigaiga (dirección norte)	Ibarrola *et al.* 1993; Ancochea *et al.* 1999; Cantagrel *et al.* 1999

1–1.1 Ma/ (0.83)	Mega deslizamiento de Güimar (avalancha de derrubios). Vinculado al tsunami que afectó el valle de Agaete en Gran Canaria (Paris *et al*. 2004; Pérez-Torrado *et al*. 2006)	1.1 Ma (Seisdedos 2006)/ <0.83Ma (Ancochea *et al*. 1990, Krastel *et al*. 2001).
<1 Ma	Vulcanismo intenso. 297 conos monogenéticos se encuentran todavía preservados en Tenerife y representan el tipo de actividad eruptiva más común que ha ocurrido en la isla durante el último millón de años	Dóniz *et al*. 2008
0.8 Ma	Pico de actividad volcánica en la Cordillera Dorsal	Llanes *et al*. 2003
0.6–1.3Ma	Avalancha de derrubios de Roques de García (dirección norte)	Cantagrel *et al*. 1999
0.54–0.69 Ma	Mega-deslizamiento de La Orotava (dirección norte)	Cantagrel *et al*. 1999
175±3 – 161±5 ka.	Avalancha de derrubios de Icod (dirección norte)	Boulesteix *et al*. 2012; 0.5-0.8 Ma en Ancochea *et al*. 1990
117±7 – 52±7 ka.	Emisiones de lavas	Boulesteix *et al*. 2012
52±7 – 18±1 ka	Inactividad volcánica	Boulesteix *et al*. 2012
18±1 ka	Se reanuda la actividad volcánica	Boulesteix *et al*. 2012
<35 Ka	Erupciones del complejo Pico Viejo. Varias erupciones de lavas en los últimos 35 ka	Ablay & Martí 2000; Martí *et al*. 2008
1.150 años	Última actividad volcánica del Teide (erupción de Lavas Negras)	Carracedo *et al*. 2007
La Gomera		
10.5– 6.4Ma	Emergencia de La Gomera	Ancochea *et al*. 2006
9.4–8.7 Ma	Colapso del LOE (Lower Old Edifice)	Llanes *et al*. 2009
8.6 – 6.4 Ma	Formación del UOL (Upper Older Edifice)	Llanes *et al*. 2009
5.7 – 4 Ma.	Construcción del nuevo edificio volcánico. El flujo de lava de este edificio no cubrió la isla entera: cubrió el centro y el sur y rellenó los barrancos profundos en el norte. Se dieron algunos eventos de colapso como las avalanchas de Tazo y de San Marcos	Ancochea *et al*. 2006
4.0 Ma	Fuerte erosión y destrucción del edificio volcánico. Fin de la principal actividad volcánica	Llanes *et al*. 2009
<2 Ma	En los últimos 2 millones de años se produce un periodo de quiescencia total	Carracedo & Day 2002
La Palma		
c. 3.5 Ma	Emergencia de La Palma	Carracedo *et al*. 2001
800–500 Ka	Formación del volcán Cumbre Nueva	Carracedo *et al*. 1999
566–549 ka	Mega-deslizamiento de Cumbre Nueva formando una gran ensenada en el centro-oeste de la isla	Carracedo *et al*. 1999
<500 años	Varias erupciones históricas se han dado en el rift principal y en las fisuras del flanco oeste de la isla.	Carracedo *et al*. 1999
El Hierro		
c. 1.1 Ma	Emergencia de El Hierro	Guillou *et al*. 1996
1.12 Ma y 882 ka	Formación del volcán Tiñor y posterior colapso en dirección noroeste. Tras el colapso del Tiñor probablemente se produjeron episodios explosivos del volcán hace unos 882 Ka.	Guillou *et al*. 1996
500–300 Ka	Colapso de El Julán (avalancha de derrubios). Formación de una ensenada de 16 km de ancho en el flanco suroeste de El Hierro.	Masson 1996
545–176 Ka	Avalancha de derrubios de San Andrés	Gee *et al*. 2001
176–145 ka	Avalancha de derrubios de Las Playas (dirección sur)	Carracedo *et al*. 2001
87±8 ka and 39±13 ka	Mega-deslizamiento de El Golfo, desplazando un tercio de la superficie terrestre de la isla dentro del mar (dirección noroeste). Formación de la espectacular ensenada de El Golfo – unos 15 kms de acantilados de hasta 1,4 km de altura	Urgeless 1997 (entre136 y 21 ka); ~15 ka en Masson 1996; Canals *et al*. 2000; Masson *et al*. 2002;
<37 ka	Erupciones de lava desde los rift principales cubriendo la mayor parte de la isla	Guillou *et al*. 1996; Carracedo *et al*. 2007

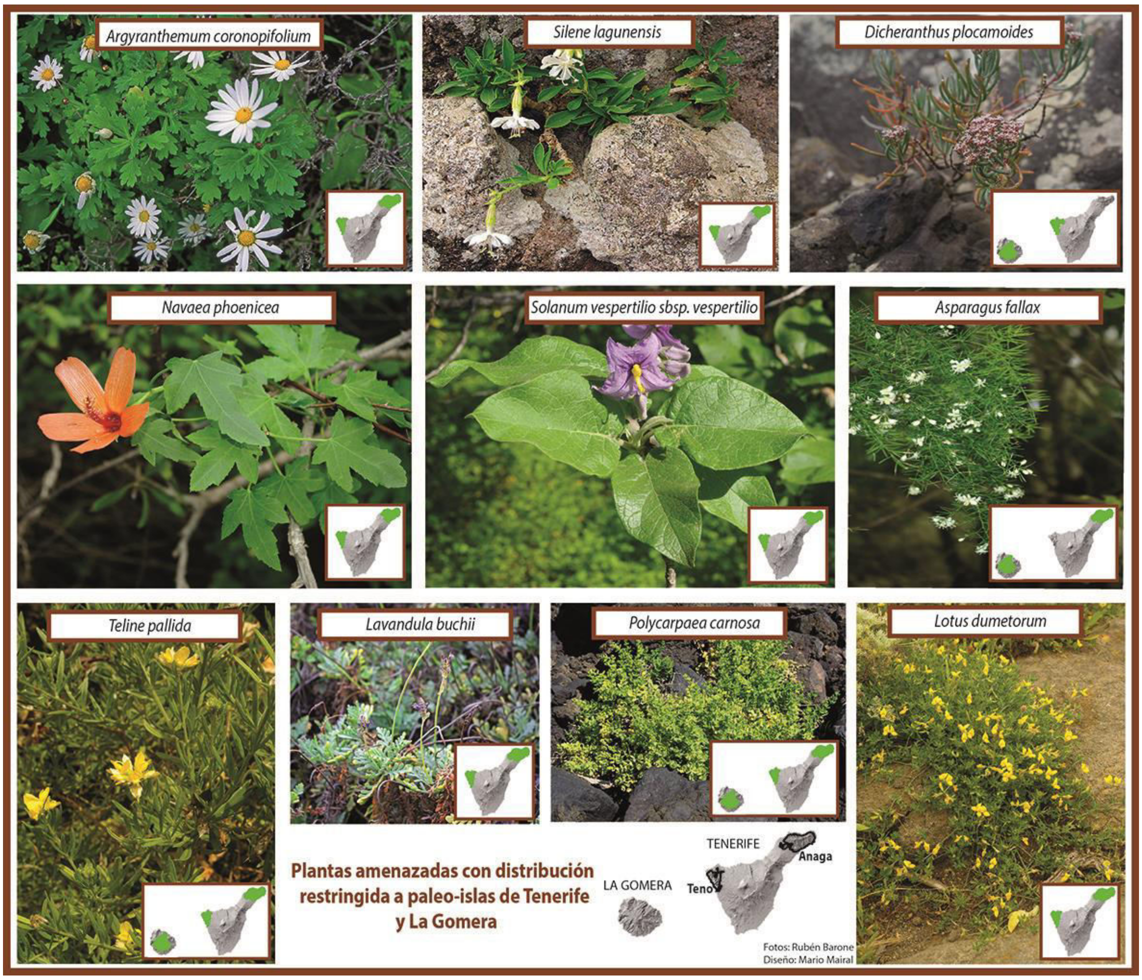


Figura 18. Ejemplos de algunas especies de plantas canarias restringidas y compartidas entre áreas estáticas geológicamente al menos los últimos 2 Ma: las tres paleo-islas de Tenerife y La Gomera.

La biodiversidad vegetal terrestre en las Islas Canarias: pisos de vegetación y afinidades biogeográficas.

En el año 1799, Alexander Von Humboldt junto a su compañero Aimé Bonpland visitaron la isla de Tenerife. En su ascensión al pico del Teide describieron por primera vez los pisos de vegetación de la isla, dando así el pistoletazo de salida de la geobotánica. Y es que en las Islas Canarias se pueden observar claramente varios pisos de vegetación con una marcada zonación altitudinal (Figura 19): 1) El cinturón costero xerofítico: dominado por la tabaiba dulce (*Euphorbia balsamifera*), la tolda (*Euphorbia aphylla*) y plantas costeras como la lechuga (*Astydamia latifolia*) y la uvilla de mar (*Zygophyllum fontanesii*). 2) El cardonal- Tabaibal: dominado por el cardón (*Euphorbia canariensis*) y otros arbustos suculentos del género *Euphorbia*, donde también destacan algunos géneros como los bejeques (*Aeonium*), las cerrajas (*Sonchus*) y las magarzas (*Argyranthemum*). 3) Los bosques termófilos: agrupan diversas formaciones, como palmerales (*Phoenix canariensis*), sabinares (*Juniperus turbinata* spp. *canariensis*), acebuchales (*Olea cerasiformis*), almacigares (*Pistacia atlantica*), lentiscares (*Pistacia lentiscus*) y retamares blancos (*Retama rhodorhizoides*) junto a otros elementos interesantes como el balo (*Plocama pendula*) y el drago (*Dracaena draco* spp. *draco*). 4) Los bosques de laurisilva: formaciones lauroides, que poseen 4 especies de lauráceas y unas 18 especies arbóreas. En las zonas de transición, por encima y por debajo de la laurisilva se da una formación más seca conocida como fayal-brezal, compuesta principalmente por la faya (*Morella faya*) y el brezo (*Erica arborea*). 5) El pinar: ecosistema endémico de Canarias donde domina totalmente el pino canario (*Pinus canariensis*). En su estrato arbustivo abundan jaras (*Cistus* sp., *Helianthemum* sp.), tomillos (*Micromeria* sp.) y codesos (*Adenocarpus* sp.). 6) Comunidades de alta montaña: comunidades arbustivas dominadas por leguminosas (*Spartocytisus supranubius*, *Teline microphylla*), donde destacan los tajinastes (*Echium* sp.) y las violetas (*Viola* sp.).

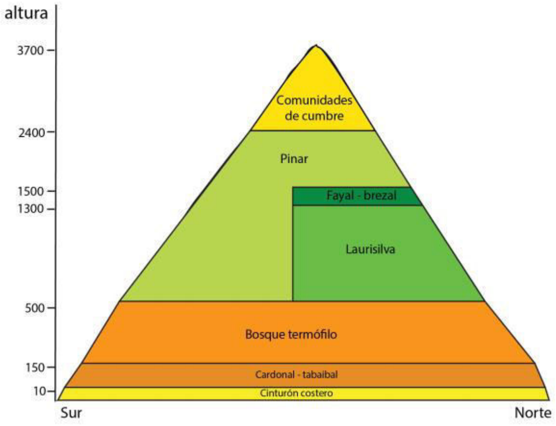


Figura 19. Pisos de vegetación en Canarias

Las afinidades biogeográficas de la flora canaria son diversas. De las actuales especies de plantas endémicas: un 35% tiene afinidades mediterráneas, un 25% tiene relaciones con el noroeste de África, un 22% está más relacionado con regiones más distantes (especialmente el este y el sur de África y el Nuevo Mundo) y, finalmente, un 18% ha derivado de ancestros macaronésicos (Caujapé-Castells *et al*. 2010). En esta tesis se hace especial hincapié en el ya comentado patrón Rand Flora; es decir, nos centramos en las afinidades de la flora macaronésica con la Península Arábiga, el este y el sur de África. Por ejemplo, algunas afinidades con la flora macaronésica que podríamos encontrar en un paseo por el sur de Yemen serían los géneros: *Aeonium*, *Campylanthus*, *Camptoloma*, *Dracaena*, *Pericallis* (*Cineraria*), *Plocama* y *Sideroxylon*. Más destacables son las afinidades macaronésicas con la isla de Socotra (Mies 1995), donde crecen especies directamente emparentadas con la flora macaronésica como: *Camptoloma lyperiifolium*, *Campylanthus spinosus*, *Plocama tinctoria* y *Campanula balfourii*. Otros géneros menos estudiados y también presentes en Socotra son: *Kleinia*, *Sideroxylon*, *Caralluma*, *Habenaria*, *Micromeria* y *Convolvulus* (hasta 37 géneros en común según Santos-Guerra 2013). Algunas especies presentes en los dos archipiélagos, y pendientes de confirmar sus relaciones filogenéticas corresponden a: *Aizoon canariense*, *Enneapogon desvauxii*, *Tricholaena teneriffae* y *Rumex lunaria*.

Algunas ideas sobre el origen de la biodiversidad canaria.

En el origen de la flora endémica del archipiélago canario han sido necesarias al menos unas 186 colonizaciones (Bramwell 1985), aunque investigaciones recientes apuntan que este número pudo haber sido mayor (Arjona *et al*. 2014). Una de las hipótesis más recientes para explicar estas colonizaciones es la hipótesis de ventanas de colonización (*Colonization Window Hypothesis*; Carine 2005), que propone que el establecimiento en las islas tiene oportunidades limitadas y está reducido a olas discretas de colonización (Carine 2005; Navarro-Pérez 2015). Además, para explicar las rutas más comunes en la colonización del archipiélago se ha propuesto la regla de progresión (*progression rule*; Whittaker & Fernández-Palacios 2007), que implica el movimiento de los propágulos desde las islas al este hacía las islas al oeste.

La flora canaria se caracteriza por un gran número de grupos monofiléticos con un alto número de endemismos (Francisco-Ortega *et al.* 2000; Silvertown 2004). Esto supone que una gran cantidad de grupos han colonizado el archipiélago en una sola ocasión, ocurriendo una diversificación posterior. Tal podría ser el caso en géneros como *Argyranthemum* (Francisco-Ortega *et al.* 1996), *Sonchus* (Kim *et al* .1996), *Echium* (Böhle *et al.* 1996), *Crambe (*Francisco-Ortega *et al.* 2002) y *Sideritis* (Barber *et al.* 2000, 2002)). De esta manera se ha calculado que 20 colonizaciones únicas han originado 269 especies en la flora canaria y que los géneros que han colonizado las islas en una sola ocasión poseen un número medio de especies endémicas siete veces mayor que los géneros que han colonizado las islas en múltiples colonizaciones (Silvertown 2004). Sin embargo, en otros grupos se han producido múltiples colonizaciones de manera independiente, explicación que parecería la más plausible debido a la proximidad de estas islas al continente (p. ej. en *Pulicaria* (Francisco-Ortega *et al.* 2001), *Navaea* (Fuertes-Aguilar *et al.* 2002), *Sonchus* (Kim *et al.* 2007) y *Scrophularia* (Navarro-Pérez *et al.* 2015)). Para explicar esta predominancia de grupos monofiléticos, Jonathan Silvertown (2004) introdujo la hipótesis *niche-preemption*. Esta hipótesis propone que los clados que han sufrido radiación adaptativa ocuparán más espacio del que podrían ocupar los colonizadores originales. Así, los miembros de un clado en el que se ha producido radiación adaptativa inhibirán el establecimiento de nuevos colonizadores más efectivamente que los colonizadores que no han radiado. Algunas alternativas a esta hipótesis han destacado el posible papel de la hibridación (Herben *et al.* 2005; Saunders & Gibson 2005). Finalmente, habría que destacar que las Islas Canarias no han actuado sólo como sumidero de linajes, sino que las últimas investigaciones también destacan su papel como cunas y fuentes de diversidad hacia otros archipiélagos (Böhle *et al.* 1996;

Barber *et al*. 2000; Trusty *et al*. 2005; Puppo *et al*. 2015) y hacia el continente (Carine *et al*. 2004; García-Verdugo *et al*. 2015; Patiño *et al*. 2015).

La dispersión entre islas habría tenido un papel más importante que las dispersiones continente-isla para explicar la diversificación en los linajes canarios (Francisco-Ortega *et al*. 2002; Sanmartín *et al*. 2008). Es este caso parece haber predominado una colonización horizontal entre ecosistemas similares, seguida de adaptación y especiación (Francisco-Ortega *et al*. 1996; 1999). Esto está apoyado por filogenias en las que diferentes taxones de un mismo clado habitan hábitats similares, lo que sugiere que la adaptación ecológica ha ocurrido solamente una vez. La diversificación entre islas se ha observado en diversos géneros como: *Sonchus* (Kim *et al*. 1996), *Aeonium* (Mort *et al*. 2002), *Adenocarpus* (Percy & Cronk 2002), *Crambe* (Francisco-Ortega *et al*. 2002) y *Lotus* (Allan *et al*. 2004). En otras ocasiones, todos los taxones forman un clado en la misma isla, por lo que la adaptación a hábitats similares en diferentes islas habría ocurrido de manera independiente, lo que se interpreta como radiaciones adaptativas. Algunos ejemplos son los géneros: *Sideritis* (Barber *et al*. 2000), *Teline* (Percy & Cronk 2002) y *Bystropogon* (Trusty *et al*. 2005).

En cuanto a la dimensión temporal, tradicionalmente se ha creído que una gran proporción de la flora canaria pertenecía a linajes antiguos que se extinguieron en el continente debido a los cambios climáticos del Terciario. Sin embargo, el auge de los estudios filogenéticos ha cambiado esa visión, revelando que una gran cantidad de linajes colonizaron las islas en un tiempo relativamente reciente (p. ej. *Sonchus* (2.8-2.3 Ma; Kim *et al*. 1996) *Echium* (0.75-30 Ma; Böhle *et al*. 1996), *Argyranthemum* (1.5-3.0 Ma; Francisco-Ortega *et al*. 1997), así como un abanico de linajes ligados a los bosques lauroides (Plio-Pleistoceno; Kondraskov *et al*. 2015)).

En Canarias la riqueza genética posee niveles relativamente altos de diferenciación interpoblacional en relación con otros archipiélagos (Francisco- Ortega *et al.* 2000; Pérez de Paz & Caujapé-Castells 2013; Melonti *et al.* en prensa), valores que hay que tomar con precaución ya que podrían ser consecuencia de un sesgado muestreo intrapoblacional (Caujapé-Castells 2009). Para explicar esta alta diversidad genética y la monofilia predominante en linajes canarios, Javier Francisco-Ortega *et al.* (2000) propusieron múltiples colonizaciones e hibridación. En relación a la diversidad genética, una de las hipótesis más elaboradas es la del "singameón surfista" de Juli Caujapé Castells (2011). Esta hipótesis argumenta que las Islas Canarias más antiguas, con menor altura y mayor superficie, albergan mayores niveles de mezcla y variación genética. Se apoya en que la ola de colonización este-oeste está asociada con un decrecimiento en los niveles de diversidad génica desde el este hacia el oeste; las islas del este podrían actuar como lugares de mezcla, donde las colonizaciones frecuentes desde el continente africano, junto a la uniformidad ecológica y el suave relieve, han dado lugar a una mayor mezcla genética y altos valores de diversidad genética. Aquí, debido a la proximidad filogenética entre genomas colonizadores, los procesos de colonización habrían creado singameones, dando lugar a especies incipientes que en ocasiones mantienen cruzamientos compatibles. Por otro lado, los procesos evolutivos en las islas occidentales implicarían especiación e interrupción del flujo génico, lo que explicaría el mayor número de endemismos en estas islas (Caujapé-Castells 2011). Hipótesis alternativas, a la vez que complementarias, predicen que las tres islas centrales (Gran Canaria, Tenerife y La gomera) son los centros de dispersión y diversificación predominantes en el archipiélago y muestran las mayores tasas de intercambio génico con el continente (Sanmartín *et al.* 2008). Aquí podrían haber afectado fenómenos de extinción provocados por fenómenos catastróficos que habrían

creado nichos ecológicos vacíos, posteriormente colonizados desde las islas centrales (p. ej. *Sideritis* (Barber *et al.* 2000) y *Micromeria* (Meimberg *et al.* 2006)).

Conservación de la diversidad vegetal canaria

Se cree que las islas fueron colonizadas hace unos 2500 años por un grupo étnico relacionado con los bereberes, los cuales introdujeron el uso del fuego, la agricultura y la ganadería, modificando así las condiciones ecológicas de las islas (de Nascimento *et al.* 2009). En el siglo XV se produjo la colonización y conquista castellana y en la actualidad las islas están habitadas por más de 2.200.000 personas. Todo esto, junto a otros impactos como el turismo de masas, ha puesto en jaque la biodiversidad del archipiélago.

En Canarias una gran cantidad de poblaciones de especies vegetales están extremadamente amenazadas y no alcanzan un tamaño mínimo poblacional viable. Existen decenas de especies que no superan las 10 poblaciones, hay alrededor de una decena de especies que han quedado reducidas a una población (p. ej. *Kunkeliella canariensis*, *Solanum vespertilio* spp. *doramae*, *Monanthes wildpretii*, *Lotus eremiticus* y *Sideritis amagroi*), mientras que otras están extintas en estado silvestre (p. ej. *Kunkeliella psilotoclada*, *Helianthemum cirae* y *Aeonium mascaense*). Además, es especialmente urgente conservar la elevada variación genética intraespecífica existente en el archipiélago. Esta variación corresponde al nivel fundamental de la diversidad que proporciona la base del cambio evolutivo y, por tanto, es crucial para mantener las capacidades de las especies para adaptarse a nuevas condiciones. Conservar esta diversidad podría verse aún más agravado por la tala masiva y la construcción de carreteras que han permitido que se establezcan *corredores genéticos*. Algo especialmente preocupante ya que facilita la conexión de poblaciones o especies aisladas históricamente, con consecuencias impredecibles para la conservación de la biodiversidad

(Francisco-Ortega & Santos-Guerra 2001). De esta manera, la reducción del tamaño poblacional y el movimiento de material vegetal endémico entre poblaciones, pueden ser críticos para la conservación del rico patrimonio vegetal canario. Además de otros múltiples factores como: destrucción del hábitat, ajardinamiento, especies invasores vegetales y animales (principalmente cabras, gatos, ratas, conejos y abejas melíferas), etc.

En una situación ideal, cualquier programa de conservación debe estar basado en estudios de diversidad génica que permitan desarrollar estrategias para la conservación del rico patrimonio vegetal canario. Han de aplicarse programas de conservación *in situ* y *ex situ* para evitar la pérdidas de formas alélicas. Un ejemplo serían los programas de conservación *ex situ* desarrollados por el banco de semillas y de ADN del Jardín Botánico `Viera y Clavijo´ en Las Palmas de Gran Canaria. También en los últimos años se están tomando medidas de conservación *in situ* (p. ej. los proyectos LIFE), para las que se recomienda mucha cautela en los programas que impliquen movimientos de plantas entre poblaciones.

El bosque Afromontano y las Islas Canarias en esta tesis

En el **Capítulo 1** se reconstruye la historia espacio-temporal del género *Canarina,* habitante del archipiélago Afromontano y las laurisilvas canarias. En el **Capítulo 2** se estudia una especie ligada al bosque Afromontano: el epífito *Canarina eminii*. En el **Capítulo 3** se trabaja con una especie ligada a los bosques lauroides: el endemismo canario *Canarina canariensis*. En el **Capítulo 4** se explora la demografía reciente de estos bosques. En los **Capítulos 5-6** se hace especial hincapié en las relaciones Afromontanas –mediterráneas/macaronésicas y en los puentes de vegetación.

4. El género Canarina

El género *Canarina* pertenece a la familia Campanulaceae, presentando una posición basal dentro de la tribu Platycodoneae (grupo hermano de las otras dos tribus, Campanuleae y Wahlenbergieae; Cosner *et al.* 2004; Roquet *et al.* 2009; Olesen *et al.* 2012; Mansion *et al.* 2012). De los nueve géneros reconocidos para Platycodoneae (Wang *et al.* 2103) sólo *Canarina* aparece en África, mientras que el resto de géneros (*Campanumoea* Blume, *Codonopsis* Wall, *Cyananthus* Wall. Ex Benth., *Cyclocodon* Griffith ex J.D. Hooker & Thomson, *Echinocodon* D.Y. Yong, *Leptocodon* Lem., *Ostrowskia* Regel y *Platycodon* Jacq. A. DC.) son todos endémicos de las montañas del centro y este de Asia (**Capítulo 1**, Figura 1a). El género *Canarina* está formado solamente por tres especies: *Canarina canariensis*, endémica de las Islas Canarias y *C. abyssinica* y *C. eminii* distribuidas en las montañas del este de África, desde Etiopía hasta Tanzania. Por lo tanto, observamos dos grandes disyunciones biogeográficas, i) a nivel de tribu, entre todos los géneros de Platycodoneae y el género *Canarina*, ii) dentro del género *Canarina*, entre las montañas del este de África y las Islas Canarias. Esta distribución actual de *Canarina* ha llevado a que se le considere un miembro de un antiguo bosque subtropical con una distribución más amplia en África y en la cuenca mediterránea, y más tarde fragmentado a causa de eventos de aridificación (Axelrod & Raven 1978; Bramwell 1985).

La primera especie del género *Canarina* fue establecida por Linneo en su obra *Species Plantarum* (1753), donde denominó a la especie canaria como *Campanula canariensis*, luego renombrada por él mismo como *Canarina campanula* (1771) y un siglo después se le dio su nombre actual *Canarina canariensis* (L.) Vatke. (1874). Más tarde fueron descritas las especies del este de África, *Canarina eminii* Aschers. ex Schweinf. (1892) y *Canarina abyssinica* Engl. (1902). Los holotipos de las especies

africanas se conservaban en el Herbario de Berlín pero fueron destruidos en un fuego durante la Segunda Guerra Mundial, lo que luego probablemente ha originado numerosos problemas para la identificación y clasificación de ejemplares recogidos posteriormente. Aunque en *Index Kewensis* aparecen hasta siete especies del género *Canarina*, posteriormente cuatro de éstas han sido descartadas (*C. zanguebar* (nombre sin resolver), *C. moluccana* (ahora *Cyclocodon lancifolius*), *C. laevigata* (sinónimo de *C. canariensis*) y *C. elegantissima* (sinónimo de *C. eminii*). Muy destacable es la monografía del género *Canarina* del botánico sueco Karl Olov Hedberg (1961b) que incluye un estudio morfológico, cariológico, descripciones, mapas de distribución, y abundante información relativa al género. A continuación se detallan las principales características de cada una de las tres especies.

Figura 20. Probablemente la primera ilustración científica de *Canarina*. Incluida en *Hortus cliffortianus* (Linnaeus 1738)

Canarina canariensis (L.) Vatke. (1874) es una herbácea terrestre y trepadora característica de zonas marginales de monteverde y en ocasiones de los bosques termófilos canarios. Crece a una altitud entre 300-1100 m. Se distingue bien por sus hojas hastadas, su corola y su cáliz a menudo reflejo. Sus grandes flores rojas o anaranjadas aparecen en invierno y poseen un gran volumen de néctar. Está distribuida en las Islas Canarias centrales y occidentales: Gran Canaria, Tenerife, La Gomera, La Palma y El Hierro (**Capítulo 1**, Figura 1a). Crece asociada a enclaves de bosque de laurisilva y monteverde, con preferencia por zonas ligeramente abiertas, donde entra la luz.

Actualmente una de sus mayores amenazas proviene de la herbivoría floral provocada por las ratas introducidas en la laurisilva.

Canarina eminii Aschers. ex Schweinf. (1892) tiene un hábito epífito, ascendiendo a las zonas más altas del dosel arbóreo del bosque Afromontano, a una altitud entre los 1500 y los 3200 metros. Son características sus hojas cordadas y su ovario acanalado. Aunque suele diferenciarse muy bien de *C. abyssinica* en vivo, en ausencia de buenos caracteres en pliegos de herbario (flores, peciolos), el mejor carácter taxonómico es la proporción de la lámina de la hoja respecto a la longitud del peciolo (Hedberg 1961b). Se encuentra restringida a los retazos de bosque Afromontano que salpican el este de África desde Etiopía hasta el sur del lago Tanganika (**Capítulo 4**, Figura 1c), donde su mayor amenaza es la destrucción, deterioro y el alarmante retroceso de estos bosques.

Canarina abyssinica Engl. (1902) tiene un hábito trepador, creciendo en matorrales montanos abiertos de las tierras altas del este de África y en zonas rocosas de selvas abiertas entre 1350 y 2000 metros de altitud. Es muy característico su pedicelo largo y enrollado, su peciolo enrollado y su ovario ancho. Su distribución se centra en el sur de Sudán, centro y sur de Etiopía, norte y este de Uganda, y norte del lago Tanganica (**Capítulo 4**, Figura 1c). Es una especie rara y poco conocida, que ha desaparecido en muchas de sus antiguas localidades y sufre probablemente un alarmante retroceso.

Aunque el área de las especies africanas pueda estar ligeramente superpuesta, nunca se han encontrado híbridos y parecen estar completamente separadas ecológicamente. En las tres especies el número básico de cromosomas es 17 y todas son diploides (Hedberg 1961b).

A pesar del interés biogeográfico de *Canarina*, los trabajos realizados hasta la fecha no incluyen una filogenia completa del género. Los estudios filogenéticos previos se han centrado en desvelar la posición del género dentro de la tribu (Platycodoneae; Zhou *et al.* 2013; Wang *et al.* 2013) o en un contexto más amplio a nivel de familia (Campanulaceae; Roquet *et al*. 2009; Mansion *et al*. 2012). El trabajo reciente de Wang *et al.* (2013) divide la tribu en cuatro grandes clados y posiciona a *Canarina* dentro de un clado que incluye los géneros del este de Asia *Platycodon*, *Cyclocodon* y *Echinocodon*. Aunque estos autores no incluyeron el género *Ostrowskia*, anteriormente propuesto como grupo hermano del género *Canarina* (Mansion *et al*. 2012). Lo más común es que estos estudios incluyan a la especie canaria *C. canariensis*, pero no a las especies africanas, más difíciles de muestrear. Olesen *et al.* (2012) ha sido el único estudio que ha incluido las tres especies de *Canarina*, aunque el trabajo estaba centrado en especies de Campanuláceas de islas oceánicas y el muestreo dentro de Platycodoneae era muy limitado.

Especial atención se ha prestado al síndrome ornitófilo de *Canarina*. Las especies de *Canarina* son polinizadas por pájaros (Vogel *et al.* 1984; Olesen 1985; Valido *et al.* 2004; Rodríguez-Rodríguez & Valido 2011; Olesen *et al*. 2012). La polinización de flores por aves es prácticamente inexistente en Europa y el noroeste de África, fenómeno que, sin embargo, sí se observa en algunas especies de Macaronesia y se ha denominado *elemento ornitófilo macaronésico* (Figura 21; Vogel *et al*. 1984; Dupont *et al*. 2004). Este síndrome ornitófilo es especialmente interesante: mientras que en el este de África las especies de *Canarina* son polinizadas por pájaros especializados en consumir néctar (familia Nectariniidae)**,** en Canarias es polinizada por pájaros más generalistas (p.ej. *Phylloscopus canariensis* y *Cyanistes caeruleus*). Además, el género *Canarina* es el único que todavía posee parientes continentales polinizados por pájaros, lo que sugiere que en *C. canariensis* la ornitofilia podría ser una condición relicta (Fuertes-Aguilar *et al.* 2002; Valido *et al*. 2004). A pesar de esto, los trabajos en biología reproductiva son escasos. Rodríguez-Rodríguez y Valido (2011) estudiaron la especie *C. canariensis* y revelaron que, a pesar de ser una especie autocompatible, tiene una capacidad muy baja para la autofecundación y una gran limitación de polen. Aunque los pájaros oportunistas aumentan la producción de semillas, estos beneficios se ven contrarrestados por la alta incidencia de la herbivoría floral a causa de babosas nativas y ratas introducidas.

Todas las características anteriores parecen señalar a *Canarina* como un candidato perfecto para testar la disyunción biogeográfica del patrón Rand Flora (**Capítulo 1**). Se espera que con el desarrollo de estudios biogeográficos, filogeográficos y de genética de poblaciones, se desvele una biodiversidad única, donde poblaciones persistentes a largo plazo con diferentes orígenes temporales revelen pistas sobre los patrones y procesos que han prevalecido en sistemas insulares. Para esto la tesis se centra en las dos especies de *Canarina* restringidas a sistemas insulares (Figura 22), *Canarina eminii* (**Capítulo 2**) y *Canarina canariensis* (**Capítulo 3**). Además, la amplia distribución de *Canarina* en los últimos bosques subtropicales del continente africano, lo convierte en un grupo ideal para realizar estudios aplicados a demografía y biología de la conservación, los cuales podrían orientarnos sobre las posibles consecuencias del cambio climático (**Capítulo 4**).

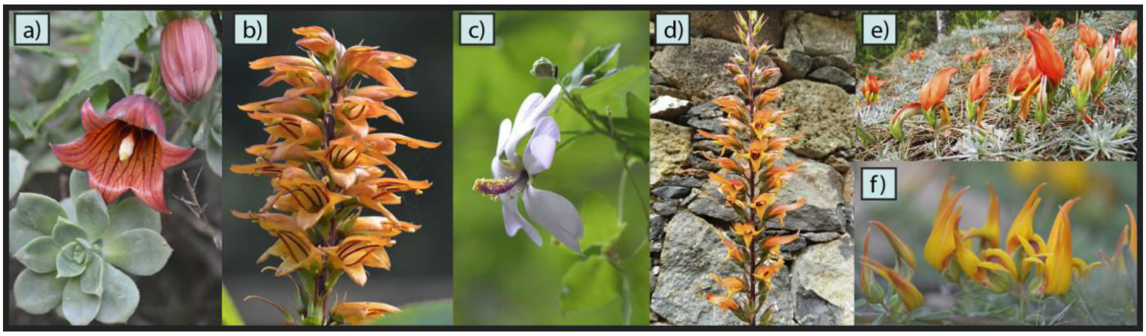


Figura 21. Algunas de las especies que integran el elemento ornitófilo macaronésico. a) *Canarina canariensis.* b) *Isoplexis canariensis*. c) *Lavatera acerifolia*. d) *Isoplexis chalcantha*. e) *Lotus pyranthus*. f) *Lotus maculatus*.

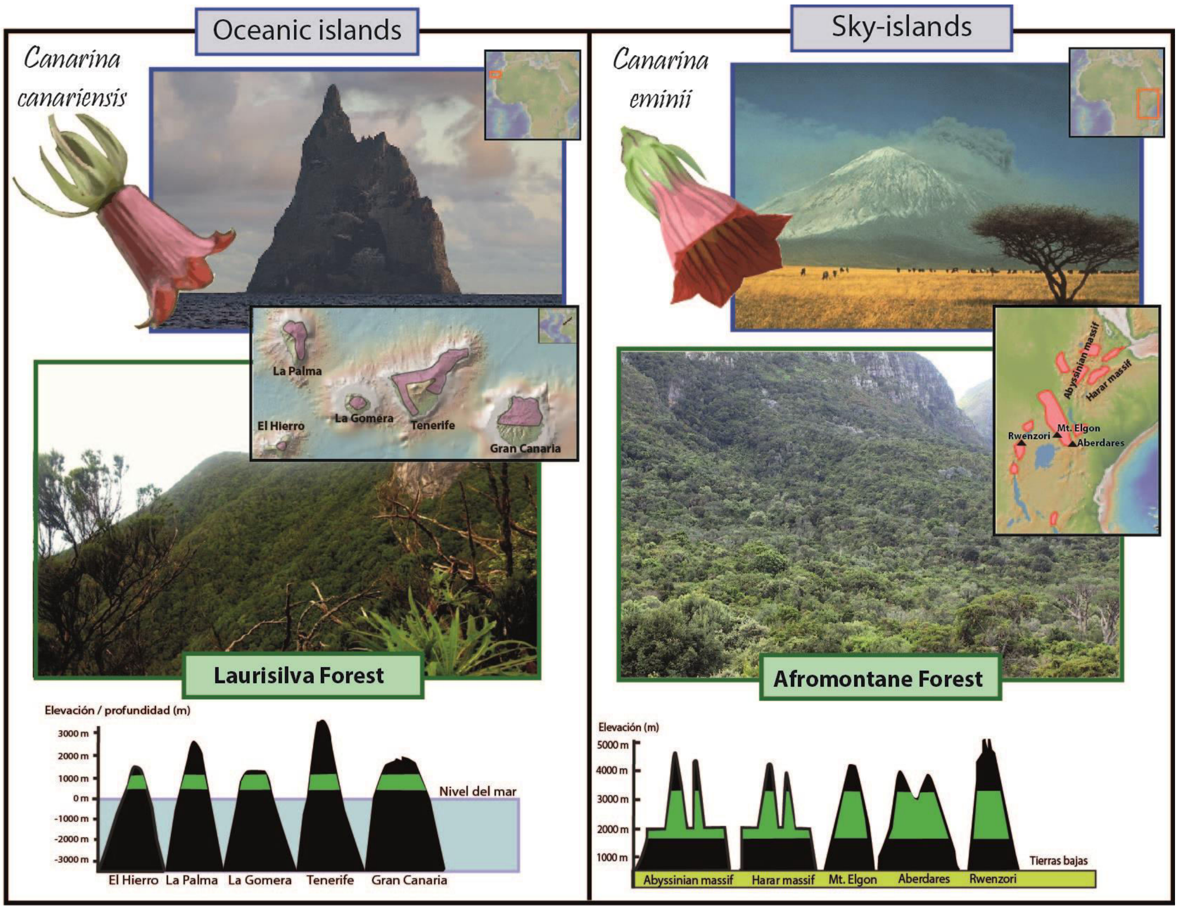


Figura 22. Sistema de estudio para el género *Canarina* en esta tesis. Fuente fotos (islands y Afromontane forest): Wikimedia commons.

IV. Objetivos e hipótesis

El objetivo de esta tesis es estudiar la historia evolutiva del género *Canarina* para tratar de comprender los patrones y procesos macroevolutivos y microevolutivos que han generado su distribución actual. La singular distribución y el hábitat del género, señalan a *Canarina* como un sistema ideal para testar patrones y procesos que han resultado de la aridificación dentro del continente africano, así como en diferentes sistemas insulares. Además, se ha extendido la investigación con *Canarina* a un conjunto de linajes con distribución Rand Flora, con el objetivo de conocer los principales factores que determinan la distribución de las plantas en respuesta a la aridificación.

Objetivos Capítulo 1: el objetivo principal es reconstruir la historia espacio-temporal del género *Canarina* y, adicionalmente, utilizar esta historia para discriminar entre las dos hipótesis establecidas para el origen del patrón Rand Flora: vicarianza con extinción climática o eventos de dispersión. Los objetivos fueron:

- Establecer las relaciones filogenéticas del género *Canarina* utilizando marcadores nucleares y cloroplásticos.
- Estimar tiempos de divergencia. Para esto se propone utilizar un nuevo método de datación "anidada" que afronte el problema de la carencia de fósiles en el grupo de estudio y la subsiguiente incertidumbre en la datación.
- Reconstruir el área ancestral y los principales eventos de migración.
- Determinar el rango ancestral de *Canarina* utilizando técnicas de modelización de nicho y datos paleoambientales.

Objetivos Capítulo 2: el objetivo principal fue realizar un estudio filogeográfico de la especie epífita de bosque Afromontano *Canarina eminii*, una especie que habita islas continentales (*sky islands*) a ambos lados de la fractura geológica del Gran Valle del Rift.

La hipótesis es que si la ruptura continental es un proceso gradual, también se esperaría una separación gradual de los linajes en el espacio, con un efecto observable a nivel intraespecífico. Además, nos centramos en las dos hipótesis clásicas de los estudios filogeográficos del bosque Afromontano: i) la hipótesis *mountain forest bridge* postula que la dispersión entre islas continentales fue más probable cuando los bosques Afromontanos estaban más extendidos, durante los previos y el presente interglaciar, antes del impacto de la agricultura; ii) la hipótesis alternativa postula eventos de dispersión aleatorios entre montañas, facilitados por vientos monzónicos o migraciones de pájaros. Otros objetivos fueron:

- Determinar la distribución geográfica de la diversidad genética en *C. eminii* mediantes marcadores cloroplásticos y nucleares (AFLPs).
- Establecer las relaciones intraespecíficas de los linajes genealógicos mediante redes de haplotipos.
- Explorar los patrones y procesos que han modelado los patrones de biodiversidad genética en el bosque Afromontano, en particular, el papel jugado por el Gran Valle del Rift como barrera para el flujo génico.
- Determinar si la influencia de las fluctuaciones climáticas del Pleistoceno han afectado en la generación de la variabilidad y estructura genética.

Objetivos Capítulo 3: el principal objetivo fue realizar un estudio filogeográfico de *Canarina canariensis*, una especie ampliamente distribuida en el archipiélago canario, ligada a los principales bosques de laurisilva. Esperamos que los patrones de variación genética sean jerárquicos en islas. En concreto esperamos que la variación genética entre islas sea mayor que la variación genética dentro de islas, ya que las barreras oceánicas son por lo general más fuertes que las barreras topográficas en provocar aislamiento en los archipiélagos. Además, esperamos que el estudio de especies ampliamente distribuidas en islas volcánicas pueda revelar la existencia de biodiversidad críptica, así como evidencias de diversificación y extinción que podrían estar ligadas a los eventos geológicos característicos de islas volcánicas (p. ej. vulcanismo activo y deslizamientos de ladera). Los objetivos fueron:

- Determinar la distribución geográfica de la diversidad genética en *C. canariensis* mediantes marcadores cloroplásticos y nucleares (AFLPs).
- Comparar los patrones de variación genética debido a barreras oceánicas (entre islas) o barreras topográficas y ecológicas (dentro de islas) en una especie ampliamente distribuida.
- Establecer las relaciones intraespecíficas de los linajes genealógicos mediante redes de haplotipos. Incorporar la incertidumbre asociada a la inferencia de estas redes para evaluar estadísticamente escenarios filogeográficos alternativos.
- Buscar áreas ancestrales y reconstruir eventos de migración entre islas.
- Inferir la historia demográfica.
- Examinar el papel de las paleo-islas de Tenerife como refugios de diversidad genética.

Objetivos Capítulo 4: en este capítulo el principal objetivo es evaluar la necesidad de conservación de los últimos relictos de vegetación subtropical que quedan alrededor del continente africano; biodiversidad altamente amenazada por el cambio global. Para ello estudiamos la demografía reciente de dos especies ampliamente distribuidas y ligadas a los ecosistemas subtropicales de África: *C. eminii* en el bosque Afromontano del este de África y *C. canariensis* en las laurisilvas de las Islas Canarias. Creemos que debido al dramático retroceso del bosque Afromontano y de la laurisilva canaria, deberíamos detectar un fuerte declive demográfico. Los objetivos fueron:

- Desarrollar marcadores microsatélites para las dos especies.
- Determinar valores de diversidad genética en poblaciones naturales
- Testar escenarios demográficos alternativos mediante técnicas de Computación Bayesiana Aproximada.
- Utilizar la información aportada por la variabilidad genética y la demografía para proponer estrategias adecuadas que se puedan utilizar para la conservación de estos ecosistemas.

Objetivos Capítulo 5: el objetivo principal es establecer un marco temporal para el patrón Rand Flora. La hipótesis principal es que una aridificación gradual por cambio climático ha relegado a los linajes Rand Flora a los extremos del continente. Si ha habido una gran extinción por cambio climático, debería poder detectarse en las reconstrucciones filogenéticas. Los objetivos detallados fueron:

- Revisar y describir la distribución geográfica y relaciones filogenéticas de algunas de las principales disyunciones que conforman el patrón Rand Flora.
- Inferir nuevas relaciones filogenéticas y tiempos de divergencia utilizando un método de datación homogéneo para 17 linajes de plantas que presentan disyunción Rand Flora.
- Inferir tasas de diversificación para cada linaje, en un intento de dilucidar el papel que los procesos evolutivos, como la extinción provocada por cambio climático, pueden haber desempeñado en la formación del patrón Rand Flora.

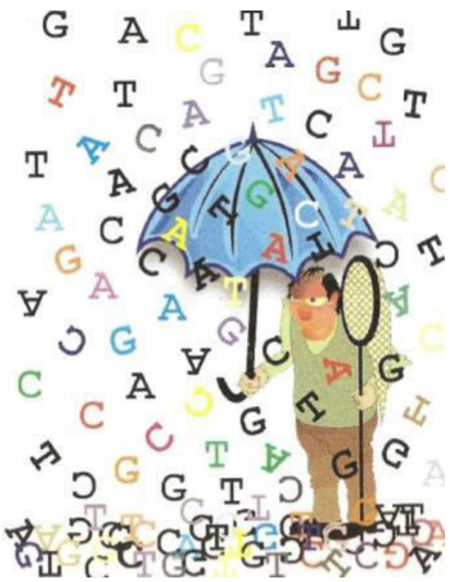
Objetivos Capítulo 6: el objetivo principal es estudiar si existe una relación entre el nicho climático ocupado por las especies y la temporalidad de la vicarianza, utilizando como modelo las disyunciones Rand Flora. La hipótesis es que si el cambio climático ha alterado el rango ancestral de las especies gradualmente (p. ej. gradientes de humedad o

temperatura), la edad de divergencia de las especies disyuntas debería estar relacionada con la tolerancia de los linajes a las condiciones climáticas. Bajo el paradigma de la conservación de nicho climático se espera que las especies cambien su rango siguiendo las condiciones climáticas adecuadas, a un ritmo establecido por sus tolerancias climáticas. Los objetivos fueron:

- Reconstruir el nicho ecológico para 16 linajes Rand Flora mediante técnicas de modelización espacial y datos paleoclimáticos.
- Cuantificar el nicho climático de cada linaje mediante variables bioclimáticas.
- Buscar correlaciones entre el nicho climático ocupado y la edad de divergencia de los linajes.

V. Metodología

En este apartado se hace una aproximación a la evolución histórica y estado actual de la metodología empleada en esta tesis: métodos de reconstrucción filogenética, marcadores moleculares empleados, datación molecular, técnicas de inferencia filogeográfica, y modelización de nicho ecológico. Información adicional del área y del grupo de estudio se ha proporcionado en el apartado III. Antecedentes. Para una descripción más detallada de los métodos y herramientas de inferencia empleados para cada objetivo, ver el apartado *Material & Métodos* (*Materials & Methods*) de cada capítulo.



Entomólogo “clásico” presenciando el aluvión molecular de este apartado.

1. Las reconstrucciones filogenéticas

El propósito de identificar los seres vivos que ha acompañado a la humanidad desde sus orígenes, junto a los últimos siglos de trabajo taxonómico, ha sentado una base sólida para la clasificación de los organismos vivos (Aristóteles 384-322 A.C.; Linnaeus 1753, 1788; Haeckel 1866; Whittaker 1969; Margulis 1974; Woese *et al*. 1990; Cavalier-Smith 1998). Si en un principio esta clasificación se basó en características de la morfología o fisiología de los organismos, la popularidad en las últimas décadas del uso de marcadores moleculares en la reconstrucción de relaciones filogenéticas ha revolucionado algunas de estas clasificaciones; ejemplos destacados son la propuesta de unir grupos tan distintos como artrópodos y nemátodos dentro de un clado común como el clado Ecdysozoa (Aguinaldo *et al*. 1997), o la adscripción de diversos géneros de escrofulariáceas dentro de la familia de las plantagináceas (Olmstead *et al*. 2001). Los métodos modernos de inferencia filogenética utilizan matrices con información molecular, a veces junto con morfológica, para reconstruir las relaciones evolutivas entre organismos mediante árboles filogenéticos. Una filogenia es un árbol que contiene nodos conectados por ramas, y en el que cada nodo representa el nacimiento de un nuevo linaje. En el caso de un árbol de especies, cada nodo representaría un evento de especiación. Si la filogenia contiene una dimensión temporal (árbol ultramétrico), cada rama representa la persistencia de un linaje genético a través del tiempo. Los métodos de inferencia filogenética han evolucionado desde la inferencia de árboles por parsimonia, donde sólo se representan las relaciones entre linajes, hacia métodos probabilísticos de evolución molecular, en los que se estima la topología del árbol y la longitud de las ramas, que representan la tasa de evolución entre linajes. Para ello, se evalúa la probabilidad de que una hipótesis evolutiva considerada haya generado un alineamiento de secuencias de ADN. En la presente tesis, los alineamientos de secuencias de ADN serían los datos observados, mientras que la incógnita sería el árbol filogenético.

1.1 Métodos de construcción de árboles.

Para construir árboles filogenéticos se utilizan principalmente dos tipos de métodos: métodos de distancias (p. ej. Mínimos cuadrados (Cavalli-Sforza & Edwards 1967) y Neighbor-Joining (Saitou & Nei 1987)) y métodos que usan caracteres discretos (p. ej. parsimonia (Fitch 1971), máxima verosimilitud (Fisher 1920) e inferencia Bayesiana (Rannala & Yang 1996)). Los métodos de distancias, máxima verosimilitud e inferencia Bayesiana utilizan un modelo de sustitución; son por tanto, métodos *model-based* (basados en modelos). Mientras que la parsimonia no utiliza un modelo explicito y sus asunciones están implícitas. La Parsimonia utiliza un método estadístico no paramétrico, se basa en la navaja de Occam que, ante hipótesis alternativas elige la que supone el mínimo número de asunciones. En este caso, el menor número de cambios entre estados de carácter. Los métodos que usan caracteres discretos comparan todas las secuencias en un alineamiento y a la vez consideran a cada carácter para calcular un valor para cada árbol. Este valor es el número mínimo de cambios en parsimonia, el *log-likelihood* en máxima verosimilitud y la probabilidad posterior en inferencia Bayesiana. En teoría, el árbol con el mejor valor debería ser identificado por la comparación de todos los árboles posibles. En la práctica, debido al gran número de árboles posibles, una búsqueda exhaustiva sólo es computacionalmente factible para sets de datos pequeños. En su lugar se utilizan algoritmos de búsqueda heurística (Yang & Rannala 2012). A continuación se hace hincapié en los métodos que usan caracteres discretos utilizados en esta tesis, la Máxima Verosimilitud y, especialmente, la Inferencia Bayesiana.

La Máxima Verosimilitud (MV) es un método de estimación que consiste en encontrar los valores de los parámetros del modelo. Se estiman los valores de cada parámetro del modelo y se comparan las verosimilitudes de los diferentes modelos, para así escoger la topología del más verosímil. Esto implica considerar todas las posibles combinaciones de estados de los nodos internos del árbol. La verosimilitud no está en función de los datos, sino que está en función de los parámetros utilizados en el modelo estadístico. El soporte de los nodos consiste en un remuestreo con repetición, medida conocida como *bootstrap*. Además se utilizan pruebas de razón de verosimilitud (*likelihood ratio test*; Goldman 1993) para detectar el modelo evolutivo al que mejor se ajustan los datos. Algunas de las ventajas de la MV sobre métodos anteriores son: el uso de modelos evolutivos, que la asunciones de los modelos son explícitas y pueden ser evaluadas y mejoradas y que se reduce el efecto de atracción de ramas largas. Sin embargo la MV es computacionalmente más costosa.

RESUMEN

La Inferencia Bayesiana (IB) se basa en el cálculo de la probabilidad posterior de un árbol a partir de unos datos observados y un modelo evolutivo; es decir, calcula la probabilidad de que un árbol sea correcto. Esto diferencia a la IB de métodos anteriores que calculaban la probabilidad de que los datos se adapten al árbol. La IB se basa en el Teorema de Bayes (Bayes 1763), en el que el conocimiento previo en forma de distribución probabilística (*prior probability*) se combina con una función de verosimilitud de los datos para generar la probabilidad posterior de los parámetros según el modelo evolutivo. Debido a la complejidad de los modelos utilizados en inferencia filogenética, que implican múltiples parámetros, y a la dificultad de evaluar todos los árboles posibles, estas probabilidades no se estiman analíticamente sino que se aproximan mediante algoritmos de simulación estocástica de Cadena de Markov (*Markov Chain Monte Carlo* en inglés, o sus siglas más conocidas MCMC), cuya implementación se popularizó en biología evolutiva con el programa MrBayes (Huelsenbeck & Ronquist 2001) y, posteriormente con el programa BEAST (Drummond *et al*. 2006). Una de las principales ventajas de IB sobre MV es que permite acomodar la incertidumbre en la estima de los parámetros del modelo, incluida la incertidumbre filogenética. Los parámetros del modelo se consideran en IB como variables aleatorias con una distribución estadística, mientras que en MV son constantes fijas desconocidas. Esto, junto con las aproximaciones MCMC, permite el uso de modelos evolutivos complejos, en los que las probabilidades posteriores de los parámetros de interés se estiman a la que vez que se integra la incertidumbre en el resto de los parámetros considerados menos importantes (*nuisance parameters*, Ronquist 2004; Ronquist & Deans 2010). Esto hace a la inferencia Bayesiana más robusta a la *sobreparametrización* (bajo ratio parámetros/datos) en relación con la inferencia MV. Por otro lado, algunas de las desventajas de IB están en la dificultad de especificar probabilidades *prior* apropiadas para cada parámetro (aunque por otro lado si estas son informativas, el uso de probabilidades previas también puede servir para reducir la incertidumbre en la estima). La dificultad para determinar si la Cadena de Markov (MCMC) ha corrido lo suficiente para alcanzar el "estado estacionario" donde las estimas de los parámetros son fiables, y el hecho de que el uso erróneo de modelos evolutivos demasiado simples puede inflar las probabilidades posteriores (Brown *et al.* 2010). Por otro lado, un modelo demasiado complejo, aunque mejora el ajuste de los datos, también puede aumentar la varianza en parámetros estimados. Los soportes de los clados en IB se conocen como *posterior probabilities*. Además, para medir el apoyo con que los datos soportan o rechazan una hipótesis se pueden utilizar los *Bayes Factors* –la proporción de las probabilidades posteriores respecto a las probabilidades previas para dos hipótesis de interés.

En esta tesis se han realizado reconstrucciones filogenéticas con datos moleculares. Se ha utilizado principalmente inferencia Bayesiana con los programas MrBayes (**Capítulo 1**) y BEAST (**Capítulos 1**, **2**, **3** y **5**). Como complemento se ha utilizado máxima verosimilitud con el programa RAxML (**Capítulo 1**). Los modelos de sustitución nucleotídica se han seleccionado con el programa MrModelTest (**Capítulos 2** y **4**). También se han escogido los *priors* con *Bayes factor* utilizando muestreos *path sampling* (PS) y *stepping-stone sampling* (SS) en BEAST (**Capítulo 1**).

2. El uso de marcadores moleculares

Como se mencionó anteriormente, la introducción de marcadores moleculares en el establecimiento de hipótesis evolutivas revolucionó el campo de la filogenia. El abaratamiento de los costes de secuenciación y la introducción de técnicas de secuenciación masiva (*Next Generation Sequencing*) ha supuesto también un incremento exponencial en el número y clase de caracteres moleculares y número de individuos analizados.

Existen numerosos tipos de marcadores moleculares (secuencias cortas, largas, fragmentos anónimos, genomas completos, transcriptomas, etc.) y su elección dependerá en gran medida de la pregunta a contestar y del nivel taxonómico al que se quiera investigar (inter- o intraespecífico). A continuación se comentan brevemente algunos de estos marcadores, con especial énfasis en los utilizados en la presente tesis.

2.1 Secuencias de ADN

Los primeros estudios filogeográficos utilizaron marcadores de herencia citoplasmática no mendeliana, como el ADN mitocondrial (ADNmt) en animales o el ADN cloroplástico (ADNp) en plantas. Ambos tipos de ADN se caracterizan por su naturaleza haploide, ser no recombinantes, y poseer una herencia exclusivamente materna (o paterna en las gimnospermas). El ADNmt presenta una alta tasa de mutación en animales, mientras que esta es baja en plantas; el ADNp tiene una mayor tasa evolutiva pero en general esta es inferior a la del ADNmt en animales (Wolfe *et al.* 1987). Esto resulta en un menor polimorfismo intraespecífico en plantas y menor información genética para examinar hipótesis filogeográficas. Otro inconveniente es que debido a la herencia materna en angiospermas, los patrones de ADNp solo reflejarían la historia de la colonización de las semillas. Por tanto, para reconstruir adecuadamente la historia evolutiva, se hace necesario el uso de otras fuentes de evidencia, como la proporcionada por el ADN nuclear. Este ADN se ha utilizado tradicionalmente menos debido a su naturaleza recombinante y por ser diploide o poliploide, o presentarse en múltiples copias (p. ej. ITS), de forma que es posible tener para cada gen historias evolutivas independientes que dificultan aún más la reconstrucción de la filogenia de las especies o la secuencia de eventos de especiación. No obstante, el uso de marcadores de ADN nuclear de copia única o de baja copia (*low-copy nuclear genes*) proporciona información interesante ya que posee una tasa de mutación más rápida que el ADNpt y añade información sobre la historia del flujo polínico (paternal). La aproximación apropiada, por tanto, es el uso integrado de marcadores del genoma plastidial y nuclear para poder reconstruir la historia evolutiva completa de la especie usando varias genealogías, en vez de un único árbol génico. Una dificultad para ello es la existencia de incongruencias entre los distintos loci, debida a fenómenos poblacionales o de coalescencia como la paralogía, evolución concertada, separación incompleta de linajes (ILS: *incomplete lineage sorting*), homoplasia, pero también a eventos de evolución reticulada como la hibridación o la introgresión.

En los **Capítulos 1** y **5** se utilizan marcadores de ADN cloroplástico y nuclear para inferir filogenias o genealogías a nivel de población; en el **Capítulo 1** se discute el efecto de fenómenos como la saturación o la homoplasia en la reconstrucción de la historia evolutiva de las especies.

2.2. Polimorfismos de ADN

Los polimorfismos de ADN, también conocidos como técnicas de huella génica o *fingerprinting*, son apropiados para el estudio de la diversidad intraespecífica. Existe un amplio abanico de marcadores polimórficos entre los que se pueden nombrar algunos como: los RAPD (*Random Amplified Polymorphic* DNA); RFLP (*Restriction Fragment Length Polymorphism*) y SNP (*Single Nucleotide Polymorphisms*). Cada marcador posee sus ventajas e inconvenientes. Algunas de las propiedades ideales para estos marcadores son: neutralismo, alto polimorfismo, herencia codominante, y alta reproducibilidad. Mientras que algunas de las desventajas habituales son: precio, alto nivel de homoplasia y baja reproducibilidad. A continuación se profundiza en los dos tipos de marcadores utilizados en esta tesis: AFLPs y microsatélites.

Los AFLPs (*Amplified Fragment Length Polymorphism*) son un tipo de marcador que permiten identificar múltiples loci polimórficos dentro de un genoma nuclear. Los microsatélites son repeticiones en tándem de 1-6 nucleótidos que se localizan en el genoma de la mayoría de organismos. Ambos marcadores se utilizan en estudios de biodiversidad (p. ej. para medir la diversidad y estructura genética poblacional) pudiendo utilizarse para abordar diferentes cuestiones ecológicas y evolutivas. En la tabla 3 se resumen las características, ventajas, e inconvenientes de estos marcadores. La principal diferencia es que los AFLPs son marcadores dominantes, mientras que los microsatélites son codominantes. A pesar de algunas desventajas frente a los SSRs o las secuencias de ADN, los AFLPs son una herramienta todavía muy utilizada en filogeografía por su fácil uso y coste relativamente bajo (ver Tabla 3). Los loci individuales de los AFLPs son menos informativos que los loci en microsatélites; sin embargo, el gran número de loci AFLP disponibles tiende a reducir la homoplasia y su distribución aleatoria a lo largo del genoma permite equilibrar este sesgo (Bonin *et al.* 2007). Se ha defendido que los AFLPs son más apropiados en estudios intraespecíficos abarcando amplias escalas geográficas, mientras que con microsatélites la estructura se borra debido a que la alta tasa de mutación genera homoplasia (Skrede *et al*. 2008). La comparación de ambos marcadores produce generalmente resultados congruentes cuando se analiza un amplio número de loci (Meudt & Clarke 2007). Así, los AFLPs pueden ser una poderosa herramienta, alternativa o complementaria a los microsatélites (Bensch & Akesson 2005), y la elección de uno u otro marcador debería hacerse de acuerdo con la pregunta que desea abordarse en el estudio.

En los **Capítulos 2** y **3** de la presente tesis se utilizan secuencias de ADN cloroplástico y marcadores nucleares AFLPs para reconstruir la estructura geográfica de la variación genética en dos especies de *Canarina*, mientras que en el **Capítulo 4** se emplean marcadores nucleares microsatélites para estudiar la historia demográfica reciente.

Tabla 3. Principales características y comparativas de marcadores AFLPs y microsatélites.

	AFLPs	**Microsatélites**
Dominio	Marcadores dominantes y anónimos. Implica que es difícil calcular las frecuencias alélicas, ya que observando la presencia de los picos no es posible distinguir entre una condición homocigota o heterocigota	Marcadores codominantes. Cuando la mayoría de individuos son homocigotos la ventaja de usar marcadores codominantes se reduce
Utilidad	i) estudios intraespecíficos, ii) introgresión e hibridación, iii) en poliploides, iv) cuando la variabilidad genética es baja	i) flujo génico entre poblaciones, ii) endogamia, iii) cuellos de botella, iv) hibridación intraespecífica e introgresión, v) asignación de individuos a poblaciones
Costes	No se necesita información genética previa por lo que el coste inicial en tiempo y dinero generalmente es menor	Si no se posee información previa, se necesita identificar las regiones flanqueantes y el desarrollo es más largo y costoso
Calidad de la muestra	Son sensibles a la calidad del ADN de partida. Requieren material bien conservado.	Se puede obtener ADN de calidad con muestras de herbario
Homología/ Homoplasia	Los fragmentos no son necesariamente homólogos, lo que disminuye su fiabilidad. Cuanto menor es el tamaño del fragmento, mayor es la probabilidad de homoplasia.	Los fragmentos se detectan con fiabilidad, sin embargo puede haber homoplasia debido a mutaciones convergentes. Se asume que no existe homoplasia al utilizar el modelo mutacional IAM.
Lectura	En algunas ocasiones las bandas pueden ser difíciles de interpretar y pueden ser homoplásicas	Lectura de las bandas claras y fáciles de interpretar
Muestro del genoma y variabilidad	Permiten monitorizar un amplio número de loci. Cubren amplias áreas del genoma, aumentando así la probabilidad de encontrar diferenciación en grupos con baja variabilidad.	Dificultad para monitorizar un amplio número de loci. Están Distribuidos principalmente en la eucromatina, mientras que en la heterocromatina son poco abundantes. Poseen una baja frecuencia de repetición en las regiones codificantes y en los telómeros. Son marcadores polimórficos, con una alta tasa de mutación (media de $5*10^{-4}$; Selkoe & Toonen 2006), producida por errores de deslizamiento en la replicación del ADN que cambia el número de repeticiones.
Mecanismo y modelo mutacional	En AFLPs no se ha podido aplicar un modelo mutacional apropiado y los métodos de estadística filogeográfica basados en simulaciones coalescentes no se han podido utilizar (sin embargo, ver Luo *et al*. 2007).	Se cree que el mecanismo mutacional predominante es el deslizamiento de ADN (DNA *slippage*). Sin embargo esto es objeto de debate y existen varios modelos mutacionales: *Infinite Allele Model* (IAM, Kimura & Crow 1964); *K-allele Model* (KAM, Crow & Kimura 1970); *Stepwise Mutation Model* (SMM, Kimura & Ohta 1978); *Two Phase Model* (TPM, Di Rienzo *et al*. 1994) y *Generalized Stepwise Model* (GSM). El debate se sitúa en torno a si predominan los SMM u otros modelos que incluyan mutaciones puntuales o eventos de recombinación, que los están reemplazando.

3. El reloj molecular y el registro fósil

En los estudios de biología evolutiva se utiliza el reloj molecular para convertir medidas de distancia genética en estimas de tiempos de divergencia. La hipótesis del reloj molecular de Zuckerkandl & Pauling asume que para una macromolécula la tasa de evolución es constante en el tiempo para todos los linajes (1962, 1965). Al incrementar el tiempo de divergencia "t", se incrementa la variabilidad interespecífica "k", siendo k/t una constante que tiene una relación lineal. De esta manera puede calcularse la edad de divergencia utilizando relojes moleculares estrictos. Sin embargo, en las últimas décadas se ha observado que el estudio de sets de datos moleculares no se ajusta a esta relación lineal, por lo que para realizar dataciones debe tenerse en cuenta la heterogeneidad de tasas de evolución entre linajes (Sanderson 1997; Thorne *et al.* 1998). Uno de los mayores avances ha siso la introducción del reloj molecular "relajado", que tiene en cuenta variaciones de tasa entre linajes para estimar los tiempos de divergencia con mayor precisión (Sanderson 1997; Kishino *et al.* 2001). Los relojes relajados permiten calcular las tasas de sustitución en las ramas de un árbol y los tiempos de divergencia utilizando inferencia estadística mediante métodos de máxima verosimilitud (PL: *Penalized Likelihood* (Sanderson 2002)) o inferencia Bayesiana (p. ej. Multidivtime (Thorne & Kishino 2002) y BEAST (Drummond & Rambaut 2007)).

El reloj molecular proporciona estimas relativas del tiempo de divergencia entre linajes. Para obtener estimas absolutas es necesario disponer de algún punto de calibración (p.ej. registro fósil o eventos geológicos) con el que asignar una edad a un nodo en el árbol filogenético (Ho & Philips 2009). En las calibraciones geológicas o biogeográficas, eventos como la apertura de una cuenca marina o la formación de una cadena montañosa, se utilizan para establecer límites máximos de edad. Sin embargo, esto puede acarrear varios problemas, como por ejemplo, que la asunción de vicarianza no tiene en cuenta una posible extinción de linajes (Waters & Craw 2006) y que no puede asegurarse que la edad de divergencia sea anterior o posterior al origen de una barrera (Heads 2005). Por otro lado, el registro fósil puede indicarnos la edad en la que un linaje logró distinción morfológica y abundancia suficiente hasta llegar a la fosilización. Sin embargo, el registro fósil no nos indica el tiempo de divergencia filogenética, si no que sirve para indicar el límite de edad mínimo en un nodo del árbol filogenético (Heads 2005). El problema derivado de esto es que pueden sobreestimarse los tiempos de divergencia (Donoghue & Benton 2007). La diferencia temporal entre la preservación del fósil y la divergencia de los linajes es desconocida y varía de un linaje a otro (Figura 23a.), dependiendo de los atributos físicos y las preferencias ambientales de los miembros de cada linaje. Así, en organismos con altas tasas de evolución morfológica y partes duras en su cuerpo (p.ej. exoesqueleto) que habitan ambientes que facilitan su conservación, se espera un tiempo pequeño entre la divergencia filogenética y la conservación del fósil (Magallón 2014). Por otro lado, los fósiles dependen de dataciones realizadas mediante interpretación estratigráfica o datación radiométrica, técnicas que también están sujetas a cierta incertidumbre y podrían acumular un sesgo en la calibración del reloj molecular. Para solucionar esto, Yang & Rannala (2006) propusieron utilizar distribuciones estadísticas para integrar la incertidumbre en la datación de los fósiles, mejorando así las estimas de divergencia. Hasta entonces se habían utilizado límites estrictos (*hard bounds*), pero se añadió una cola de probabilidad más allá del límite permitiendo el uso de límites relajados (*soft bounds*) en métodos de datación Bayesianos (Yang & Rannala 2006). Asimismo, dependiendo del tipo de calibración utilizada, se han propuesto diferentes distribuciones probabilísticas (*priors*) para incorporar en las dataciones la incertidumbre asociada a la edad del fósil o al evento geológico usado en la calibración (Ho & Phillips 2009; Figura 24).

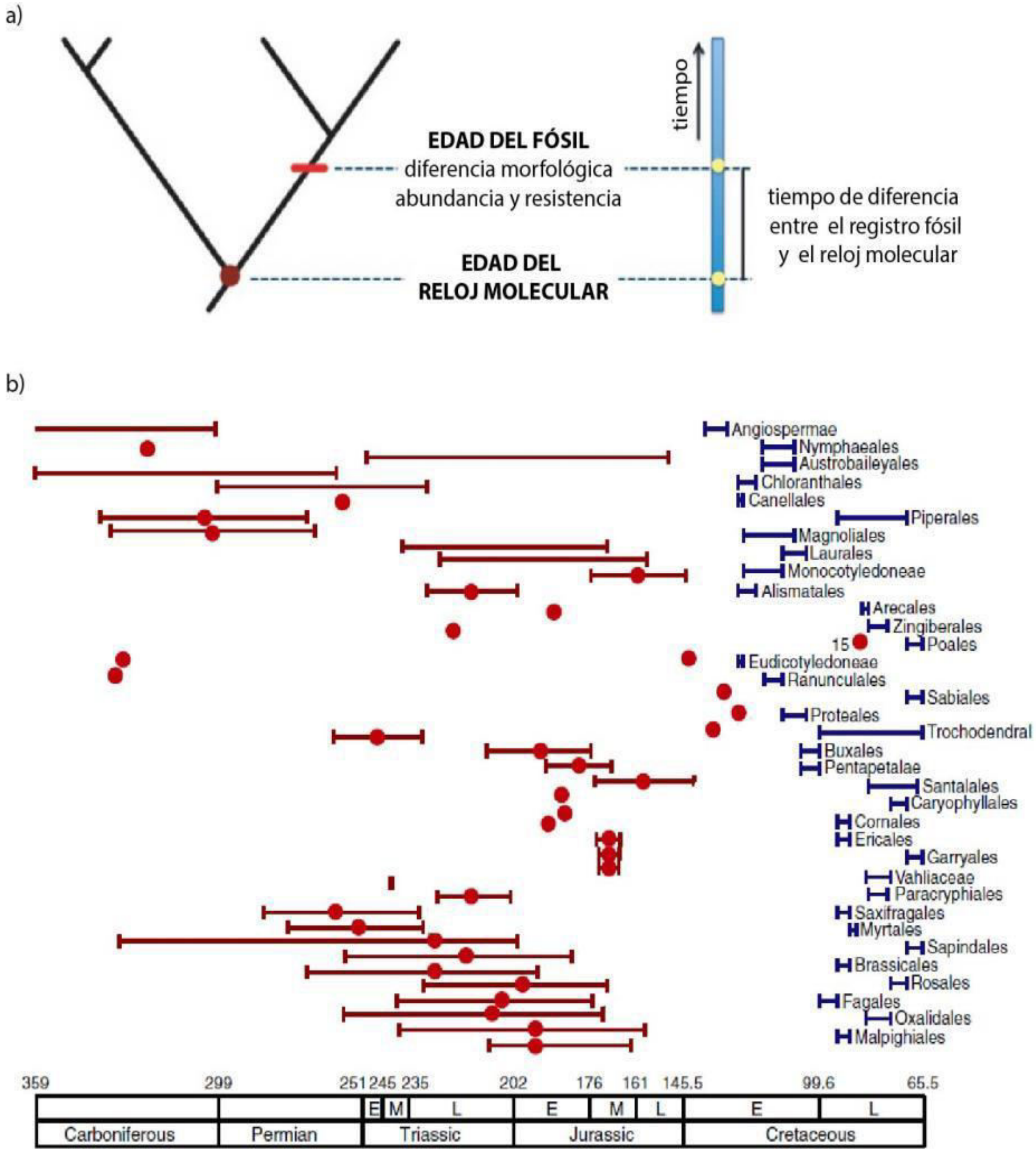


Figura 23. a) Estimas de edad del reloj molecular y del registro fósil (modificado de Magallón 2014). Los relojes moleculares permiten estimar el tiempo de divergencia entre dos linajes en un árbol filogenético. El fósil de mayor edad de un linaje indica el tiempo en el que el linaje adquirió una morfología diferente y fue suficientemente abundante para entrar en el registro fósil. b) Estimas fósiles (en azul) y moleculares (en rojo) de la edad de diferentes linajes de angiospermas (modificado de Magallón *et al*. 2015).El nombre del orden de cada clado se muestra a la derecha. Mientras que el registro fósil es consistente con un comienzo de las angiospermas en el Cretácico Temprano, las estimas moleculares proporcionan una edad más antigua del origen de las angiospermas.

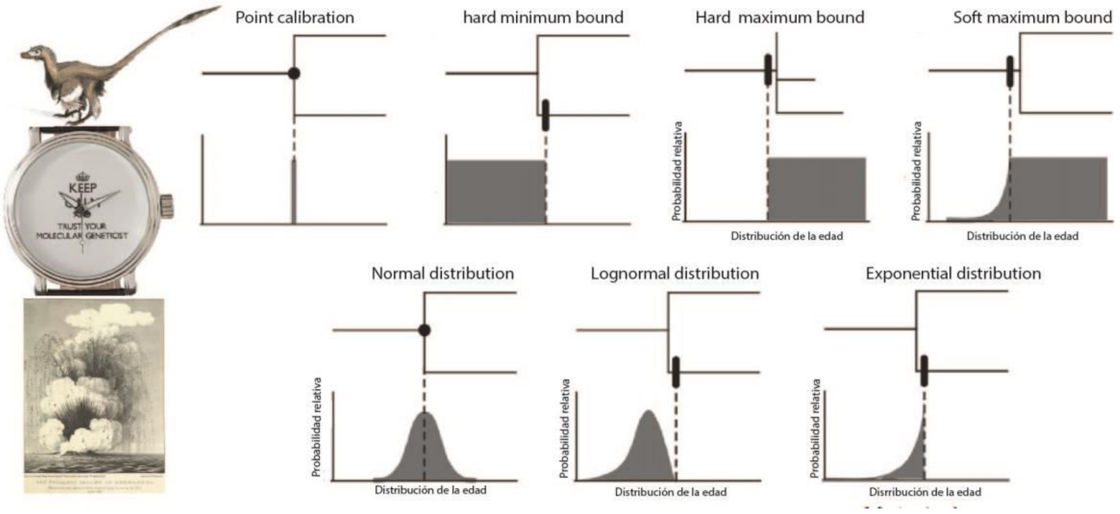


Figura 24. Métodos para incorporar las calibraciones en las dataciones (modificado de Ho & Phillips 2009).

En angiospermas, el registro fósil ofrece una visión general de la diversificación evolutiva; sin embargo, para linajes particulares, el registro fósil está mayormente incompleto (Magallón 2014). Lo más común por tanto es utilizar calibraciones secundarias basadas en el registro fósil; en este tipo de calibraciones, la falta de fósiles se suple utilizando estimas de edades obtenidas de la datación de un linaje de mayor rango taxonómico que incluye el grupo de estudio; a su vez, este linaje se databa utilizando el mayor número posible de linajes y de fósiles. Las estimas iniciales se basaron en un principio en métodos que asumían autocorrelación en las tasas de evolución molecular entre linajes y que usaban puntos de calibración fijos (Wikström *et al.* 2001). Posteriormente, se han utilizado relojes moleculares relajados que no asumen correlación entre tasas (Drummond *et al.* 2006) y tienen en cuenta la heterogeneidad de tasas entre linajes. Por ejemplo, Charles Bell *et al.* (2010) utilizaron un reloj Bayesiano relajado para obtener estimas de divergencia en angiospermas (45 órdenes, 355 familias y 657 taxones) utilizando numerosos puntos de calibración fósil (36). Más recientemente, Magallón *et al.* (2015) han estimado un nuevo marco temporal de la evolución de angiospermas (con representación del 87% de familias de angiospermas y 137 puntos de calibración) que promete ser el nuevo punto de referencia para realizar calibraciones.

En los **Capítulos 1**, **2**, **3** y **5** utilizamos el programa BEAST para implementar relojes moleculares relajados y estimar tasas de cambio y tiempos de divergencia mediante Inferencia Bayesiana; adicionalmente, en el **Capítulo 2** utilizamos un método de máxima verosimilitud, *Penalized Likelihood*. En los **Capítulos 1** y **2**, **3** y **5** el trabajo de Bell *et al.* es utilizado para proporcionar puntos de calibración en nuestras filogenias.

4. En busca de la inferencia filogeográfica…

4.1 La teoría de coalescencia

La aparición de la teoría de la coalescencia interpretada bajo el prisma de la genética de poblaciones clásica ha sido fundamental a la hora de examinar hipótesis filogeográficas dentro de un marco estadístico e inferencial. La teoría de coalescencia se basa en las formulaciones matemáticas descritas por John Kingman en 1982 y se considera una extensión de la teoría clásica de la genética de poblaciones. Esta teoría establece que: en una población de tamaño constante y a lo largo de varias generaciones, alelos nuevos aparecen por mutación y otros se pierden por deriva génica, de manera que todos los alelos proceden de un único alelo ancestral. De este modo, partiendo de los haplotipos presentes podemos rastrear hacia atrás la historia de estos linajes para localizar donde se unen o coalescen en un ancestro común. Aquí cobra relevancia la relación entre el tiempo de coalescencia y el tamaño poblacional ya que, cuanto más pequeño sea el tamaño poblacional, más rápido coalescen las secuencias. Además, el tiempo de coalescencia también puede verse afectado por otros factores como la inmigración o la selección natural, lo que hace necesario considerar otros parámetros. La teoría de coalescencia acepta la teoría de la neutralidad de Kimura y asume que las mutaciones neutras no afectan al éxito reproductivo de los individuos, por lo que tampoco se verá afectada la topología en las genealogías; asimismo, asume que el conflicto entre árboles de genes y de especies sólo se debe a la separación incompleta de linajes, por lo que no tiene en cuenta otros procesos como el flujo génico o la hibridación. Para reconstruir el árbol de especies se necesita información de múltiples árboles de genes que estarán incluidos dentro de la historia del árbol de especies mientras no violen la siguiente condición: el tiempo del ancestro común de un gen no puede ser más reciente que el tiempo de divergencia de las respectivas especies (Figura 25).

En definitiva, la teoría de coalescencia proporciona un riguroso marco matemático-estadístico que permite discriminar entre hipótesis alternativas para así interpretar conjuntamente los árboles de genes y la demografía poblacional –la coalescencia ha permitido unir los conceptos de genética de poblaciones y filogenia a través del estudio del árbol de genes (Nielsen & Beaumont 2009).

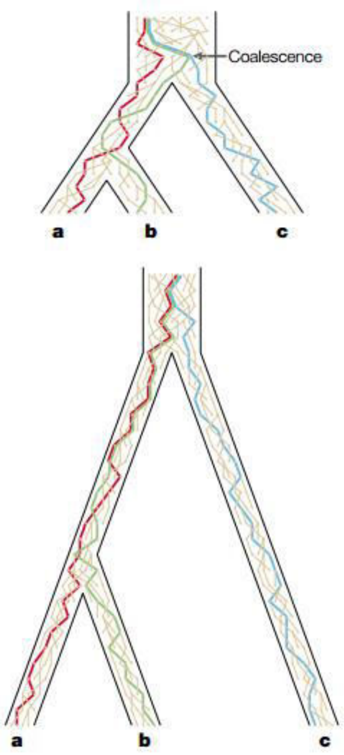


Figura 25. Conflicto entre árbol de genes y árbol de especies debido a reparto incompleto de copias génicas entre linajes, *incomplete lineage sorting* (tomado de Rosenberg & Nordborg 2002). Se muestran dos árboles de especies y embebido en cada uno de ellos se muestra el árbol de genes. En la imagen de arriba, el árbol de genes no refleja el árbol de especies; un gen muestreado para la *especie b* (verde) y otro muestreado para la *especie c* (azul) están más relacionados que el gen muestreado para la *especie a* (rojo), aunque *a* y *b* sean las últimas especies en divergir. En la imagen inferior, sin embargo, todos los linajes génicos coalescen siguiendo el orden de los eventos de especiación en el árbol de especies, y por tanto apoyan la misma topología de éste. En el árbol de genes de la imagen superior, el gen *c* muestra una coalescencia profunda (*deep coalescence*) que contradice el árbol de especies; este tipo de discordancia entre el árbol de genes y el árbol de especies es más común cuanto mayor sea el tamaño poblacional de la población ancestral (reperesentado por el grosor de las ramas) y más reciente la divergencia entre las especies (ramas cortas).

4.2 Estimadores de la diversidad genética

La filogeografía utiliza los estadísticos descriptivos para determinar el grado de variación genética poblacional. A continuación se comentan algunos de los estadísticos utilizados en esta tesis y su interpretación para la búsqueda de refugios, ya que reaccionan de forma diferente a procesos como los cuellos de botella genéticos (Widmer & Lexer 2001). i) La riqueza alélica: es el número de alelos por locus. Es muy dependiente del tamaño poblacional, por lo tanto es más útil para identificar procesos históricos como cuellos de botella y mezcla poblacional. ii) La diversidad genética de Nei, H (heterocigosidad esperada): es la probabilidad de que dos alelos muestreados aleatoriamente en una población sean diferentes. En los cuellos de botella, la reducción de H es limitada en comparación con la riqueza alélica. iii) La proporción de loci polimórficos: Es el número de loci polimórficos dividido por el número total de loci. Es insensible a los cuellos de botella, porque los alelos en común pueden persistir incluso en cuellos de botella severos.

Para estudiar la distribución de la variabilidad genética en el espacio hemos utilizados los estadísticos F, desarrollados originalmente por Wright (1931) y posteriormente extendidos a múltiples loci por Nei (1973). Además hemos utilizado inferencia Bayesiana para detectar la estructura genética basada en modelos de *ancestría mixta* (Pritchard *et al.* 2000) implementados en el programa STRUCTURE. En los **Capítulos 2-4** se puede encontrar más información sobre los estadísticos descriptivos.

4.3 Aproximaciones de vanguardia en filogeografía

En 2010, Bloomquist *et al.* en su artículo "*Three roads diverged: routes to phylogeographic inference*", revisaron los que consideraron como los tres principales enfoques metodológicos en filogeografía: i) la filogeografía comparativa (*comparative* *phylogeography*), ii) la filogeografía basada en modelos Bayesianos de difusión espacial (*Bayesian spatial difussion models*), y iii) la filogeografía basada en procesos de genética de poblaciones (*population-based coalescent models*). Estos enfoques tienen en común que pretenden inferir los procesos filogeográficos que han determinado la estructura espacial de la diversidad génica a partir de datos moleculares con localización geográfica conocida (Bloomquist *et al.* 2010). Una de las primeras aproximaciones que se desarrollaron en filogeografía fue la comparativa, que consistía en la construcción de árboles de haplotipos con un único locus (Templeton *et al*. 1992, Templeton & Sing 1993); la más popular es *Nested Clade Phylogeographic Analysis* (NCPA), una aproximación basada en remuestreo, utilizada para inferencia y testeo de hipótesis filogeográficas (Templeton *et al*. 1992; Templeton 1998). Una década después se demostró que realizar estas inferencias con un único locus tenía una alta tasa de falsos positivos (Knowles & Maddison 2002), lo que se intentó solucionar introduciendo una aproximación multi-locus (Templeton 2004). Esta aproximación, vigente a día de hoy, es objeto de un acalorado debate entre los partidarios de la aproximación comparativa y los partidarios de una filogeografía estadística basada en modelos. La filogeografía estadística surge como una síntesis de la genética de poblaciones y la filogenética. Se basa en un modelo probabilístico explícito basado en procesos (parámetros) que describen cómo se generan los datos (Knowles & Maddison 2002) y utiliza métodos de inferencia estadística (Máxima Verosimilitud, Inferencia Bayesiana) para estimar los parámetros y testar la fiabilidad del modelo. Esta sería la principal diferencia respecto a los métodos comparativos como NCPA, ya que estos no incorporan el error o la incertidumbre en la estimación, lo que puede aumentar el Error Tipo I; rechazar erróneamente la hipótesis nula. En general las aproximaciones basadas en modelos de procesos y en genética de poblaciones están ganando peso, ya que pueden incorporar evidencia externa adicional en la inferencia como la distancia geográfica, el área de distribución, o la presencia de barreras que limitan el flujo génico. Estos métodos basados en modelos se han ido imponiendo sobre los métodos comparativos en el análisis de hipótesis filogeográficas (Beaumont *et al*. 2010); sin embargo, métodos como NCPA son todavía muy usados para explorar la estructura geográfica de la variación genética y han sido recientemente refinados con la incorporación de aproximaciones probabilísticas que permiten incorporar la incertidumbre en la estimación (Manolopoulou *et al.* 2011; Manolopoulou & Emerson 2012: ver más abajo).

En la presente tesis se utilizan los tres enfoques o aproximaciones de inferencia filogeográfica, los cuales se describen brevemente a continuación.

i) Filogeografía Comparativa (comparative approach)

La filogeografía comparativa asume que la geografía es un carácter fenotípico y lo relaciona con la historia evolutiva, reconstruyendo así las relaciones evolutivas entre secuencias infraespecíficas o entre haplotipos (Bloomquist *et al*. 2010). Entre los métodos más populares se encuentran los basados en Parsimonia Estadística (Templeton *et al.* 1992) como las redes estimadas con el programa TCS (Clement *et al*. 2000) y los ya comentados NCPA (Templeton 1998, 2004).

Estos métodos, a diferencia de los clásicos árboles bifurcados, permiten representar las conexiones entre los haplotipos en forma de red. Así se tienen en cuenta procesos propios de niveles intrapoblacionales (p. ej. recombinación) y se pueden obtener estimas de las relaciones genealógicas entre secuencias a nivel poblacional. TCS permite reconstruir árboles o redes de haplotipos utilizando la parsimonia estadística de Templeton *et al*. 1992. Las conexiones entre haplotipos en el árbol representan eventos coalescentes y en su interpretación se utilizan las siguientes predicciones o asunciones: i) los haplotipos más antiguos (los de mayor frecuencia) tienen mayor probabilidad de ser interiores, y los más jóvenes, exteriores ii) los alelos más antiguos, tendrán una amplia distribución geográfica, iii) los haplotipos con mayor frecuencia tenderán a tener más conexiones, iv) los haplotipos con una única conexión es más probable que estén conectados a haplotipos con múltiples conexiones, v) los haplotipos con una única conexión es más probable que estén conectados a haplotipos de la misma población (Posada & Krandall 2001). En los **Capítulos 5** y **6**, se utiliza TCS para reconstruir árboles de haplotipos.

Como alternativa a TCS hemos utilizado el método BPEC (*Bayesian Phylogeographic and Ecological Clustering*; Manolopoulou *et al.* 2011; Manolopoulou & Emerson 2012). Este método introduce el marco Bayesiano en las aproximaciones comparativas y corrige algunos de los problemas estadísticos atribuidos a métodos como TCS o NCPA. BPEC permite estimar las probabilidades posteriores de las redes de haplotipos utilizando un modelo *markoviano* de migración-mutación similar al desarrollado por Sanmartín *et al.* 2008 para filogenias de especies, pero basado en coalescencia. BPEC permite estimar la probabilidad posterior de cada uno de los árboles de coalescencia mediante Inferencia Bayesiana MCMC y así integrar la incertidumbre en la inferencia filogeográfica. Este marco estadístico probabilístico también permite a BPEC examinar explicaciones alternativas sobre la historia evolutiva de las poblaciones. Finalmente, a diferencia de la parsimonia estadística, se pueden inferir eventos de migración no observados pero que están predichos por el modelo, a los que asigna una probabilidad posterior; el método, sin embargo, no es reversible, por lo tanto no es útil para detectar eventos de migración en ambas direcciones como en Sanmartín *et al.* 2008. Este método se utiliza en el **Capítulo 3** y se compara con los árboles obtenidos con TCS.

ii) Filogeografía basada en modelos Bayesianos de difusión espacial (Bayesian spatial diffusion approach)

Ofrece información sobre la historia ancestral de nuestros datos infiriendo los procesos geográficos discretos que han configurado la estructura genética actual. Para ello la superficie geográfica se divide en áreas discretas y el movimiento entre áreas (*diffusion process*) se modela con un proceso estocástico o probabilístico denominado Cadena de Markov (*Markov-chain*), con estados discretos de manera homogénea en el tiempo. Este tipo de modelo probabilístico es semejante al que se utiliza para modelar la evolución de un carácter en una filogenia o a los modelos de substitución nucleotídica empleados en sistemática molecular. Su desarrollo constituye una de las fronteras más prometedoras en el campo de la biogeografía. Fue primero propuesto para biogeografía histórica por Sanmartín *et al.* en 2008 (*Bayesian Island Biogeographic* (BIB) *approach*, Sanmartín 2012), y de forma independiente por Lemey *et al.* (2009) para inferencia filogeográfica. El modelo BIB consiste en utilizar una cadena de Markov homogénea en el tiempo (*Continuous-time Markov Chain model*; CTMC) para modelar la evolución biogeográfica, donde los estados discretos corresponden a las localizaciones geográficas de las secuencias y las tasas de transición entre estados a las tasas de migración o dispersión entre áreas (Sanmartín *et al.* 2008). Mediante aproximaciones MCMC de Inferencia Bayesiana, permite estimar los estados ancestrales en cada nodo de la filogenia (la localización geográfica de los ancestros) y las tasas de migración entre áreas a partir de secuencias de ADN y su localización geográfica. Una de sus ventajas es su flexibilidad a la hora de incorporar evidencia externa como la distancia geográfica o el área (Ronquist & Sanmartín, 2011). El enfoque de Lemey *et al.* (2009) permite estimar también los eventos de migración en el espacio y en el tiempo, este modelo se utiliza en el **Capítulo 1**.

iii) Filogeografía basada en procesos de genética de poblaciones (population-based genetic approach)

También denominados métodos de filogeografía estadística. Permiten realizar inferencias utilizando el marco de la teoría de coalescencia, para revelar procesos que han modelado la historia evolutiva a nivel poblacional (Figura 26). A diferencia de la filogeografía comparada, la inferencia se realiza en un marco estadístico bajo un modelo explícito de evolución filogeográfica o génica, que incorpore procesos poblacionales como selección, migración, cambios en el tamaño poblacional o recombinación, con otros que afectan a la historia del árbol de genes (*gene tree*) y que pueden causar discordancia con el árbol de especies (por ejemplo, un modelo que incorpore *incomplete lineage sorting* o eventos de duplicación y extinción de linajes génicos (*gene duplication and gene*

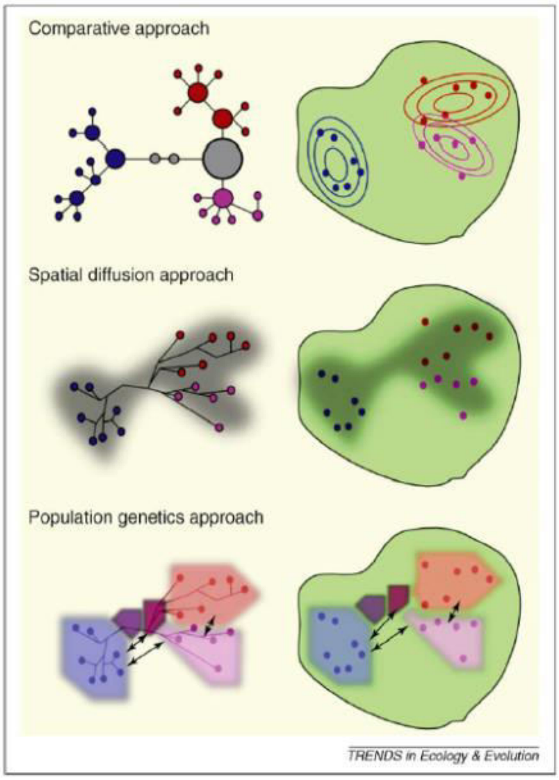


Figura 26. Las tres aproximaciones filogeográficas según Bloomquist *et al*. 2010. Los datos representan tres especies (rojo, púrpura y azul) muestreadas en una isla (verde). En la aproximación comparativa, a la izquierda se muestra el árbol de haplotipos, mientras que los círculos concéntricos en la isla representan la distribución geográfica de las tres especies. En la aproximación de difusión espacial, a la izquierda se muestra el árbol filogenético superpuesto sobre la localización geográfica de las muestras. El árbol filogenético se muestra en negro para reforzar que no se infiere la información poblacional. En la isla, las áreas grises detrás del árbol representan una alta probabilidad para la localización de los ancestros de los taxones empleados. En la aproximación de genética de poblaciones, los cinco polígonos coloreados representan las poblaciones ancestrales y sus tamaños. Las flechas representan las migraciones entre poblaciones.

loss) o eventos de transferencia horizontal. Uno de los modelos más utilizados es el llamado "modelo coalescente de múltiples especies" (*multispecies coalescent model*, Rannala & Yang 2003) e implementado en *BEAST (Heled & Drummond 2010). Este modelo permite reconstruir el árbol de especies a partir de múltiples genes o loci en especies estrechamente relacionadas, incorporando fenómenos de coalescencia relacionados con la separación incompleta de linajes (ILS: *incomplete lineage sorting*) para explicar la discrepancia entre árboles de genes y de especies (ver Figura 25). El modelo *multispecies coalescent*, sin embargo, no permite la transferencia horizontal, es decir, no incorpora la recombinación entre loci, y asume que entre las especies no se da flujo génico, transferencia horizontal e introgresión. Otros métodos a destacar son los que utilizan Computación Bayesiana Aproximada (*Approximate Bayesian Computing*: ABC) para simular por coalescencia historias demográficas complejas y comparar los datos observados con los obtenidos según el modelo. El uso de ABC se ha extendido para realizar inferencias de la historia demográfica y adaptación local (Beaumont 2010; Csilléry *et al*. 2010).

En el **Capítulo 1** utilizamos una alternativa a *BEAST, mientras que en el **Capítulo 4** utilizamos técnicas ABC implementadas en el programa DIYABC, para inferir historia demográfica reciente.

5. Los modelos de distribución de especies

Los modelos de distribución de especies (*Species distribution modeling*, SDM) predicen la distribución de las especies en el espacio y en el tiempo. A partir de los datos de presencia de un organismo se utilizan herramientas estadísticas y matemáticas para inferir la distribución en función de diferentes variables ambientales. Los SDM se pueden combinar con simulaciones paleoclimáticas para predecir la distribución de las especies a través de grandes regiones (Svenning *et al*. 2011). Además pueden utilizarse para hacer proyecciones futuras y predecir la respuesta de las especies al cambio climático (Pagel & Schurr 2012). El uso de SDM en escenarios donde se han producido grandes cambios climáticos, como en África, es especialmente interesante para detectar refugios de biodiversidad (Gavin *et al*. 2014). Sin embargo, los SDM todavía poseen algunas limitaciones y asunciones que limitan su uso, por ejemplo: asumen estabilidad de nicho ecológico, no acomodan los procesos a nivel poblacional y no son precisos para detectar microclimas generados por el efecto de la topografía.

En la presente tesis se combina el uso de modelos de distribución de especies junto a simulaciones climáticas para proyectar la distribución de las especies en diferentes intervalos desde el presente hasta el Mioceno Tardío. En el **Capítulo 1** se modela la distribución del género *Canarina*, mientras que en el **Capítulo 6** se modela la distribución de varios linajes Rand Flora.

VI. Conclusions (in English)

1. Genus *Canarina* is monophyletic. *Canarina abyssinica* is sister to a clade formed by the disjunct species *C. canariensis* and *C. eminii*. The age for interspecific divergences within the genus (Late Miocene) stands in contrast with the much younger intraspecific divergences, which date back to the Pleistocene-Holocene.
2. Ancestors of *Canarina* migrated from Asia (where its sister group, the monotypic genus *Ostrowskia*, is distributed) to East Africa in the Middle Miocene, probably through the Arabian Peninsula. The first divergence isolated *C. abyssinica* from the ancestor of *C. eminii* and *C. canariensis* and could be explained by habitat specialization driven by Late Miocene climate aridification. This was followed by an expansion across northern Africa, probably before the formation of the Sahara Desert at the end of the Miocene, and a subsequent colonization of the Canary Islands in the Pleistocene.
3. The current disjunct distribution of *Canarina* might be explained by a process of climatic vicariance. Ancestors of *Canarina* achieved in the past a widespread distribution across North Africa, which was fragmented by climatic events, associated with the Neogene gradual aridification of the continent. Subsequent extinction of the intermediate populations led to the current distribution, with remaining populations surviving in climatically stable "island" refugees at both sides of the African continent: Macaronesia and the Eastern African mountains.
4. We propose a “nested-dating” method to infer divergence times for phylogeographic analysis in those groups which lack a phylogenetically close fossil record. In this method, a higher-level dataset calibrated with external evidence is used to constrain the molecular clock rate of additional linked datasets containing population-level data. This allows us to apply different tree growth evolutionary models (birthdeath vs.

coalescent) to different taxonomic levels (from above-species to populations), thus linking the macroevolutionary and microevolutionary approach.

5. *Canarina eminii* exhibits a strong phylogeographic structure, with reciprocally monophyletic lineages on each side of the Great Rift Valley, which seemingly acted as a major barrier to gene flow. Geological processes such as incipient rifting along the Rift Valley (volcanic-tectonic dynamics), and concomitant climatic changes, seem to have been determinant in the assembly of intraspecific diversity patterns in *Canarina* and other Afromontane species.
6. Populations of *Canarina eminii* are characterized by high levels of genetic exclusivity, restricted to different Afromontane forests patches, and critically endangered. The large and rapid climatic oscillations of the Pleistocene may be responsible for this strong geographic structure among populations, where the *sky islands* would have played a key role as refugia of genetic diversity.
7. The strong structure observed in the chloroplast genome of *Canarina eminii* suggests restricted seed migration among *sky islands*, while the nuclear structure indicates a stronger connection among populations east of the Great Rift Valley that among west populations. Our data cannot confirm the existence of vegetation bridges among the *sky islands* but suggests the existence of wider forest coverage in the Ethiopian highlands, as well as the possibility of short-range or stepping-stone dispersal among the various mountain ranges.
8. *Canarina canariensis* shows a strong geographic genetic structure, with Tenerife as its central axis, which divides populations into an eastern clade, grouping the paleo-island of Anaga with Gran Canaria, and a western clade, including the paleo-island of Teno and the islands of La Gomera, La Palma and El Hierro.

9. For a multiple island endemic species such as *Canarina canariensis*, oceanic barriers appear to have played a lesser role in structuring levels of genetic variation than the topographic barriers within islands. In the case of Tenerife this is explained by catastrophic events (volcanic eruptions and landslides) after the merging of the paleo-islands 3.5 million years ago.
10. Genetic diversity was greatest in the geologically stable "paleo-islands" of Tenerife: Anaga, Teno, and Roque del Conde. These areas have apparently preserved genotypes or allelic varieties that went extinct in other geologically less stable areas, and therefore can be considered as refugia of genetic diversity. In addition, the topographic complexity of the paleo-islands might have favoured its role as cradles of genetic diversity, acting as sources of new allelic combinations to other areas of Tenerife or adjacent islands. This confers these areas the characteristics of "phylogeographical hotspots", regions whose conservation should be prioritised.
11. On both oceanic islands and *sky islands*, geological dynamics and topography seem to be important drivers in the generation and structuring of genetic diversity, supporting the axiom that geographic barriers and biota evolve together.
12. The dramatic regression of the last remnants of subtropical ecosystems around Africa (Afromontane and laurel forest) has led to population decline in *C. canariensis* and *C. eminii*. This confirms the need for conservation actions on the habitats of these relict species.
13. The Rand Flora pattern is a clear example of biogeographic pseudocongruence: the same biogeographic pattern would have originated at different geological times from the Middle Miocene to the Late Pliocene coinciding with increasing aridification in Africa.

14. The existence of "climate corridors" in the Miocene (and Pliocene) might have allowed Rand Flora lineages to expand or migrate through Africa following their ancestral climatic preferences. The disappearance of these climatic corridors relegated populations to its current distribution, relicts within refugia or climatic islands in the margins of the African continent. The age of vicariance in these lineages seems to depend on their climatic tolerances; those lineages with sub-humid or sub-tropical preferences diverged first, while more xeric lineages exhibit younger divergences.
15. Extinction mediated by climate change has been important in the assembly of regional and global biodiversity patterns.

VII. Conclusiones (en español)

1. El género *Canarina* es monofilético. *Canarina abyssinica* es la especie hermana del clado formado por las especies *C. canariensis* y *C. eminii*. La datación de las divergencias interespecíficas en el género (Mioceno Superior) contrastan con las divergencias intraespecíficas mucho más recientes, datadas en el Pleistoceno-Holoceno.
2. Los ancestros de *Canarina* migraron desde Asia (donde se distribuye su grupo hermano, el género monotípico *Ostrowskia*) hacia el este de África en el Mioceno Medio, probablemente a través de la Península Arábiga. La primera divergencia aisló a *C. abyssinica* del ancestro de *C. eminii* y *C. canariensis*, y podría explicarse por la especialización de hábitat a causa de la aridificación climática del Mioceno Tardío. Posteriormente *Canarina* se habría expandido a través del norte de África, probablemente antes de la formación del desierto del Sáhara a finales del Mioceno, para finalmente colonizar las Islas Canarias en el Pleistoceno.
3. La actual disyunción del género *Canarina* podría explicarse por un proceso de vicarianza climática. Los ancestros de *Canarina* alcanzaron en el pasado una amplia distribución a lo largo del Norte de África, la cual fue fragmentada por eventos climáticos asociados con la aridificación gradual del continente en el Neogeno. Esto habría producido la extinción de las poblaciones intermedias dando lugar a la actual distribución, en la que las poblaciones actuales sobreviven en refugios insulares climáticamente estables a ambos lados del continente africano: Macaronesia y región Afromontana del este de África.
4. Se propone un método de datación "anidado" para inferir tiempos de divergencia en análisis filogeográficos de grupos que carecen de un registro fósil cercano filogenéticamente. En este método, un *dataset* de nivel taxonómico superior calibrado

con evidencia externa, se utiliza para informar a la tasa de mutación del reloj molecular de otro *dataset* que contiene información a nivel poblacional. Esto además nos permite aplicar diferentes modelos evolutivos de construcción de árboles (*birthdeath* vs. coalescencia) a cada nivel taxonómico (desde el nivel interespecífico al poblacional), vinculando así las aproximaciones macroevolutivas y microevolutivas.

5. *Canarina eminii* posee una fuerte estructura filogeográfica, con linajes que poseen una *monofilia recíproca* a ambos lados del Gran Valle del Rift, que al parecer actuó como la principal barrera para el flujo génico. Los procesos geológicos como el agrietamiento incipiente a lo largo del Gran Valle del Rift (dinámicas volcánico-tectónicas) y el consiguiente cambio climático, parecen haber sido determinantes en el ensamblaje de los patrones de diversidad intraespecífica en *Canarina* y otras especies Afromontanas.
6. Las poblaciones de *Canarina eminii* se caracterizan por altos niveles de exclusividad genética, restringida a diferentes parches del bosque Afromontano, y críticamente amenazada. Las grandes y rápidas oscilaciones climáticas del Pleistoceno parecen ser las responsables de esta fuerte estructura geográfica entre poblaciones, donde las *sky islands* habrían jugado un papel determinante como refugios de diversidad genética.
7. La fuerte estructura observada en el genoma cloroplástico de *C. eminii* sugiere una migración de semillas restringida entre *sky islands*, mientras que la estructura nuclear muestra una mayor conexión entre las poblaciones al este del Gran Valle del Rift, que entre las poblaciones al oeste. Nuestros datos no permiten confirmar la existencia de puentes de vegetación entre *sky islands*, pero sugieren la existencia de áreas forestales más extensas en las mesetas etíopes, así como una dispersión a corta distancia o mediante *stepping-stones* entre diferentes sistemas montañosos.
8. *Canarina canariensis* muestra una fuerte estructura geográfica de la variación genética con Tenerife como eje central, que divide a las poblaciones en un clado al este, que

agrupa a la paleo-isla de Anaga y la isla de Gran Canaria, y otro clado al oeste, que agrupa la paleo-isla de Teno y las islas de La Gomera, La Palma y El Hierro).

9. En una especie con amplia distribución insular como *C. canariensis*, las barreras oceánicas parecen haber tenido un papel menor en la estructuración de los niveles de variación genética que las barreras topográficas dentro de cada isla. En el caso de Tenerife esto se explicaría por eventos catastróficos (erupciones volcánicas y mega-deslizamientos) posteriores a la formación de la isla a partir de la unión de las tres paleo-islas volcánicas iniciales.
10. Los mayores niveles de diversidad se localizaron en las tres paleo-islas de Tenerife geológicamente estables (Anaga, Teno y Roque del Conde). Estas áreas aparentemente han conservado genotipos o variedades alélicas que se han extinguido en otras áreas geológicamente menos estables, por lo que las podemos considerar refugios de diversidad genética. Además, la complejidad topográfica de las paleo-islas podría haber favorecido su papel como cunas de diversidad genética, que habrían actuado como fuentes de diversidad de nuevas combinaciones alélicas hacia otras áreas de Tenerife o de islas adyacentes. Esto otorga a estas áreas la característica de *hotspots filogeográficos*, regiones cuya conservación debería ser prioritaria.
11. Tanto en islas oceánicas como en *sky islands*, las dinámicas geológicas y la topografía parecen tener un importante papel en la generación y estructuración de la diversidad genética, lo que apoya el axioma de que las barreras geográficas y las biotas evolucionan juntas.
12. El dramático retroceso de los últimos retazos de ecosistemas subtropicales alrededor de África (bosque Afromontano y laurisilva) ha provocado un declive demográfico en *C. canariensis* y *C. eminii*. Esto confirma la necesidad de realizar acciones de conservación en el hábitat de estas especies relictas.

13. El patrón Rand Flora es un claro ejemplo de *pseudocongruencia biogeográfica*: el mismo patrón biogeográfico se habría originado en épocas geológicas distintas desde el Mioceno Medio hasta el Plioceno Tardío, lo que coincidiría con el incremento de la aridificación en África.
14. La existencia de *corredores climáticos* en el Mioceno y Plioceno podría haber permitido a los linajes Rand Flora expandirse o migrar a través de África siguiendo sus preferencias ecológicas. La desaparición de estos corredores climáticos habría relegado a las poblaciones a su distribución actual, relictos en refugios o islas climáticas en los márgenes del continente africano. La edad de la vicarianza en estos linajes parece depender de sus tolerancias climáticas: los linajes con preferencias sub-húmedas o sub-tropicales divergen primero, mientras que los linajes más xéricos presentan divergencias más jóvenes.
15. La extinción mediada por cambio climático habría sido determinante en el ensamblaje de patrones regionales y globales de biodiversidad.

VIII. Bibliography

Abebe B, Acocella V, Korme T, Ayalew D. 2007. Quaternary faulting and volcanism in the Main Ethiopian Rift. *Journal of African Earth Sciences* 48:115–124.

Ablay G, Martí J. 2000. Stratigraphy, structure, and volcanic evolution of the Pico Teide–Pico Viejo formation, Tenerife, Canary Islands. *Journal of Volcanology and Geothermal Research* 103: 175–208.

Acosta J, Uchupi E, Muñoz A, Herranz P, Palomo C, Ballesteros M, Group Z. 2005. Geologic evolution of the Canarian Islands of Lanzarote, Fuerteventura, Gran Canaria and La Gomera and comparison of landslides at these islands with those at Tenerife, La Palma and El Hierro. *Marine Geophysical Researches* 26: 77–82.

Aguinaldo AMA, Turbeville JM, Linford LS, Rivera MC, Garey JR, Raff RA, Lake JA. 1997. Evidence for a clade of nematodes, arthropods and other moulting animals. *Nature* 387: 489–493.

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

Allen M, Armstrong H. 2008. Arabia-Eurasia collision and the forcing of mid-Cenozoic global cooling. *Palaeogeography, Palaeoclimatology, Palaeoecology* 265: 52–58.

Ancochea E, Fuster J, Ibarrola E, Cendrero A, Coello J, Hernan F, Cantagrel JM, Jamond C. 1990. Volcanic evolution of the island of Tenerife (Canary Islands) in the light of new K- Ar data. *Journal of Volcanology and Geothermal Research* 44: 231–249.

Ancochea E, Huertas MJ, Cantagrel JM, Coello J, Fúster JM, Arnaud N, Ibarrola E. 1999. Evolution of the Cañadas edifice and its implications for the origin of the Cañadas Caldera (Tenerife, Canary Islands). *Journal of Volcanology and Geothermal Research* 88: 177–199.

Ancochea E, Hernán F, Huertas M, Brändle J, Herrera R. 2006. A new chronostratigraphical and evolutionary model for La Gomera: Implications for the overall evolution of the Canarian Archipelago. *Journal of Volcanology and Geothermal Research* 157: 271–293.

Araújo MB, Nogués-Bravo D, Diniz-Filho JAF, Haywood AM, Valdes PJ, Rahbek C. 2008. Quaternary climate changes explain diversity among reptiles and amphibians. *Ecography* 31: 8–15.

Arctander P, Johansen C, Coutellec-Vreto M. 1999. Phylogeography of three closely related African bovids (tribe Alcelaphini). *Molecular Biology and Evolution* 16: 1724–1739.

Arjona Y, Nogales M, Vargas P. 2014. Phylogenetic correction to infer number of colonization events to oceanic islands. Poster presented in the conferece *Modern Phylogenetic Comparative Methods*. Seville, Spain.

Assefa A, Ehrich D, Taberlet P, Nemomissa S, Brochmann C. 2007. Pleistocene colonization of afro-alpine *sky islands* by the arctic-alpine *Arabis alpina*. *Heredity* 99: 133–142.

Avise JC, Arnold J, Ball RM, Bermingham E, Lamb T, Neigel JE, Reeb C, Saunders NC. 1987. Intraspecific Phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology, Evolution, and Systematics* 18: 489–522.

Avise J. 1994. *Molecular markers, natural history and evolution*. Chapman & Hall, New York.

Avise J. 2000. *Phylogeography. The history and formation of species*. Harvard University Press, Cambridge, MA.

Axelrod DI. 1975. Evolution and biogeography of Madrean-Tethyan sclerophyll vegetation. *Annals of the Missouri Botanical Garden* 62: 280-334.

Axelrod DI, Raven PH. 1978. Late Cretaceous and Tertiary vegetation history of Africa. In: *Biogeography and ecology of southern Africa* (ed. Werger MJA) pp. 77-130. Springer Netherlands.

Ayele T, Gailing O, Umer MFinkeldey R. 2009. Chloroplast DNA haplotype diversity and postglacial recolonization of *Hagenia abyssinica* (Bruce) J.F. Gmel. in Ethiopia. *Plant Systematics and Evolution* 280: 175–185.

Barber J, Ortega J, Santos-Guerra A, Marrero A, Jansen R. 2000. Evolution of Endemic *Sideritis* (Lamiaceae) in Macaronesia: Insights from a Chloroplast DNA Restriction Site Analysis. *Systematic Botany* 25: 633–647.

Barber J, Francisco-Ortega J, Santos-Guerra A, Turner K, Jansen R. 2002. Origin of Macaronesian *Sideritis* L. (Lamioideae: Lamiaceae) inferred from nuclear and chloroplast sequence datasets. *Molecular Phylogenetics and Evolution* 23: 293–306.

Bayes T. 1763. A Letter from the late reverend Mr. Thomas Bayes, F. R. S. to John Canton, M. A. and F. R. S. *Philosophical Transactions of the Royal Society of London* 53: 269–271.

Beaumont MA. 2010. Approximate Bayesian Computation in evolution and ecology. *Annual Review of Ecology, evolution and Systematics* 41: 379–406.

Belay G, Mori A. 2006. Intraspecific phylogeographic mitochondrial DNA (D-loop) variation of Gelada baboon, *Theropithecus gelada*, in Ethiopia. *Biochemical Systematics and Ecology* 34: 554–561.

Bell CD, Soltis DE, Soltis PS. 2010. The age and diversification of the angiosperms re-revisited. *American Journal of Botany* 97: 1296–1303.

Bensch S, Akesson M. 2005. Ten years of AFLP in ecology and evolution: why so few animals? *Molecular Ecology* 14: 2899–2914.

Bloomquist EW, Lemey P, Suchard MA. 2010. Three roads diverged? Routes to phylogeographic inference. *Trends in Ecology & Evolution* 25: 626–632.

Bobe R, Behrensmeyer A. 2004. The expansion of grassland ecosystems in Africa in relation to mammalian evolution and the origin of the genus *Homo*. *Palaeogeography, Palaeoclimatology, Palaeoecology* 207: 399–420.

Bobe R. 2006. The evolution of arid ecosystems in eastern Africa. *Journal of Arid Environments* 66: 564–584.

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

Bonin A, Ehrich D, Manel S. 2007. Statistical analysis of amplified fragment length polymorphism data: a toolbox for molecular ecologists and evolutionists. *Molecular Ecology* 16: 3737–3758.

Bonnefille R, Roeland J, Guiot J. 1990. Temperature and rainfall estimates for the past 40,000 years in equatorial Africa. *Nature* 346: 347–349.

Boulesteix T, Hildenbrand A, Gillot P, Soler V. 2012. Eruptive response of oceanic islands to giant landslides: New insights from the geomorphologic evolution of the Teide–Pico Viejo volcanic complex (Tenerife, Canary). *Geomorphology* 138: 61–73.

Bowie R, Fjeldså J, Hackett S, Crowe T. 2004. Molecular evolution in space and through time: mtDNA phylogeography of the Olive Sunbird (*Nectarinia olivacea*/*obscura*) throughout continental Africa. *Molecular Phylogenetics and Evolution* 33: 56–74.

Bramwell D. 1972. Endemism in the flora of the Canary Islands. In: *Taxonomy, Phytogeography and Evolution* (ed. Valentine DH) pp. 141–159. Academic Press, London.

Bramwell D. 1985. Contribución a la biogeografía de las islas Canarias. *Botánica Macaronésica* 14: 3–34.

Braun EL. 1955. The phytogeography of unglaciated Eastern United States and its interpretation. The *Botanical Review* 21: 297–375.

Brito JC, Godinho R, Martínez-Freiría F, Pleguezuelos JM, Rebelo H, Santos X, Vale C, Velo-Antón G, Boratyński Z, Carvalho S, Ferreira S, Gonçalves D, Silva T, Tarroso P, Campos J, Leite J, Nogueira J, Álvares F, Sillero N, Sow A, Fahd S, Crochet P, Carranza S. 2013. Unravelling biodiversity, evolution and threats to conservation in the Sahara-Sahel. *Biological Reviews* 89: 215–231.

Brown JM, Hedtke SM, Lemmon AR, Lemmon EM. 2010. When trees grow too long: investigating the causes of highly inaccurate Bayesian branch-length estimates. *Systematic Biology* 59: 145–161.

Buffon GLL de. 1787. *Histoire naturelle*. Imprimerie Royale, Paris.

Bussmann R. 2001. Succession and Regeneration Patterns of East African Mountain Forests. A Review. *Systematics and Geography of Plants* 71: 959–974.

Canals M, Urgeles R, Masson DG, Bermúdez JLC. 2000. Los deslizamientos submarinos de las Islas Canarias. *Makaronesia: Boletín de la Asociación de Amigos del Museo de Ciencias Naturales de Tenerife* 2: 57–69.

Candolle AP de. 1820. *Essai élémentaire de géographie botanique*. Levrault, Strasbourg.

Cantagrel J, Arnaud N, Ancochea E, Fúster J, Huertas M. 1999. Repeated debris avalanches on Tenerife and genesis of Las Cañadas caldera wall (Canary Islands). *Geology* 27: 739–742.

Carine M, Russell S, Santos-Guerra A, Francisco-Ortega J. 2004. Relationships of the Macaronesian and Mediterranean floras: molecular evidence for multiple colonizations into Macaronesia and back-colonization of the continent in *Convolvulus* (Convolvulaceae). *American Journal of Botany* 91: 1070–1085.

Carine M. 2005. Spatio-Temporal Relationships of the Macaronesian Endemic Flora: A Relictual Series or Window of Opportunity? *Taxon* 54: 895.

Carlquist S. 1970. Wood anatomy of insular species of *Plantago* and the problem of raylessness. *Bulletin of the Torrey Botanical Club* 97: 353–361.

Carlquist S. 1974. *Island biology*. Columbia University Press, New York & London.

Carracedo JC, Day S, Guillou H, Rodríguez Badiola E, Canas JA, Pérez Torrado FJ. 1998. Hotspot volcanism close to a passive continental margin: the Canary Islands. *Geological Magazine* 135: 591–604.

Carracedo J. 1999. Growth, structure, instability and collapse of Canarian volcanoes and comparisons with Hawaiian volcanoes. *Journal of Volcanology and Geothermal Research* 94: 1–19.

Carracedo J, Badiola E, Guillou H, De la Nuez J, Pérez Torrado F. 2001. Geology and volcanology of La Palma and El Hierro. Western Canaries. *Estudios Geológicos* 57: 1–24.

Carracedo JC, Day S. 2002. *Canary Islands*. Dunedin Academic Press Ltd.

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

Caujapé-Castells J. 2009. General G_{ST} and θ inflation due to biased intra-population sampling, and its consequences for the conservation of the Canarian Flora. *Conservation Genetics* 11: 709–720.

Caujapé-Castells J, Santos-Guerra A, Jardim R, Gouveia L, Melo J, Rodrigues NP, Bramwell D, Gómes I. 2010. The status of plant conservation on the Macaronesian archipelagos. In: *Proceedings of the 4th Global Botanic Gardens Congress* (pp. 13-18). IUCN editions, London.

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

Cavalier-Smith T. 1998. A revised six-kingdom system of life. *Biological Reviews* 73: 203–266.

Cavalli-Sforza LL, Edwards AW. 1967. Phylogenetic Analysis: Models and Estimation Procedures. *Evolution* 21: 233-257.

Cerling T, Harris J, MacFadden B, Leakey M, Quade J, Eisenmann V, Ehleringer J. 1997. Global vegetation change through the Miocene/Pliocene boundary. *Nature* 389: 153–158.

Chorowicz J. 2005. The East African rift system. *Journal of African Earth Sciences* 43: 379–410.

Clement M, Posada D, Crandall KA. 2000. TCS: a computer program to estimate gene genealogies. *Molecular Ecology* 9: 1657–1659.

Coetzee JA. 1964. Evidence for a considerable depression of the vegetation belts during the Upper Pleistocene on the East African Mountains. *Nature* 204: 564–566.

Cohen A, Soreghan M, Scholz C. 1993. Estimating the age of formation of lakes: An example from Lake Tanganyika, East African Rift system. *Geology* 21: 511–514.

Comps B, Gömöry D, Letouzey J, Thiébaut B, Petit RJ. 2001. Diverging trends between heterozygosity and allelic richness during postglacial colonization in the European beech. *Genetics* 157: 389–397.

Cosner ME, Raubeson LA, Jansen RK. 2004. Chloroplast DNA rearrangements in Campanulaceae: phylogenetic utility of highly rearranged genomes. *BMC Evolutionary Biology* 4: 27, doi:10.1186/1471-2148-4-27.

Crisp MD, Cook LG. 2007. A congruent molecular signature of vicariance across multiple plant lineages. *Molecular Phylogenetics and Evolution* 43: 1106–1117.

Croizat L. 1958. *Panbiogeography*. Publicado por el autor, Caracas.

Cronk QCB. 1997. Islands: stability, diversity, conservation. *Biodiversity & Conservation* 6: 477–493.

Crow JF, Kimura M. 1970. *An introduction to population genetics theory*. Harper & Row, New York.

Csilléry K, Blum MG, Gaggiotti OE, François O. 2010. Approximate Bayesian computation (ABC) in practice. *Trends in ecology & evolution* 25: 410–418.

Darlington P. 1965. *Biogeography of the southern end of the world: distribution and history of far-southern life and land, with and assessment of continental drift*. Harvard University Press, Cambridge, MA.

Darwin C. 1859. *On the origin of species by means of natural selection: or, the preservation of favoured races in the struggle for life*. Murray J, Albemarle Street, London.

de Nascimento L, Willis K, Fernández-Palacios J, Criado C, Whittaker R. 2009. The long-term ecology of the lost forests of La Laguna, Tenerife (Canary Islands). *Journal of Biogeography* 36: 499–514.

Désamoré A, Laenen B, Devos N, Popp M, González-Mancebo JM, Carine MA, Vanderpoorten A. 2011. Out of Africa: north-westwards Pleistocene expansions of the heather *Erica arborea*. *Journal of Biogeography* 38: 164–176.

Di Rienzo A, Peterson AC, Garza JC, Valdes AM, Slatkin M, Freimer NB. 1994. Mutational processes of simple-sequence repeat loci in human populations. *Proceedings of the National Academy of Sciences* 91: 3166–3170.

Diamond JM. 1972. Biogeographic Kinetics: estimation of relaxation times for avifaunas of southwest Pacific islands. *Proceedings of the National Academy of Sciences* 69: 3199–3203.

Dóniz J, Romero C, Coello E, Guillén C, Sánchez N, García-Cacho L, García A. 2008. Morphological and statistical characterisation of recent mafic volcanism on Tenerife (Canary Islands, Spain). *Journal of Volcanology and Geothermal Research* 173: 185–195.

Donoghue PCJ, Benton MJ. 2007. Rocks and clocks: calibrating the Tree of Life using fossils and molecules. *Trends in Ecology & Evolution* 22: 424–431.

Drummond AJ, Ho SY, Phillips MJ, Rambaut A. 2006. Relaxed phylogenetics and dating with confidence. *PLoS Biology* 4: e88.

Drummond AJ, Rambaut A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology* 7: 214, doi:10.1186/1471-2148-7-214.

Dupont YL, Hansen DM, Rasmussen JT, Olesen JM. 2004. Evolutionary changes in nectar sugar composition associated with switches between bird and insect pollination: the Canarian bird-flower element revisited. *Functional Ecology* 18: 670–676.

Ebinger CJ. 1989a. Tectonic development of the western branch of the East African rift system. *Geological Society of America Bulletin* 101: 885–903.

Ebinger CJ, Deino AL, Drake RE, Tesha AL. 1989b. Chronology of volcanism and rift basin propagation: Rungwe Volcanic Province, East Africa. *Journal of Geophysical Research: Solid Earth* 94: 15785–15803.

Ebinger C, Yemane T, Woldegabriel G, Aronson J, Walter R. 1993. Late Eocene–Recent volcanism and faulting in the southern main Ethiopian rift. *Journal of the Geological Society* 150: 99-108.

Ehrich D, Gaudeul M, Assefa A, Koch M, Mummenhoff K, Nemomissa S, ... Brochmann C. 2007. Genetic consequences of Pleistocene range shifts: contrast between the Arctic, the Alps and the East African mountains. *Molecular Ecology* 16: 2542–2559.

Emerson BC. 2003. Genes, geology and biodiversity: faunal and floral diversity on the island of Gran Canaria. *Animal biodiversity and conservation* 26: 9–20.

Engler A. 1892. *Über die Hochgebirgsflora des tropischen Afrika*. Königliche Akademie der Wissenschaften zu Berlin, Berlin.

Engler A. 1879. *Versuch einer Eintwicklungsgeschichte, insbesondere der Florengebiete seit der Tertiärperiode. I. Die extra-tropischen Gebiete der nördlischen Hemisphäre*. Engelmann, Leipzig, Germany.

Engler A. 1904. Plants of the northern temperate zone in their transition to the high mountains of Tropical Africa. *Annals of Botany* 4: 523–540.

FAO. 2001. Food and Agricultural Organization of the United Nations. Forest Resource Assessment, 2000 (FAO, Rome). Forestry Paper 140.

Fernández-Palacios J, de Nascimento L, Otto R, Delgado J, García-del-Rey E, Arévalo J, Whittaker R. 2011. A reconstruction of Palaeo-Macaronesia, with particular reference to the long-term biogeography of the Atlantic island laurel forests. *Journal of Biogeography* 38: 226–246.

Fernández-Palacios JM, Rijsdijk KF, Norder SJ, Otto R, de Nascimento L, Fernández-Lugo S, Tjørve E, Whittaker RJ. 2015. Towards a glacial-sensitive model of island biogeography. *Global Ecology and Biogeography* doi: 10.1111/geb.12320.

Fernandes C, Rohling E, Siddall M. 2006. Absence of post-Miocene Red Sea land bridges: biogeographic implications. *Journal of Biogeography* 33: 961–966.

Fisher RA. 1920. A mathematical examination of the methods of determining the accuracy of observation by the mean error, and by the mean square error. *Monthly Notices of the Royal Astronomical Society* 80: 758–770.

Fitch WM. 1971. Toward defining the course of evolution: minimum change for a specific tree topology. *Systematic Biology* 20: 406–416.

Fitch F, Hooker P, Miller J, Mitchell J, Watkins R. 1985. Reconnaissance potassium–argon geochronology of the Suregei–Asille district, northern Kenya. *Geological Magazine* 122: 609–622.

Fjeldså J, Lovett JC. 1997. Geographical patterns of old and young species in African forest biota: the significance of specific montane areas as evolutionary centres. *Biodiversity & Conservation* 6: 325–346.

Flannery J, Rosendahl B. 1990. The seismic stratigraphy of Lake Malawi, Africa: implications for interpreting geological processes in lacustrine rifts. *Journal of African Earth Sciences (and the Middle East)* 10: 519–548.

Flenley JR. 1979. *The equatorial rain forest: a geological history*. Butterworth & Co, London/Boston.

Francisco-Ortega J, Jansen R, 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: 4085–4090.

Francisco-Ortega J, Santos-Guerra A, Hines A, Jansen R. 1997. Molecular Evidence for a Mediterranean Origin of the Macaronesian Endemic Genus *Argyranthemum* (Asteraceae). *American Journal of Botany* 84: 1595–1613.

Francisco-Ortega J, Goertzen L, Santos-Guerra A, Benabid A, Jansen R. 1999. Molecular Systematics of the *Asteriscus* Alliance (Asteraceae: Inuleae) I: Evidence from the Internal Transcribed Spacers of Nuclear Ribosomal DNA. *Systematic Botany* 24: 249–266.

Francisco-Ortega J, Santos-Guerra A, Kim S, Crawford D. 2000. Plant Genetic Diversity in the Canary Islands: A Conservation Perspective. *American Journal of Botany* 87: 909–919.

Francisco-Ortega J, Park S, Santos-Guerra A, Benabid A, Jansen R. 2001. Origin and evolution of the endemic Macaronesian Inuleae (Asteraceae): evidence from the internal transcribed spacers of nuclear ribosomal DNA. *Biological Journal of the Linnean Society* 72: 77–97.

Francisco-Ortega J, Santos-Guerra A. 2001. Genes y conservación de plantas vasculares. In: *Naturaleza de las Islas Canarias: ecología y conservación* (eds. Fernández–Palacios JM, Martín–Esquivel JL) pp. 357–365. Publicaciones Turquesa, Santa Cruz de Tenerife.

Francisco-Ortega J, Fuertes-Aguilar J, Kim S, Santos-Guerra A, Crawford D, Jansen R. 2002. Phylogeny of the Macaronesian endemic *Crambe* section Dendrocrambe (Brassicaceae) based on internal transcribed spacer sequences of nuclear ribosomal DNA. *American Journal of Botany* 89: 1984–1990.

Frankham R, Briscoe DA, Ballou JD. 2002. *Introduction to conservation genetics*. Cambridge University Press, Cambridge, UK.

Freitag S, Robinson T. 1993. Phylogeographic Patterns in Mitochondrial DNA of the Ostrich (*Struthio camelus*). *The Auk* 110: 614–622.

Fuertes-Aguilar J, Ray MF, Francisco-Ortega J, Santos-Guerra A, Jansen RK. 2002. Molecular evidence from chloroplast and nuclear markers for multiple colonizations of *Lavatera* (Malvaceae) in the Canary Islands. *Systematic Botany* 27: 74–83.

Galley C, Bytebier B, Bellstedt D, Linder H. 2007. The Cape element in the Afrotemperate flora: from Cape to Cairo? *Proceedings of the Royal Society B: Biological Sciences* 274: 535–543.

García-Verdugo C, Sajeva M, La Mantia T, Harrouni C, Msanda F, Caujapé-Castells J. 2015. Do island plant populations really have lower genetic variation than mainland populations? Effects of selection and distribution range on genetic diversity estimates. *Molecular Ecology* 24: 726–741.

Gavin DG, Fitzpatrick MC, Gugger PF, Heath KD, Rodríguez-Sánchez F, Dobrowski SZ,…Williams JW. 2014. Climate refugia: joint inference from fossil records, species distribution models and phylogeography. *New Phytologist* 204: 37–54.

Gee M, Watts A, Masson D, Mitchell N. 2001. Landslides and the evolution of El Hierro in the Canary Islands. *Marine Geology* 177: 271–293.

Gehrke B, Linder H. 2009. The scramble for Africa: pan-temperate elements on the

African high mountains. *Proceedings of the Royal Society B: Biological Sciences* 276: 2657–2665.

Gehrke B, Linder H. 2014. Species richness, endemism and species composition in the tropical Afroalpine flora. *Alpine Botany* 124: 165–177.

Geldmacher J, Hoernle K, van den Bogaard P, Zankl G, Garbe-Schönberg D. 2001. Earlier history of the ≥70-Ma-old Canary hotspot based on the temporal and geochemical evolution of the Selvagen Archipelago and neighboring seamounts in the eastern North Atlantic. *Journal of Volcanology and Geothermal Research* 111: 55–87.

Geldmacher J, Hoernle K, Bogaard P, Duggen S, Werner R. 2005. New 40Ar/39Ar age and geochemical data from seamounts in the Canary and Madeira volcanic provinces: Support for the mantle plume hypothesis. *Earth and Planetary Science Letters* 237: 85–101.

Gillett JB. 1955. The relation between the highland floras of Ethiopia and British east Africa. Webbia 11: 459–469.

Givnish TJ, Renner SS. 2004. Tropical intercontinental disjunctions: Gondwana breakup, immigration from the boreotropics, and transoceanic dispersal. *International Journal of Plant Sciences* 165: S1–S6.

Gizaw A, Kebede M, Nemomissa S, Ehrich D, Bekele B, Mirré V, ... Brochmann C. 2013. Phylogeography of the heathers *Erica arborea* and *E. trimera* in the afro-alpine 'sky islands' inferred from AFLPs and plastid DNA sequences. *Flora-Morphology, Distribution, Functional Ecology of Plants* 208: 453–463.

Goldman N. 1993. Simple diagnostic statistical tests of models for DNA substitution. *Journal of Molecular Evolution* 37: 650–661.

Gómez A, Lunt DH. 2007. Refugia within refugia: patterns of phylogeographic concordance in the Iberian Peninsula. In: *Phylogeography of southern European refugia* (eds. Weiss S, Ferrand N) pp. 155–188. Springer Netherlands.

Gottelli D, Marino J, Sillero-Zubiri C, Funk SM. 2004. The effect of the last glacial age on speciation and population genetic structure of the endangered Ethiopian wolf (*Canis simensis*). *Molecular Ecology* 13: 2275–2286.

Goudie A. 2005. The drainage of Africa since the Cretaceous. *Geomorphology* 67: 437–456.

Guillou H, Carracedo J, Torrado F, Badiola E. 1996. K-Ar ages and magnetic stratigraphy of a hotspot-induced, fast grown oceanic island: El Hierro, Canary Islands. *Journal of Volcanology and Geothermal Research* 73: 141–155.

Hackman B, Charsley T, Key R, Wilkinson A. 1990. The development of the East African Rift system in north-central Kenya. *Tectonophysics* 184: 189–211.

Haeckel EH. 1866. *Generelle Morphologie der Organismen*. Reimer G, Berlin.

Hanski IA, Simberloff D. 1997. The metapopulation approach, its history, conceptual domain, and application to conservation biology. In: *Metapopulation Biology* (eds. Hanski IA, Gilipin ME) pp. 5–26. Academic Press, San Diego, California.

Harter DE, Irl SD, Seo B, Steinbauer MJ, Gillespie R, Triantis KA, Fernández-Palacios JM, Beierkuhnlein C. 2015. Impacts of global climate change on the floras of oceanic islands – Projections, implications and current knowledge. *Perspectives in Plant Ecology, Evolution and Systematics* 17: 160–183.

Hawando T. 1997. Desertification in Ethiopian highlands. *Rala Report* (200).

Hawkins BA. 2001. Ecology's oldest pattern? *Endeavour* 25: 133–134.

Heads M. 1999. Vicariance biogeography and terrane tectonics in the South Pacific: analysis of the genus *Abrotanella* (Compositae). *Biological Journal of the Linnean Society* 67: 391–432.

Heads M. 2005. Dating nodes on molecular phylogenies: a critique of molecular biogeography.*Cladistics* 21: 62–78.

Hedberg O. 1951. Vegetation belts of the East African mountains. *Svensk Botanisk Tidskrift Utgifven af Svenska Botaniska Foreningen. Stockholm* 45: 140–202.

Hedberg O. 1961a. The phytogeographical position of the Afroalpine flora. *Recent*

Advances in Botany 1: 914–919.

Hedberg O. 1961b. Monograph of the genus *Canarina* L. (Campanulaceae). *Svensk Botanisk Tidskrift* 55: 17–62.

Hedberg O. 1965. Afroalpine flora elements. *Webbia* 19: 519–529.

Hedberg O. 1969. Evolution and speciation in a tropical high mountain flora. *Biological Journal of the Linnean Society* 1: 135–148.

Hedberg O. 1970. Evolution of the afroalpine flora. *Biotropica* 2: 16–23.

Hedberg O. 1986. Evolution of the Afroalpine flora. In: *High altitude tropical biogeography* (eds. Vuilleumier F, Monastero M) pp. 443–468. Oxford University Press, New York.

Heled J, Drummond AJ. 2010. Bayesian inference of species trees from multilocus data. *Molecular Biology and Evolution* 27: 570–580.

Herben T, Suda J, Munclinger P. 2005. The ghost of hybridization past: niche pre-emption is not the only explanation of apparent monophyly in island endemics. *Journal of Ecology* 93: 572–575.

Heuertz M, Teufel J, González-Martínez SC, Soto A, Fady B, Alía R, Vendramin GG. 2010. Geography determines genetic relationships between species of mountain pine (*Pinus mugo* complex) in western Europe. *Journal of biogeography 37*: 541-556.

Hewitt G. 2000. The genetic legacy of the Quaternary ice ages. *Nature* 405: 907–913.

Hewitt G. 2001. Speciation, hybrid zones and phylogeography - or seeing genes in space and time. *Molecular Ecology* 10: 537–549.

Hewitt G. 2004. The structure of biodiversity - insights from molecular phylogeography. *Frontiers in Zoology* 1: 1–16.

Ho SYW, Phillips MJ. 2009. Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times. *Systematic Biology* 58: 367–380.

Hofmann C, Courtillot V, Feraud G, Rochette P, Yirgu G, Ketefo E, Pik R. 1997. Timing of the Ethiopian flood basalt event and implications for plume birth and global change. *Nature* 389: 838–841.

Hooker JD. 1853. *Introductory essay to the flora of New Zealand*. Reeve L, London.

Hooker JD. 1860. *The botany of the Antarctic voyage of H.M. discovery ships 'Erebus' and 'Terror', in the years 1839–1843 under the command of Captain Sir James Clark Ross*. Reeve Brothers, London.

Huelsenbeck JP, Ronquist F. 2001. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 17: 754–755.

Hughes GW, Varol O, Beydoun ZR. 1991. Evidence for Middle Oligocene rifting of the Gulf of Aden and for Late Oligocene rifting of the southern Red Sea. *Marine and Petroleum Geology* 8: 354–358.

Ibáñez M. 1994. Los moluscos terrestres endémicos del archipiélago canario. In: *Actas del X Congreso nacional de Macalogía*, pp. 23-27. Barcelona.

Ibarrola E, Ancochea E, Fuster JM, Cantagrel JM, Coello J, Snelling NJ, Huertas MJ. 1993. Cronoestratigrafía del Macizo de Tigaiga: Evolución de un sector del edificio Cañadas (Tenerife, Islas Canarias). *Boletín de la Real Sociedad Española de Historia Natural. Sección Geológica* 88: 57–72.

Juan C, Emerson BC, Oromí P, Hewitt GM. 2000. Colonization and diversification: towards a phylogeographic synthesis for the Canary Islands. *Trends in Ecology and Evolution* 15: 104–109.

Kadereit JW, Baldwin BG. 2012. Western Eurasian–western North American disjunct plant taxa: The dry-adapted ends of formerly widespread north temperate mesic lineages —and examples of long-distance dispersal. *Taxon* 61: 3–17.

Kadu C, Schueler S, Konrad H, Muluvi G,Eyog-Matig O, Muchugi A, ...Geburek T. 2011. Phylogeography of the Afromontane *Prunus africana* reveals a former migration corridor between East and West African highlands. *Molecular Ecology* 20: 165–178.

Kahindo C, Bowie RC, Bates JM. 2007. The relevance of data on genetic diversity for the conservation of Afromontane regions. *Biological Conservation* 134: 262–270.

Kampunzu A, Bonhomme M, Kanika M. 1998. Geochronology of volcanic rocks and evolution of the Cenozoic Western Branch of the East African Rift System. *Journal of African Earth Sciences* 26: 441–461.

Kebede M, Ehrich D, Taberlet P, Nemomissa S, Brochmann C. 2007. Phylogeography and conservation genetics of a giant lobelia (*Lobelia giberroa*) in Ethiopian and Tropical East African mountains. *Molecular Ecology* 16: 1233–1243.

Kim S, Crawford D, Francisco-Ortega J, Santos-Guerra A. 1996. A common origin for woody *Sonchus* and five related genera in the Macaronesian islands: molecular evidence for extensive radiation. *Proceedings of the National Academy of Sciences* 93: 7743–7748.

Kim SC, Chunghee L, Mejías JA. 2007. Phylogenetic analysis of chloroplast DNA matK gene and ITS of nrDNA sequences reveals polyphyly of the genus *Sonchus* and new relationships among the subtribe Sonchinae (Asteraceae: Cichorieae). *Molecular phylogenetics and evolution* 44: 578-597.

Kimura M, Crow JF. 1964. The number of alleles that can be maintained in a finite population. *Genetics* 49: 725–738.

Kimura M, Ohta T. 1978. Stepwise mutation model and distribution of allelic frequencies in a finite population. *Proceedings of the National Academy of Sciences* 75: 2868–2872.

Kingdon J. 1990. *Island Africa: the evolution of Africa's rare animals and plants*. London, Collins.

Kingdon J. 2003. *Lowly origins: when, where, and why our ancestors first stood Up*. Princeton University Press, Princeton, NJ.

Kingman JFC. 1982. The coalescent. *Stochastic Processes and their Applications* 13: 235–248.

Kishino H, Thorne JL, Bruno WJ. 2001. Performance of a divergence time estimation method under a probabilistic model of ratenielsen evolution. *Molecular Biology and Evolution* 18: 352–361.

Kloos H, Legesse W. 2010. *Water resources management in Ethiopia: implications for the Nile basin*. Cambria Press, New York.

Knowles LL, Maddison WP. 2002. Statistical phylogeography. *Molecular Ecology* 11: 2623–2635.

Koch M, Kiefer C, Ehrich D, Vogel J, Brochmann C, Mummenhoff K. 2006. Three times out of Asia Minor: the phylogeography of *Arabis alpina* L. (Brassicaceae). *Molecular Ecology* 15: 825–839.

Kondraskov P, Schuetz N, Schuessler C, Sequeira M, Santos guerra A, Caujapé-castells J,…Thiv M. 2015. Biogeography of Mediterranean hotspot biodiversity: Re-evaluating the `Tertiary relict´ hypothesis of Macaronesian laurel forests. *PLoS ONE* doi: 10.1371/journal.pone.0132091.

Krastel S, Schmincke H, Jacobs C, Rihm R, Le Bas T, Alibés B. 2001. Submarine landslides around the Canary Islands. *Journal of Geophysical Research* 106: 3977–3997.

Kuper R, Kröpelin S. 2006. Climate-controlled Holocene occupation in the Sahara: motor of Africa's evolution. *Science* 313: 803–807.

Lehmann T. 2000. The Rift Valley complex as a barrier to gene flow for *Anopheles gambiae* in Kenya: the mtDNA perspective. *Journal of Heredity* 91: 165–168.

Lemey P, Rambaut A, Drummond AJ, Suchard MA. 2009. Bayesian phylogeography finds its roots. *PLoS Computational Biology* 5: e1000520. doi:10.1371/journal.pcbi.1000520.

Leroy SAG, Arpe K. 2007. Glacial refugia for summer-green trees in Europe and south-west Asia as proposed by ECHAM3 time-slice atmospheric model simulations. *Journal of Biogeography* 34: 2115–2128.

Levins R. 1969. Some demographic and genetic consequences of environmental heterogeneity for biological control. *Bulletin of the Entomological Society of America* 15: 237–240.

Levins R. 1970. Extinction. *Lectures on mathematics in the life sciences* 2: 75–107.

Lezzar K, Tiercelin J, Batist M, Cohen A, Bandora T, Rensbergen P, ... Klerkx J. 1996. New seismic stratigraphy and Late Tertiary history of the North Tanganyika Basin, East African Rift system, deduced from multichannel and high-resolution reflection seismic data and piston core evidence. *Basin Research* 8: 1–28.

Linder HP. 2014. The evolution of African plant diversity. *Frontiers in Ecology and Evolution* 2:1–14.

Linnaeus CV. 1738. *Hortus cliffortianus*. Salomonem Schouten, Amstelaedami.

Linnaeus CV. 1753. *Species Plantarum*. Ed. Salvii, Estocolmo.

Linnaeus CV. 1788. *Systema naturae per regna tria naturae secundum classes, ordines, genera, species* (Vol. 1). Holmiae: Impensis Direct. Laurentii Salvii.

Llanes P, Muñoz A, Muñoz-Martín A, Acosta J, Herranz P, Carbó A, Palomo C. 2003. Morphological and structural analysis in the Anaga offshore massif, Canary Islands: fractures and debris avalanches relationships. *Marine Geophysical Researches* 24: 91–112.

Llanes P, Herrera R, Gómez M, Muñoz A, Acosta J, Uchupi E, Smith D. 2009. Geological evolution of the volcanic island La Gomera, Canary Islands, from analysis of its geomorphology. *Marine Geology* 264: 123–139.

Lomolino MK. 2000. A call for a new paradigm of island biogeography. *Global Ecology and Biogeography* 9: 1–6.

Lomoschitz A, Meco J, Corominas J. 2002. The Barranco de Tirajana basin, Gran Canaria (Spain). A major erosive landform caused by large landslides. *Geomorphology* 42: 117–130.

Luo R, Hipp AL, Larget B. 2007. A Bayesian model of AFLP marker evolution and phylogenetic inference. *Statistical Applications in Genetics and Molecular Biology* 6, article 11. Available at http://www.bepress.com/sagmb/vol6/iss1/art11.

Lyell SC. 1830. *Principles of Geology*. Murray, London.

MacArthur RH, Wilson EO. 1963. An equilibrium theory of insular zoogeography. *Evolution* 17: 373–387.

MacArthur RH, Wilson EO. 1967. *The theory of island biogeography*. Princeton University Press, Princeton.

Carrillo AM, Oromí P. 2000. *Elenco de los coleópteros de las Islas Canarias*. Instituto de Estudios Canarios, La Laguna.

Magallón S. 2014. Una revisión del efecto de los métodos de reloj molecular relajado, ramas largas, genes y calibraciones en la estimación de la edad de las angiospermas. *Botanical Sciences 92*: 1–22.

Magallón S, Gómez-Acevedo S, Sánchez-Reyes LL, Hernández-Hernández T. 2015. A metacalibrated time-tree documents the early rise of flowering plant phylogenetic diversity. *New Phytologist* doi: 10.1111/nph.13264.

Maley J. 1996. The African rain forest –main characteristics of changes in vegetation and climate from the Upper Cretaceous to the Quaternary. *Proceedings of the Royal Society of Edinburgh. Section B. Biological Sciences* 104: 31–73.

Manolopoulou I, Legarreta L, Emerson BC, Brooks S, Tavaré S. 2011. A Bayesian approach to phylogeographic clustering. *Interface focus*, rsfs20110054.

Manolopoulou I, Emerson BC. 2012. Phylogeographic ancestral inference using the coalescent model on haplotype trees. *Journal of Computational Biology* 19: 745–946.

Mansion G, Parolly G, Crowl AA, Mavrodiev E, Cellinese N, Oganesian M,… Borsch T. 2012. How to handle speciose clades? Mass taxon-sampling as a strategy towards illuminating the Natural History of *Campanula* (Campanuloideae). *PLoS ONE* doi: 10.1371/journal.pone.0050076.

Margulis L. 1974. Five-kingdom classification and the origin and evolution of cells. *Evolutionary Biology* 7: 45–78.

McCormack JE, Huang H, Knowles LL. 2009. Sky islands. In: *Encyclopedia of Islands* (eds. Gillespie RG, Clague DA) pp. 841–843. University of California Press, Berkeley and Los Angeles.

Marrero A, Francisco-Ortega J. 2001. Evolución en islas: la metáfora especio-tiempo-forma. In: *Naturaleza de las Islas Canarias: ecología y conservación* (eds.

Fernández–Palacios JM, Martín–Esquivel JL) pp. 133–140. Publicaciones Turquesa, Santa Cruz de Tenerife.

Martí J, Geyer A, Andujar J, Teixidó F, Costa F. 2008. Assessing the potential for future explosive activity from Teide–Pico Viejo stratovolcanoes (Tenerife, Canary Islands). *Journal of Volcanology and Geothermal Research* 178: 529–542.

Masson D, Watts A, Gee M, Urgeles R, Mitchell N, Le Bas T, Canals M. 2002. Slope failures on the flanks of the western Canary Islands. *Earth-Science Reviews* 57: 1–35.

Meco J, Petit-Maire N, Guillou H, Carracedo JC, Lomoschitz A, Ramos AJG, Ballester J. 2003. Climatic changes over the last 5,000,000 years as recorded in the Canary Islands. *Episodes* 6: 133–134.

Meco J, Ballester J, Betancort JF, Cilleros A, Scaillet S, Guillou H, Carracedo JC, Lomoschitz A, Petit-Maire N, Ramos AJG, Perera N, Meco JM. 2006. *Paleoclimatología del Neógeno en las Islas Canarias. Geliense, Pleistoceno y Holoceno*. Ministerio de Medio Ambiente, Universidad de Las Palmas de Gran Canaria, Las Palmas de Gran Canaria.

Médail F, Diadema K. 2009. Glacial refugia influence plant diversity patterns in the Mediterranean Basin. *Journal of Biogeography* 36: 1333–1345.

Mège D, Korme T. 2004. Dyke swarm emplacement in the Ethiopian Large Igneous Province: not only a matter of stress. *Journal of Volcanology and Geothermal Research* 132: 283–310.

Meimberg H, Abele T, Bräuchler C, McKay JK, de Paz PLP, Heubl G. 2006. Molecular evidence for adaptive radiation of *Micromeria* Benth. (Lamiaceae) on the Canary Islands as inferred from chloroplast and nuclear DNA sequences and ISSR fingerprint data. *Molecular Phylogenetics and Evolution* 41: 566–578.

Melonti M, Reid A, Caujapé-Castells J, Soto M, Fernández-Palacios JM, Conti E. 2015. High genetic diversity and population structure in the endangered Canarian endemic *Ruta oreojasme* (Rutaceae). *Genetica* doi:10.1007/s10709-015-9855-0.

Meseguer AS, Aldasoro JJ, Sanmartín I. 2013. Bayesian inference of phylogeny, morphology and range evolution reveals a complex evolutionary history in St.

John’s wort (*Hypericum*). *Molecular phylogenetics and evolution* 67: 379–403.

Meseguer AS, Lobo JM, Ree R, Beerling DJ, Sanmartin I. 2015. Integrating fossils, phylogenies, and niche models into biogeography to reveal ancient evolutionary history: the case of *Hypericum* (Hypericaceae). *Systematic Biology* 64: 215–232.

Meudt HM, Clarke AC. 2007. Almost forgotten or latest practice? AFLP applications, analyses and advances. *Trends in Plant Science* 12: 106–117.

Mies BA. 1995. On the comparison of the flora and vegetation of the island groups of Socotra and Macaronesia. *Boletim* do *Museu* Municipal do *Funchal* 4: 455–471.

Milne RI, Abbott RJ. 2002. The origin and evolution of Tertiary relict floras. *Advances in Botanical Research* 38: 281–314.

Milne RI. 2006. Northern Hemisphere Plant Disjunctions: A Window on Tertiary Land Bridges and Climate Change? *Annals of Botany* 98: 465–472.

Mittelbach GG, Schemske DW, Cornell HV, Allen AP, Brown JM, Bush MB,… Turelli M. 2007. Evolution and the latitudinal diversity gradient: speciation, extinction and biogeography. *Ecology Letters* 10: 315–331.

Mittermeier R, Gil P, Hoffmann M, Pilgrim J, Brooks T, Mittermeier C, ... Fonseca G. 2004. *Hotspots revisited. CEMEX* (pp. 218-229). ISBN 968-6397-77-9.

Moreau R. 1966. *The Bird Faunas of Africa and its Islands*. Academic Press, London.

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

Myers N, Mittermeier RA, Mittermeier CG, Da Fonseca G, Kent J. 2000. Biodiversity hotspots for conservation priorities. *Nature* 403: 853–858.

Navarro-Pérez M, Vargas P, Fernández-Mazuecos M, López J, Valtueña F, Ortega-Olivencia A. 2015. Multiple windows of colonization to Macaronesia by the dispersal-unspecialized *Scrophularia* since the Late Miocene. *Perspectives in Plant Ecology, Evolution and Systematics* 17: 263–273.

Nei M. 1973. Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences* 70: 3321–3323.

Nielsen R, Beaumont MA. 2009. Statistical inferences in phylogeography. *Molecular Ecology* 18: 1034–1047.

Nieto-Feliner G. 2014. Patterns and processes in plant phylogeography in the Mediterranean Basin. A review. *Perspectives in Plant Ecology, Evolution and Systematics* 16: 265–278.

Oberprieler C. 2005. Temporal and Spatial Diversification of Circum-Mediterranean Compositae-Anthemideae. *Taxon* 54:951–966.

Olesen JM. 1985. The Macaronesian bird-flower element and its relation to bird and bee opportunists. *Botanical Journal of the Linnean Society* 91: 395–414.

Olesen JM, Alarcón M, Ehlers BK, Aldasoro JJ, Roquet C. 2012. Pollination, biogeography and phylogeny of oceanic island bellflowers (Campanulaceae). *Perspectives in Plant Ecology, Evolution and Systematics* 14: 169–182.

Olmstead RG, Wolfe AD, Young ND, Elisons WJ, Reeves PA. 2001. Disintegration of the Scrophulariaceae. *American Journal of Botany* 88: 348–361.

Pagel J, Schurr FM. 2012. Forecasting species ranges by statistical estimation of ecological niches and spatial population dynamics. *Global Ecology and Biogeography* 21: 293–304.

Patiño J, Carine M, Mardulyn P, Devos N, Mateo RG, Gonzalez-Mancebo JM, Shaw A, Vanderpoorten A. 2015. Approximate Bayesian Computation reveals the crucial role of oceanic islands for the assembly of continental biodiversity. *Systematic Biology* doi: 10.1093/sysbio/syv013.

Pelser P, Nordenstam B, Kadereit J, Watson L. 2007. An ITS Phylogeny of Tribe Senecioneae (Asteraceae) and a New Delimitation of *Senecio* L. *Taxon* 56: 1077–1104.

Percy D, Cronk Q. 2002. Different fates of island brooms: contrasting evolution in *Adenocarpus*, *Genista*, and *Teline* (Genisteae, Fabaceae) in the Canary Islands and Madeira. *American Journal of Botany* 89: 854–864.

Pérez de Paz J, Caujape-Castells J. 2013. A review of the allozyme data set for the Canarian endemic flora: causes of the high genetic diversity levels and implications for conservation. *Annals of Botany* 111: 1059–1073.

Pérez-Torrado F, Carracedo J, Mangas J. 1995. Geochronology and stratigraphy of the Roque Nublo Cycle, Gran Canaria, Canary Islands. *Journal of the Geological Society* 152: 807–818.

Petit RJ, Aguinagalde I, de Beaulieu JL, Bittkau C, Brewer S, Cheddadi R,… Vendramin GG. 2003. Glacial refugia: hotspots but not melting pots of genetic diversity. *Science* 300: 1563-1565.

Petit RJ, Hampe A, Cheddadi R. 2005. Climate changes and tree phylogeography in the Mediterranean. *Taxon* 54: 877-885.

Pickford M. 1982. New higher primate fossils from the middle Miocene deposits at Majiwa and Kaloma, western Kenya. *American Journal of Physical Anthropology* 58: 1–19.

Pickford M, Senut B, Hadoto D. 1993. *Geology and Palaeobiology of the Albertine Rift Valley, Uganda-Zaire: Geology* (Vol. 1). CIFEG: Orléans, France.

Pole M. 1994. The New Zealand flora-entirely long-distance dispersal? *Journal of Biogeography* 21: 625–635.

Popp M, Gizaw A, Nemomissa S, Suda J, Brochmann C. 2008. Colonization and diversification in the African "sky islands" by Eurasian *Lychnis* L. (Caryophyllaceae). *Journal of Biogeography* 35: 1016–1029.

Posada D, Crandall KA. 2001. Intraspecific gene genealogies: trees grafting into networks. *Trends in Ecology & Evolution* 16: 37–45.

Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155: 945–959.

Puff C, Nemomissa S. 2001. The Simen Mountains (Ethiopia): Comments on Plant Biodiversity, Endemism, Phytogeographical Affinities and Historical Aspects. *Systematics and Geography of Plants* 71: 975–991.

Puppo P, Curto M, Gusmão-Guedes J, Cochofel J, de Paz PLP, Bräuchler C, Meimberg H. 2015. Molecular phylogenetics of *Micromeria* (Lamiaceae) in the Canary Islands, diversification and inter-island colonization patterns inferred from nuclear genes. *Molecular Phylogenetics and Evolution* 89: 160–170.

Quammen D. 1996. *The song of the dodo: island biogeography in an age of extinction*. Hutchinson, London.

Rannala B, Yang Z. 1996. Probability distribution of molecular evolutionary trees: A new method of phylogenetic inference. *Journal of Molecular Evolution* 43: 304–311.

Rannala B, Yang Z. 2003. Bayes estimation of species divergence times and ancestral population sizes using DNA sequences from multiple loci. *Genetics* 164: 1645–1656.

Raven PH. 1972. Plant species disjunctions: a summary. *Annals of the Missouri Botanical Garden* 59: 234–246.

Raven PH, Axelrod DI. 1978. *Origin and relationships of the California flora* (Vol. 72). University of California Press, Berkeley and Los Angeles, California.

Rijsdijk KF, Hengl T, Norder SJ, Otto R, Emerson BC, Ávila SP,…Fernández-Palacios JM. 2014. Quantifying surface-area changes of volcanic islands driven by Pleistocene sea-level cycles: biogeographical implications for the Macaronesian archipelagos. *Journal of Biogeography* 41: 1242–1254.

Roberts E, Stevens N, O'Connor P, Dirks P, Gottfried M, Clyde W, ... Hemming S. 2012. Initiation of the western branch of the East African Rift coeval with the eastern branch. *Nature Geoscience* 5: 289–294.

Rodriguez-Gonzalez A, Fernandez-Turiel J, Perez-Torrado F, Paris R, Gimeno D, Carracedo J, Aulinas M. 2012. Factors controlling the morphology of monogenetic basaltic volcanoes: The Holocene volcanism of Gran Canaria (Canary Islands, Spain). *Geomorphology* 136: 31–44.

Rodríguez-Rodríguez MC, Valido A. 2011. Consequences of plant-pollinator and floral-herbivore interactions on the reproductive success of the Canary Islands endemic *Canarina canariensis* (Campanulaceae). *American Journal of Botany* 98: 1465–1474.

Ronquist F. 2003. Parsimony analysis of coevolving species associations (ed. Page RDM) pp. 22–64. In: *Tangled trees: phylogeny, cospeciation and coevolution*.

Ronquist F. 2004. Bayesian inference of character evolution. *Trends in Ecology & Evolution* 19: 475–481.

Ronquist F, Deans AR. 2010. Bayesian phylogenetics and its influence on insect systematics. *Annual Review of Entomology* 55: 189–206.

Ronquist F, Sanmartín I. 2011. Phylogenetic methods in biogeography. *Annual Review of Ecology, Evolution, and Systematics* 42: 441–464.

Roquet C, Sanmartín I, Garcia-Jacas N, Sáez L, Susanna A, Wikström N, Aldasoro JJ. 2009. Reconstructing the history of Campanulaceae with a Bayesian approach to molecular dating and dispersal–vicariance analyses. *Molecular Phylogenetics and Evolution* 52: 575–587.

Rosenberg NA, Nordborg M. 2002. Genealogical trees, coalescent theory and the analysis of genetic polymorphisms. *Nature Reviews Genetics* 3: 380–390.

Roy MS, Sponer R, Fjeldså J. 2001. Molecular Systematics and Evolutionary History of Akalats (Genus *Sheppardia*): A Pre-Pleistocene Radiation in a Group of African Forest Birds. *Molecular Phylogenetics and Evolution* 18: 74–83.

Ruddiman W. 2003. Orbital insolation, ice volume, and greenhouse gases. *Quaternary Science Reviews* 22: 1597–1629.

Saitou N, Nei M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular biology and evolution* 4: 406–425.

Sanderson MJ. 1997. A nonparametric approach to estimating divergence times in the absence of rate constancy. *Molecular Biology and Evolution* 14: 1218–1231.

Sanderson MJ. 2002. Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. *Molecular Biology and Evolution* 19: 101–109.

Sanmartín. 2003. Dispersal vs. vicariance in the Mediterranean: historical biogeography of the Palearctic Pachydeminae (Coleoptera, Scarabaeoidea). *Journal of Biogeography* 30 1883–1897.

Sanmartín I, Ronquist F. 2004. Southern Hemisphere biogeography inferred by event-based models: plant versus animal patterns. *Systematic Biology* 53: 216–243.

Sanmartín I, van der Mark P, Ronquist F. 2008. Inferring dispersal: a Bayesian approach to phylogeny-based island biogeography, with special reference to the Canary Islands. *Journal of Biogeography* 35: 428–449.

Sanmartín I, Anderson Cl, Alarcón M, Ronquist F, Aldasoro JJ. 2010. Bayesian island biogeography in a continental setting: the Rand Flora case. *Biology Letters* 6: 703–707.

Sanmartín I. 2012. Historical biogeography: evolution in time and space. *Evolution: Education and Outreach* 5: 555–568.

Scharf M. 1993. Itching due to bites, stings and infestations. In *Itch: Mechanisms and Management of Pruritus* (eds. Jeffrey D, Bernhard MD) pp. 135–145. McGraw-Hill Companies, New York.

Santos-Guerra A. 1999. Origen y evolución de la flora canaria. In: *Ecología y Cultura en Canarias* (eds. Fernández-Palacios JM, Bacallado JJ, Belmonte JA) pp. 107–129. Museo de las Ciencias y el Cosmos, Santa Cruz de Tenerife.

Santos-Guerra 2013. Floristic relationships between Macaronesia, mainly Canary Islands, and East Africa. In: *Proceedings of the Amurga international conferences on island biodiversity 2011* (eds. Caujapé-Castells J, Nieto Feliner G, Fernández Palacios JM) pp. 192–200. Fundación Canaria Amurga-Maspalomas, Las Palmas de Gran Canaria, Spain.

Saunders N, Gibson D. 2005. Breeding system, branching processes, hybrid swarm theory, and the humped-back diversity relationship as additional explanations for apparent monophyly in the Macaronesian island flora. *Journal of Ecology* 93: 649–652.

Schmincke H, Klügel A, Hansteen T, Hoernle K, van den Bogaard P. 1998. Samples from the Jurassic ocean crust beneath Gran Canaria, La Palma and Lanzarote (Canary Islands). *Earth and Planetary Science Letters* 163: 343–360.

Schneider J, Pérez Torrado F, Gimeno Torrente D, Wassmer P, del Carmen Cabrera Santana M, Carracedo J. 2004. Sedimentary signatures of the entrance of coarse-grained volcaniclastic flows into the sea: the example of the breccia units of the Las Palmas Detritic Formation (Mio–Pliocene, Gran Canaria, Eastern Atlantic, Spain). *Journal of Volcanology and Geothermal Research* 138: 295–323.

Schönswetter P, Stehlik I, Hholderegger R, Tribsch A. 2005. Molecular evidence for glacial refugia of mountain plants in the European Alps. *Molecular Ecology* 14: 3547–3555.

Seisdedos J. 2006. Los grandes paleo-deslizamientos de Güimar y La Orotava (Tenerife): análisis geológico, mecanismos de inestabilidad y modelización geomecánica. *Boletín Geológico y Minero* 117: 545-550.

Seisdedos J, Ferrer M, González de Vallejo L. 2012. Geological and geomechanical models of the pre-landslide volcanic edifice of Güímar and La Orotava mega-landslides (Tenerife). *Journal of Volcanology and Geothermal Research* 239: 92–110.

Selkoe KA, Toonen RJ. 2006. Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology Letters* 9: 615–629.

Senut B, Pickford M, Ségalen L. 2009. Neogene desertification of Africa. *Comptes Rendus Geoscience* 341: 591–602.

Sepulchre P, Ramstein G, Fluteau F, Schuster M, Tiercelin JJ, Brunet M. 2006. Tectonic uplift and Eastern Africa aridification. *Science* 313: 1419–1423.

Seto KC, Güneralp B, Hutyra, LR. 2012. Global forecasts of urban expansion to 2030 and direct impacts on biodiversity and carbon pools. *Proceedings of the National Academy of Sciences*, 109: 16083–16088.

Silvertown J. 2004. The ghost of competition past in the phylogeny of island endemic plants. *Journal of Ecology* 92: 168–173.

Skrede I, Borgen L, Brochmann C. 2008. Genetic structuring in three closely related circumpolar plant species: AFLP versus microsatellite markers and high-arctic versus arctic–alpine distributions. *Heredity* 102: 293–302.

Smith M, Mosley P. 1993. Crustal heterogeneity and basement influence on the development of the Kenya Rift, East Africa. *Tectonics* 12: 591–606.

Stebbins GL. 1974. *Flowering plants: evolution above the species level*. Belknap, Cambridge, MA.

Stillman C. 1999. Giant Miocene landslides and the evolution of Fuerteventura, Canary Islands. *Journal of Volcanology and Geothermal Research* 94: 89–104.

Sunding P. 1979. Origins of the Macaronesian flora. In: *Plants and islands* (ed. Bramwell D) pp. 13–40. Academic Press, London, UK.

Svenning JC, Fløjgaard C, Marske KA, Nógues-Bravo D, Normand S. 2011. Applications of species distribution modeling to paleobiology. *Quaternary Science Reviews* 30: 2930–2947.

Taberlet P, Fumagalli L, Wust-saucy AG, Cosson JF. 1998. Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology* 7: 453–464.

Takhtajan A. 1969. *Flowering Plants, Origin and Dispersal*. Oliver & Boyd, Edinburgh.

Templeton AR, Crandall KA, Sing CF. 1992. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics* 132: 619–633.

Templeton AR, Sing CF. 1993. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping. IV. Nested analyses with cladogram uncertainty and recombination. *Genetics* 134: 659–669.

Templeton AR. 1998. Nested clade analyses of phylogeographic data: testing hypotheses about gene flow and population history. *Molecular Ecology* 7*:* 381-397.

Templeton AR. 2004. Statistical phylogeography: methods of evaluating and minimizing inference errors. *Molecular Ecology* 13: 789–809.

Thorne J, Kishino H, Painter I. 1998. Estimating the rate of evolution of the rate of molecular evolution. *Molecular Biology and Evolution* 15: 1647–1657.

Thorne JL, Kishino H. 2002. Divergence time and evolutionary rate estimation with multilocus data. *Systematic Biology* 51: 689–702.

Tiffney BH. 1985. The Eocene North Atlantic land bridge: its importance in Tertiary and modern phytogeography of the Northern Hemisphere. *Journal of the Arnold Arboretum* 66: 243–273.

Treviranus GR. 1803. *Biologie oder Philosophie der lebenden Natur für Naturforscher und Aerzte*. Band I, Göttingen.

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

Tzedakis PC. 2009. Museums and cradles of Mediterranean biodiversity. *Journal of Biogeography* 36: 1033–1034.

Urgeles R, Canals M, Baraza J, Alonso B, Masson D. 1997. The most recent megalandslides of the Canary Islands: El Golfo debris avalanche and Canary debris flow, west El Hierro Island. *Journal of Geophysical Research* 102: 20305–20323.

Valido A, Dupont YL, Olesen JM. 2004. Bird-flower interactions in the Macaronesian islands. *Journal of Biogeography* 31: 1945–1953.

Vanderpoorten A, Rumsey F, Carine M. 2007. Does Macaronesia exist? Conflicting signal in the bryophyte and pteridophyte floras. *American Journal of Botany* 94: 625–639.

Vargas P, Valente LM, Blanco-Pastor JL, Liberal I, Guzmán B, Cano E, Forrest A, Fernández-Mazuecos M. 2014. Testing the biogeographical congruence of palaeofloras using molecular phylogenetics: snapdragons and the Madrean-Tethyan flora. *Journal of Biogeography* 41: 932–943.

Vogel S, Westerkamp C, Thiel B, Gessner K. 1984. Ornithophilie auf den Canarischen Inseln. *Plant Systematics and Evolution* 146: 225–248.

Wallace AR. 1864. The origin of human races and the antiquity of man deduced from the theory of "Natural Selection". *Journal of the Anthropological Society of London* 2: CLVIII–CLXXXVII.

Wallace AR. 1876. Lecture on the comparative antiquity of continents, as indicated by the distribution of living and extinct animals. *Proceedings of the Royal Geographical Society of London* 21: 505.

Wallace AR. 1880. *Island life, or, the phenomena and causes of insular faunas and floras*. Harper & Brothers, New York.

Walter T, Troll V, Cailleau B, Belousov A, Schmincke H, Amelung F, Vd Bogaard P. 2005. Rift zone reorganization through flank instability in ocean island volcanoes: an example from Tenerife, Canary Islands. *Bulletin of Volcanology* 67: 281–291.

Wang Q, Zhou SL, Hong DY. 2013. Molecular phylogeny of the platycodonoid group (Campanulaceae s. str.) with special reference to the circumscription of *Codonopsis*. *Taxon* 62: 498-504.

Warren BH, Simberloff D, Ricklefs RE, Aguilée R, Condamine FL, Gravel D,… Thébaud C. 2015. Islands as model systems in ecology and evolution: prospects fifty years after MacArthur-Wilson. *Ecology Letters* 18: 200–217.

Waters JM, Craw D. 2006. Goodbye Gondwana? New Zealand biogeography, geology, and the problem of circularity. *Systematic Biology* 55: 351–356.

Wen J. 1999. Evolution of eastern Asian and Eastern North American disjunct distributions in flowering plants. *Annual Review of Ecology, Evolution, and Systematics* 30: 421–455.

Wen J, Ickert-bond SM. 2009. Evolution of the Madrean-Tethyan disjunctions and the North and South American amphitropical disjunctions in plants. *Journal of Systematics and Evolution* 47: 331–348.

White F. 1978. The afromontane region. In: *Biogeography and ecology of southern Africa* (pp. 463-513). Springer, Netherlands.

White F. 1983. *Vegetation of Africa*. Unesco, Paris.

Whittaker RH. 1961. Vegetation history of the Pacific coast states and the central significance of the Klamath region. *Madrono* 16: 5–23.

Whittaker RH. 1969. New concepts of kingdoms of organisms. *Science* 163: 150–160.

Whittaker R, Fernández-Palacios J. 2007. *Island biogeography: ecology, evolution, and conservation*. Oxford: Oxford University Press.

Wickens GE. 1976. The flora of Jebel Marra (Sudan Republic) and its geographical affinities. *Kew Bulletin Additional Series* 5: 1–368.

Widmer A, Lexer C. 2001. Glacial refugia: sanctuaries for allelic richness, but not for gene diversity. *Trends in Ecology & Evolution* 16: 267–269.

Wiens JJ. 2004. Speciation and ecology revisited: phylogenetic niche conservatism and the origin of species. *Evolution* 58: 193–197.

Wikström N, Savolainen V, Chase MW. 2001. Evolution of the angiosperms: calibrating the family tree. *Proceedings of the Royal Society B: Biological Sciences* 268: 2211–2220.

Wilfert L, Kaib M, Durka W, Brandl R. 2006. Differentiation between populations of a termite in eastern Africa: implications for biogeography. *Journal of Biogeography* 33: 1993–2000.

Woese CR, Kandler O, Wheelis ML. 1990. Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proceedings of the National Academy of Sciences* 87: 4576–4579.

Wolfe KH, Li WH, Sharp PM. 1987. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proceedings of the National Academy of Sciences* 84: 9054–9058.

Wolfenden E, Ebinger C, Yirgu G, Deino A, Ayalew D. 2004. Evolution of the northern Main Ethiopian rift: birth of a triple junction. *Earth and Planetary Science Letters* 224: 213–228.

Wolfenden E, Ebinger C, Yirgu G, Renne PR, Kelley SP. 2005. Evolution of a volcanic rifted margin: Southern Red Sea, Ethiopia. *Geological Society of America Bulletin* 117: 846.

Wondimu T, Gizaw A, Tusiime F, Masao C, Abdi A, Gussarova G, ... Brochmann C. 2014. Crossing barriers in an extremely fragmented system: two case studies in the afro-alpine sky island flora. *Plant systematics and evolution* 300: 415–430.

Wright S. 1931. Evolution in Mendelian populations. *Genetics* 16: 97–159.

Wunderlich J. 1991. Die spinnen-Fauna der Makaronesischen Inseln. Taxonomie, Ökologie, Biogeographie und Evolution. *Beitrage zürAraneologie* 1: 1–619.

Xiang QY, Manchester SR, Thomas DT, Zhang W, Fan C. 2005. Phylogeny, biogeography, and molecular dating of cornelian cherries (*Cornus*, cornaceae): tracking tertiary plant migration. *Evolution* 59: 1685–1700.

Yang Z, Rannala B. 2006. Bayesian estimation of species divergence times under a molecular clock using multiple fossil calibrations with soft bounds. *Molecular Biology and Evolution* 23: 212–226.

Yang Z, Rannala B. 2012. Molecular phylogenetics: principles and practice. *Nature Reviews Genetics* 13: 303–314.

Yannic G, Pellissier L, Ortego J, Lecomte N, Couturier S, Cuyler C,… Côté SD. 2013. Genetic diversity in caribou linked to past and future climate change. *Nature Climate change* 4: 132–137.

Zaczek K, Troll VR, Cachao M, Ferreira J, Deegan FM, Carracedo JC, Soler V, Meade F, Burchardt S. 2015. Nannofossils in 2011 El Hierro eruptive products reinstate plume model for Canary Islands. *Scientific Reports* 5: 7945, doi:10.1038/srep07945.

Zhengyi W. 1983. On the significance of Pacific intercontinental discontinuity. *Annals of the Missouri Botanical Garden* 70: 577–590.

Zhou Z, Hong D, Niu Y, Li G, Nie Z, Wen J, Sun H. 2013. Phylogenetic and biogeographic analyses of the Sino-Himalayan endemic genus *Cyananthus*

(Campanulaceae) and implications for the evolution of its sexual system. *Molecular Phylogenetics and Evolution* 68: 482–497.

Zuckerkandl E, Pauling L. 1962. Molecular disease, evolution and genetic heterogeneity In: *Horizons in Biochemistry* (eds. Kasha M, Pullman B) pp. 189–225. Academic Press, New York and London.

Zuckerkandl E, Pauling L. 1965. Molecules as documents of evolutionary history. *Journal of Theoretical Biology* 8: 357–366.

IX. List of papers included in this phD

(*=corresponding author)

*__Mario Mairal__, Lisa Pokorny, Juan José Aldasoro, Marisa Alarcón, Isabel Sanmartín. **2015**. Ancient vicariance and climate-driven extinction explain continental-wide disjunctions in Africa: the case of the Rand Flora genus *Canarina* (Campanulaceae). ***Molecular Ecology*** 24: 1335–1354. Evolutionary Biology, **Q1** (5/46), **IF=6.494**. https://onlinelibrary.wiley.com/doi/abs/10.1111/mec.131144

*__Mario Mairal__, Isabel Sanmartín, Juan José Aldasoro, Ioanna Manolopoulou, Victoria Culshaw, Marisa Alarcón. **2015**. Palaeo-islands as refugia and cradles of diversity in volcanic archipelagos: The case of the widespread endemic *Canarina Canariensis*. ***Molecular Ecology*** 24: 3944–3963. Evolutionary Biology, **Q1** (5/46), **IF=6.494**. https://onlinelibrary.wiley.com/doi/abs/10.1111/mec.13282

*__Mario Mairal__, Juan José Aldasoro, Isabel Sanmartín, Pablo Vargas, Marisa Alarcón. **2017**. Geographic barriers and Pleistocene climate change shaped patterns of genetic variation in the Eastern Afromontane biodiversity hotspot. ***Scientific Reports*, 7**. Multidisciplinary Sciences, **Q1** (7/61), **IF=5.228**. https://www.nature.com/articles/srep45749

*__Mario Mairal__, Isabel Sanmartín, Loïc Pellissier. **2017**. Lineage-specific climatic niche drives the tempo of vicariance in the Rand Flora. ***Journal of Biogeography*** 44: 911–923. Geography, physical **Q1** (5/49), **IF = 4.154**. https://onlinelibrary.wiley.com/doi/full/10.1111/jbi.12930

*__Mario Mairal__, Juli Caujapé, Loïc Pellissier, Myriam Heuertz, Ruth Jaén Molina, Nadir Álvarez, Isabel Sanmartín. **2018**. A tale of two forests: ongoing aridification drives population decline and genetic diversity loss at continental scale in Afro-Macaronesian evergreen-forest archipelago endemics. ***Annals of Botany*** 122: 1005–1017. Plant Sciences **Q1** (22/211), **IF=4.04**. https://academic.oup.com/aob/advance-article/doi/10.1093/aob/mcy107/5037756

RESUMEN

Lisa Pokorny, Ricarda Riina, **Mario Mairal**, Victoria Culshaw, Andrea Sánchez-Meseguer, Miguel Serrano, Rodrigo Carbajal, Santiago Ortiz, Myriam Heuertz, Isabel Sanmartín. **2015**. Living on the edge: timing of Rand Flora pattern disjunctions compatible with ongoing aridification in Africa. ***Frontiers in Genetics*** 6: 1–15. Genetics, **Q1**, **IF=4.08**. https://www.frontiersin.org/articles/10.3389/fgene.2015.00154/full

Mario Mairal**, Andrea Sánchez-Meseguer. **2012**. Resolviendo una incógnita biogeográfica, el caso de la Rand Flora Afro-mediterránea. ***Chronica naturae 2: 53–63. http://www.hombreyterritorio.org/chronica_naturae/num2/archivos/chronicanaturae2_53_2012.pdf

Isabel Sanmartín, Lisa Pokorny & **Mario Mairal**. **2016**. Origen evolutivo de una enigmática distribución geográfica: La “Rand Flora” Africana. *Investigación y ciencia* (***Scientific American***).

Presentation of the research of this dissertation at scientific conferences

Date	Conference	Place	Authors	Title	Type
3 – 4 February 2011	VIII Annual Ecoflor Meeting	Toledo, Spain	**Mario Mairal**, Isabel Sanmartín, Juan José Aldasoro, Marisa Alarcón	*Canarina* phylogeographic study, a genus of Afro-Macaronesian distribution	Oral (Mario Mairal)
17 October 2011	V RJB-CSIC Scientific Marathon	Madrid, Spain	**Mario Mairal**, Juan José Aldasoro, Marisa Alarcón, Isabel Sanmartín	A biogeographical mystery: the disjunction of Afro-Macaronesian genus *Canarina* (Campanulaceae)	Oral (Mario Mairal) *Awarded*
5 – 8 September 2012	FloraMac 2012	Funchal, Madeira, Portugal	**Mario Mairal,** Juan José Aldasoro, Marisa Alarcón, Isabel Sanmartín	Long-distance dispersal or large-scale extinction? Explaining the extreme Rand Flora disjunction in the bellflower genus *Canarina*	Poster
28 – 30 November 2012	II UNESCO Chair Meeting for Plant Biodiversity Conservation in Macaronesia and Western Africa.	Las Palmas de Gran Canaria, Canary Islands, Spain	**Mario Mairal**, Juli Caujapé, Ruth Jaén, Marisa Alarcón, Juan José Aldasoro, Isabel Sanmartín	Solving the mystery of the Rand Flora pattern: islands in the sky 'vs'. oceanic islands	Oral (Mario Mairal)
9 – 13 January 2013	VI Biennial Meeting of the International Biogeography Society	Miami, FL, USA	**Mario Mairal**, Lisa Pokorny, Ruth Jaén, Juli Caujapé, Isabel Sanmartín	Long-distance dispersal or large-scale extinction? Explaining the extreme Rand Flora disjunction in the bellflower genus *Canarina*	Poster
31 January – 1 February 2013	X Annual Ecoflor Meeting.	Sevilla, Spain	**Mario Mairal**, Juan José Aldasoro, Isabel Sanmartín, Pablo Vargas, Marisa Alarcón	Genetic structure of *Canarina eminii*, an epiphytic species of the East African *Sky island* forests	Oral (Marisa Alarcón)
25 – 27 September 2013	II Iberian Congress of Biological Systematics (CISA 2013).	Barcelona, Spain	**Mario Mairal**, Ruth Jaén, Juli Caujapé, Isabel Sanmartín.	Plant evolution in island systems: population study of oceanic islands and continental *sky islands* in the genus *Canarina* (Campanulaceae)	Oral (Mario Mairal) *Awarded*
15 November 2013	VII RJB-CSIC Scientific Marathon	Madrid, Spain	**Mario Mairal**, Juan José Aldasoro, Marisa Alarcón, Isabel Sanmartín	Historia de dos islas: extinción, dispersión y cambio climático en el género Afro-Macaronésico *Canarina*	Oral (Mario Mairal)
27 – 29 November 2013	IV SESBE Meeting	Barcelona, Spain	**Mario Mairal**, Lisa Pokorny, Ruth Jaén, Juli Caujapé, and Isabel Sanmartín.	Disentangling the role of stepping-stone dispersal and large-scale extinction on wide ranging disjunctions: a case study on African *Canarina* flowers.	Poster
20 – 24 June 2014	Evolution	North Carolina, USA	Lisa Pokorny, Ricarda Riina, **Mario Mairal**, Andrea Briega, Jon Cendoya, Isabel Sanmartín	The role of extinction in the assembly of large-scale biodiversity patterns	Oral (Lisa Pokorny)
28 November 2014	VIII RJB-CSIC Scientific Marathon	Madrid, Spain	**Mario Mairal**, Juan José Aldasoro, Marisa	Macroevolution and microevolution in island systems: the Rand Flora pattern in the Afro-Macaronesian genus *Canarina*	Oral (Mario Mairal)

			Alarcón, Isabel Sanmartín		
24 – 27 March 2015	FloraMac 2015	Las Palmas de Gran Canaria, Canary Islands, Spain	**Mario Mairal**, Marisa Alarcón, Juan José Aldasoro, Lisa Pokorny, Isabel Sanmartín	A long history of climate-driven extinction has led Rand Flora lineages to find refugium in Macaronesia: a deeper look at the Tenerife paleo-islands	Oral (Mario Mairal)

CAPÍTULOS /

CHAPTERS

CHAPTER 1

Ancient vicariance and climate-driven extinction explain continental-wide disjunctions in Africa: the case of the Rand Flora genus *Canarina* (Campanulaceae).

Mario Mairal, Lisa Pokorny, Juan José Aldasoro, Marisa Alarcón,
Isabel Sanmartín

This chapter has been published in the peer-review journal
Molecular Ecology (2015) 24: 1335-1354.
DOI: 10.1111/mec.13114

MOLECULAR ECOLOGY

Molecular Ecology (2015) **24**, 1335–1354 doi: 10.1111/mec.13114

Ancient vicariance and climate-driven extinction explain continental-wide disjunctions in Africa: the case of the Rand Flora genus *Canarina* (Campanulaceae)

M. MAIRAL,* L. POKORNY,* J. J. ALDASORO,† M. ALARCÓN† and I. SANMARTÍN*
*Real Jardín Botánico (RJB-CSIC), 28014, Madrid, Spain, †Institut Botànic de Barcelona (IBB-CSIC), 08038 Barcelona, Spain

Abstract

Transoceanic distributions have attracted the interest of scientists for centuries. Less attention has been paid to the evolutionary origins of 'continent-wide' disjunctions, in which related taxa are distributed across isolated regions within the same continent. A prime example is the 'Rand Flora' pattern, which shows sister taxa disjunctly distributed in the continental margins of Africa. Here, we explore the evolutionary origins of this pattern using the genus *Canarina*, with three species: *C. canariensis*, associated with the Canarian laurisilva, and *C. eminii* and *C. abyssinica*, endemic to the Afromontane region in East Africa, as case study. We infer phylogenetic relationships, divergence times and the history of migration events within *Canarina* using Bayesian inference on a large sample of chloroplast and nuclear sequences. Ecological niche modelling was employed to infer the climatic niche of *Canarina* through time. Dating was performed with a novel *nested* approach to solve the problem of using deep time calibration points within a molecular dataset comprising both above-species and population-level sampling. Results show *C. abyssinica* as sister to a clade formed by disjunct *C. eminii* and *C. canariensis*. Miocene divergences were inferred among species, whereas infraspecific divergences fell within the Pleistocene–Holocene periods. Although *C. eminii* and *C. canariensis* showed a strong genetic geographic structure, among-population divergences were older in the former than in the latter. Our results suggest that *Canarina* originated in East Africa and later migrated across North Africa, with vicariance and aridification-driven extinction explaining the 7000 km/7 million year divergence between the Canarian and East African endemics.

Keywords: Bayesian biogeography, climate-driven extinction, continental islands, long-distance dispersal, *nested* phylogenetic dating, vicariance

Received 22 October 2014; revision received 10 February 2015; accepted 11 February 2015

Introduction

Transoceanic disjunct distributions have long attracted the attention of biogeographers (von Humboldt & Bonpland 1805; Hooker 1867; Raven & Axelrod 1972; Donoghue & Smith 2004). A prime example is the Gondwanan distribution exhibited by groups like ratites or marsupials, in which sister lineages are scattered across continents now isolated by thousands of kilometres of oceanic waters (Treviranus 1803; Hooker 1853). Fragmentation of an ancient widespread distribution by plate tectonics (vicariance) and long-distance dispersal events have alternatively been postulated to explain this pattern (Givnish & Renner 2004; Sanmartín & Ronquist 2004).

In contrast, less attention has been paid to the evolutionary origins of 'continent-wide' disjunctions, in which related taxa are distributed across geographically isolated regions within the same continent. Transoceanic disjunctions are explained either by tectonic-induced vicariance (i.e. continental drift) followed by biotic division (Raven & Axelrod 1972; Sanmartín *et al.* 2001) or by LDD (Renner 2004). Within-continent disjunctions, on the other hand, can be explained by LDD

Correspondence: Mario Mairal and Isabel Sanmartín, Fax: +34914200157; E-mails: mariomairal@gmail.com; isanmartin@rjb.csic.es

(Coleman *et al.* 2003; Pelser *et al.* 2012) but are often attributed to large-scale climatic events, such as global climate cooling or aridification that would have extirpated a once continuous biota from part of its area of distribution, leaving relict taxa in refugia or 'continental islands' (Axelrod & Raven 1978; Wiens & Donoghue 2004; Crisp & Cook 2007). The barrier that caused the range division in this case is not the opening of an ocean basin, but an environmental change that creates stretches of inhospitable land that are outside the climatic tolerances of the organism (Wiens & Donoghue 2004). Within-continent disjunctions are thus interesting to explore the role of climate-driven extinction in the assembly of biodiversity patterns (Linder 2014).

A prime example of this type of disjunction is the African 'Rand Flora' pattern (from the German word meaning *rim*, aka 'flora from the edge'), in which distantly related plant families show a similar disjunct distribution, with sister taxa inhabiting geographically isolated regions in the continental margins of Africa—that is north-west Africa, Horn of Africa–southern Arabia, eastern Africa and Southern Africa, and adjacent islands, Macaronesia, Socotra, Madagascar (Christ 1892; Engler 1910; Lebrun 1961; Quézel 1978; Andrus *et al.* 2004; Sanmartín *et al.* 2010 for a historical review). Although they differ in aspects such as morphology, habit or phenology, Rand Flora lineages share some degree of adaptation to subtropical or temperate environments, so that the tropical lowlands of central Africa or the arid terrains of the Sahara and Sino-Arabic Deserts in the north and the Kalahari desert in the south constitute for them effective climatic barriers to dispersal. Traditionally, this pattern has been explained by vicariance, the fragmentation of an ancient widespread African flora by aridification events during the Late Neogene, leaving relict taxa that survived and diversified in 'climatic refuges' at the margins of the continent (Axelrod & Raven 1978; Bramwell 1985). However, the advent of molecular phylogenetics and the possibility of obtaining estimates of divergence times have shown that, for some lineages, these disjunctions can be better explained in terms of recent independent dispersal events among the Rand Flora regions, followed by *in situ* diversification (Fiz *et al.* 2008; Meseguer *et al.* 2013). Because continental disjunct patterns such as the Rand Flora are explained by the appearance of a climatic barrier that causes range division (e.g. the formation of the Sahara desert in the Late Miocene), ecological niche modelling techniques (ENMs) might also be useful to examine the evolutionary origins of Rand Flora lineages. By reconstructing the potential climatic niche of a species and projecting it backwards in time, we can identify areas that were in the past within the organism's range of climatic tolerances but are inhospitable today due to large-scale climate change (Yesson & Culham 2006; Smith & Donoghue 2010; Meseguer *et al.* 2014).

One of the strongest connections within the Rand Flora pattern links the Macaronesian Islands to East Africa. Genera such as *Camptoloma* (Kornhall *et al.* 2001), *Aeonium* (Mort *et al.* 2002), *Campylanthus* (Thiv *et al.* 2010) or *Euphorbia* (Riina *et al.* 2013) harbour Macaronesian endemics, whose sister groups are found along eastern Africa and southern Arabia. In a recent meta-analysis of Rand Flora lineages, Sanmartín *et al.* (2010) found a comparatively high rate of historical dispersal between these two regions (i.e. NW Africa vs. E Africa/S Arabia), suggesting a long history of biotic connections across the Sahara. Here, we focus on one of the most striking examples of this disjunction, which has never been studied before. The bellflower genus *Canarina* (family Campanulaceae) is a small angiosperm genus of three species, one endemic to the Canary Islands (*Canarina canariensis* (L.) Vatke. (1874)) and two other distributed exclusively in the montane regions of eastern Africa: *Canarina eminii* Aschers. ex Schweinf. (1892) and *Canarina abyssinica* Engl. (1902). *Canarina canariensis* is associated with the Canarian laurisilva, the highly endemic laurel forest present in the western and central Canary Islands. *Canarina eminii* is an epiphyte endemic to the forests belts of the Afromontane region, while *C. abyssinica* occurs in the upland open forests of eastern Africa (Fig. 1; see Supplementary Text 'Study Group' for a more detailed description of the morphology, biology and geographic distribution of each species). Both the Canarian laurisilva and the Afromontane region – a series of isolated areas forming an archipelago-like centre of endemism in the mountains of East and West Africa (White 1983) – are traditionally considered as examples of the *refugium–fragmentation* theory: the remnants of a subtropical flora that once was widespread through Africa but became later extinct due to climatic aridification events (Axelrod & Raven 1978; Bramwell 1985). Therefore, *Canarina* represents not only a wide continental disjunction of nearly 7000 km across the Sahara, but also a potential relict of an 'ancient pan-African flora' (Axelrod & Raven 1978) and a prime candidate to test the climatic vicariance theory in the origins of the Rand Flora pattern. Moreover, the particular distribution of *Canarina* in the Canary Islands and in the fragmented Afromontane forests offers us a unique opportunity to study patterns of colonization in *true* 'oceanic islands' vs. ecological 'mountain islands' (aka 'sky islands', McCormack *et al.* 2009). The high-altitude mountain regions in the Afromontane region of East Africa have often been equated to ecological islands (Hedberg 1961; Popp *et al.* 2008; McCormack

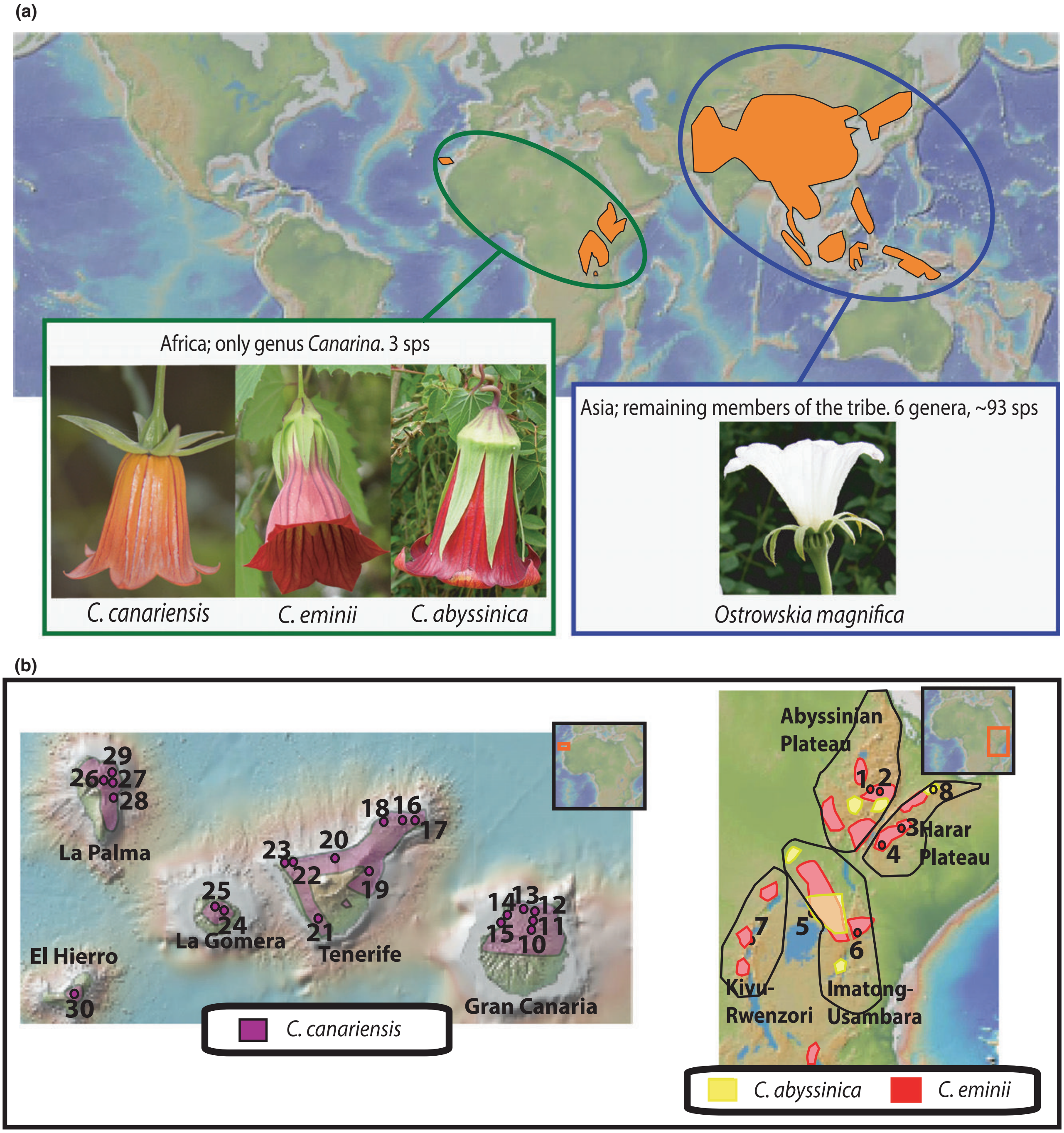


Fig. 1 (a) Worldwide distribution of tribe Platycodoneae (Campanulaceae) shows the geographic disjunction between the single African genus (*Canarina*) and the remaining members of the tribe, which are endemic to the mountains of Asia. (b) Geographic distribution of the three species of *Canarina*; the distribution of the East African species, *Canarina eminii* and *Canarina abyssinica* has been modified from Hedberg (1961). Numbers correspond to the sampled populations, with codes given in Table S1. Maps have been modified from GeoMapApp (Ryan *et al.* 2009; www.geoMapApp.org).

et al. 2009), isolated from one another by stretches of dry land or tropical lowlands.

Canarina belongs to tribe Platycodoneae, a basal lineage within family Campanulaceae (Roquet *et al.* 2009) that includes six other genera endemic to the mountains of central and eastern Asia. Recent molecular studies have reconstructed the phylogeny of the tribe (Wang *et al.* 2013; Zhou *et al.* 2013), but a phylogeny of *Canarina* is still missing due to the difficulty obtaining material for the East African endemic species poses. Here, we present the first species-level phylogeny of the genus using evidence from the nuclear ribosomal (nrDNA) ITS region and seven noncoding chloroplast (cpDNA) markers and a large sample of infraspecific sequences – covering the entire geographic range of *C. eminii* and *C. canariensis* – as well as a representative sample of genera within Platycodoneae. Bayesian inference methods were used to estimate lineage divergence times and to infer ancestral ranges and the main history of migration events within *Canarina*. Lack of fossils in plant phylogenetic studies often results in deep calibration points being applied to a broader data set, which sometimes includes both above-species and below-species level sampling (Blanco-Pastor *et al.* 2013; Nolasco-Soto *et al.* 2014). This can result in uncertain or even biased time estimates as we move from species level to coalescent dynamics (Ho *et al.* 2008, 2011). Here, we used a novel 'nested-dating' Bayesian approach to accommodate the expected change in molecular rates and tree growth model resulting from heterogeneous species-population sampling schemes. Finally, we used ENMs and paleoenvironmental data to estimate the climatic niche of *Canarina*, which when projected backward in time allowed us to detect climatically suitable areas that might have formed part of its geographic distribution in the past or acted as climatic dispersal corridors.

Materials and methods

Taxon sampling and DNA sequencing

Throughout several field campaigns in eastern Africa and the Canary Islands (2009–2012), fresh material for 29 individuals, representing different populations within *Canarina*, was collected and included in the analysis (Table S1, Supporting information): one sample of *Canarina abyssinica* (from the Ethiopian Highlands), seven samples of *Canarina eminii* and 21 of *Canarina canariensis*, covering the entire range of distribution of the last two species. The low number of samples in *C. abyssinica* reflects the difficulty to collect this species, which has apparently disappeared from many of the original localities where it was first described by Hedberg (1961; see Data S1 for a description on the sampling effort and current conservation status of *C. abyssinica*). Nine species representing additional genera within tribe Platycodoneae (*Campanumoea*, *Codonopsis*, *Cyananthus*, *Cyclocodon*, *Ostrowskia* and *Platycodon*), as well as related tribe Campanuleae (*Campanula* L.) and Campanulaceae subfamilies Lobelioideae (*Lobelia* L.) and Cyphioideae (*Cyphia* P.J. Bergius) were used as alternative outgroups in the phylogenetic and biogeographic analyses. In all, 256 sequences were generated for this study and 11 downloaded from GenBank. Species names, voucher information and GenBank accession numbers for all sequences are provided in Table S1 (Supporting information).

We selected seven noncoding plastid regions exhibiting high levels of genetic variation, the intergenic spacers *psb*J–*pet*A, *rpl*32–*trn*L, *trn*L–*trn*F, *trn*S–*trn*G and 3′*trn*V–*ndh*C (Shaw *et al.* 2005, 2007) and the *trn*G and the *pet*D group II introns (*pet*B–*pet*D, Borsch *et al.* 2009). Details on PCR amplification and sequence editing and alignment are given in Data S1 and Table S2 (Supporting information). Two data sets were constructed to address different objectives. The '*Platycodoneae* data set' ($n = 12$) included samples of all aforementioned outgroup genera plus one accession of each *Canarina* species and was used to reconstruct the phylogeny of the tribe and provide additional calibration points in the dating analyses. The '*Canarina* data set' ($n = 29$) included one accession of each population sampled within *Canarina*, plus one sequence of genera *Ostrowskia* and *Cyclocodon*, which were identified in a prior Campanulaceae study as closely related to *Canarina* (Mansion *et al.* 2012). This latter data set was used to infer the population and phylogeographic history of *C. canariensis* and *C. eminii*.

Phylogenetic inference

Phylogenetic relationships were estimated for each marker separately using Bayesian inference implemented in MrBayes (Ronquist *et al.* 2012). Additional analyses were run using maximum likelihood implemented in the software RAxML (Stamatakis *et al.* 2008). The *Platycodoneae* data set was rooted using *Lobelia* as the outgroup taxon; the *Canarina* data set was rooted using *Ostrowskia* as the outgroup, except for the interspacer 3′*trn*V-*ndh*C and the *trn*G intron for which *Ostrowskia* and *Cyclocodon* sequences were missing, in which case we used *Platycodon*. Details on these analyses are provided in the Data S1.

Before concatenating the genes into a combined data set, we checked for topological congruence in the inferred relationships by examining the Bayesian consensus trees and searching for well-supported clades ($PP > 0.95$) in the consensus tree of one marker that were not present in the consensus trees of the other markers (Antonelli & Sanmartín 2011). All analysed genes recovered similar phylogenetic relationships at the generic level, but significant incongruence was found in species relationships within *Canarina* for both the *Platycodoneae* (Fig. S1, Supporting information) and the *Canarina* data sets (Fig. S2, Supporting information) for the plastid markers. Three cpDNA genes (*rpl*32, *trn*SG and *trn*V-*ndh*C) grouped *C. eminii* with *C. abyssinica* with *C. canariensis* as their sister group, while the rest of markers either failed to resolve relationships (*trn*G2G) or placed *C. abyssinica* as sister to a clade formed by *C. canariensis* and East African *C. eminii* (*pet*BD, *trn*LF, *psb*J-*pet*A). The latter relationship was also recovered by the single nuclear marker ITS (Figs S1–S2, Supporting information). The same relationships were also obtained using ML although with lower support values (Fig. S2, Supporting information).

Incongruent relationships between gene trees can be attributed to different phenomena, including paralogy, concerted evolution, incomplete lineage sorting (ILS), homoplasy or noise resulting from substitutional saturation or PCR artefacts. Paralogy and concerted evolution are not expected in plastid markers since, unlike multiple-copy nuclear markers like ITS, cpDNA genes are thought to be single copy and behave as a single, linked genome. Multispecies coalescent models (Heled & Drummond 2010) can address ILS but require infraspecific sampling for each species, whereas we only had one sequence for *C. abyssinica* and all outgroup genera. Instead, we used BUCKy (Larget *et al.* 2010) to estimate the Bayesian support for alternative topologies among different genes when analysed in a concatenated data set. BUCKy makes no assumption about the reason for discordance among gene trees but instead estimates the dominant history of sampled individuals and how much of the genome supports each relationship, using Bayesian concordance analysis. Groups of genes sharing the same tree are detected (while accounting for uncertainty in gene tree estimates) and then combined to gain more resolution on their common tree (Ané *et al.* 2007; Larget *et al.* 2010). Using BUCKy with default settings (α = infinity, allowing genes to evolve independently) showed that inclusion of *rpl*32 in a concatenated cpDNA *Platycodoneae* data set was responsible for significant topological changes in the phylogeny of *Canarina*, but that this was not the case with other incongruent markers such as *trn*SG, which consistently grouped *C. eminii* with *C. canariensis*, and *C. abyssinica* as their sister species (Table S3, Supporting information).

Noncoding intergenic spacer regions, such as *rpl*32-*trn*L, have become very popular for solving relationships at low taxonomic levels because of their high sequence variability (Shaw *et al.* 2007), but recent studies have pointed out that this variability is not necessarily correlated with phylogenetic usefulness and can lead to higher levels of homoplasy (Korotkova *et al.* 2011). To test whether higher levels of homoplasy and substitutional saturation might explain topology differences among cpDNA genes, we plotted uncorrected pairwise distances against maximum-likelihood distances among sequences estimated in PAUP* v4.0b10 (Swofford 2002) and looked for deviation from linearity in saturation plots (Fig. S3, Supporting information). All plots showed a strong fit to a linear regression but *rpl*32 showed slight levels of saturation at the deepest divergences (Fig. S3, Supporting information). Furthermore, a MrBayes analysis of a cpDNA concatenate data set of *Canarina* estimated gene-specific rate multiplier that was four times higher in *rpl*32 than in any other region (Table S3, Supporting information), while the total tree length was two times higher in *rpl*32 (TL = 1.374) compared with other plastid markers (*trn*SG: TL = 0.668; *pet*BD: TL = 0.448, Table S3, Supporting information), suggesting faster higher mutation rates. These phenomena were not observed in *trn*V or *trn*SG, which showed rate multipliers and tree-length estimates similar to *pet*BD (Table S3, Supporting information). Moreover, plastid *rpl*32 also exhibited the largest proportion of indels in relation to substitutions than any other marker (35.06%; Table 1).

Given this possible level of homoplasy, we decided to exclude *rpl*32 from further analyses. Additionally, we excluded *trn*G2G because of lack of variability (Figs S1 and S2, Supporting information), and the *3′trn*V-*ndh*C interspacer because it showed slight levels of saturation (Fig. S3, Supporting information) and we lacked sequences for all outgroup taxa except *Platycodon* (Figs S1 and S2, Supporting information); it has been shown that outgroup composition can have a strong influence on the ingroup topology and support values (Rothfels *et al.* 2012). On the other hand, we kept the *trn*SG gene in our analyses because – although it supported the same species topology as *rpl32* – it did not show evidence of saturation or accelerated substitution rates like the latter marker (Table 1, Fig. S3, Supporting information). Therefore, for the final analyses of the *Canarina* data sets, we concatenated the four regions, *psb*J-*pet*A, *pet*B-*pet*D, *trn*L-*trn*F and *trn*S-*trn*G into a combined cpDNA matrix, which was analysed in conjunction with the nuclear ITS, as the latter marker supported the same topology as the combined cpDNA data set and no evidence of multiple copies were found. The concatenated data matrix was analysed under the GTR + G model, partitioned by gene and allowing the overall mutation rate to differ among partitions using the MrBayes command *prset rate = variable*.

Divergence time estimation

Lineage divergence times were estimated using the Bayesian relaxed-clock models implemented in BEAST v.1.7 (Drummond & Rambaut 2007). Choice of model priors was based on Bayes factor comparisons using the path sampling (PS) and stepping stone (SS) sampling methods in BEAST, which have been shown to

Table 1 Summary statistics of the chloroplast and nuclear regions analysed here for the *Canarina* data set (no outgroups). Fragment length is given in base pairs (bp); alignment length includes the indels

	*rpl*32-*trn*L	3′*trn*V-*ndh*C	*psb*J-*pet*A	*pet*B-*pet*D	*Trn*L-*trn*F	*Trn*S-*trn*G	*trn*G intron	ITS
Fragment length	581–647	756–855	822–840	753–798	683–918	666–688	661–676	604–734
Alignment length	647	855	841	808	933	690	677	734
Constant sites	590	822	814	766	902	666	666	648
Variable sites	57	33	27	42	31	24	11	86
Indel (%)	35.06	–	20.13	11.2	8.13	6.98	–	10.28

outperform the harmonic mean estimator in terms of consistency and reduced variance (Baele *et al.* 2012). The Yule model and the uncorrelated lognormal distribution (UCLD) were selected, respectively, as the tree and clock model priors for all the analyses (Table S4, Supporting information). Two MCMCs were run for 20 million generations, sampling every 1000th generation. We used Tracer v.1.6 (Rambaut *et al.* 2013) to monitor convergence and EES values (>200) for all parameters, and TreeAnnotator v. 1.7 (Rambaut & Drummond 2013) to construct a maximum clade credibility tree from the posterior distribution after discarding 10% samples as burn-in.

There are no known fossils of *Canarina*, so we relied on two approaches to estimate lineage divergence times. First, we used a standard 'secondary calibration approach' in which a more inclusive, higher-level data set is used to estimate divergence times within the ingroup. We estimated divergence times among Platycodoneae genera using the cpDNA data set with a GTR + G model (we did not include ITS to avoid potential artefacts derived from simultaneously dating plastid and nuclear genomes, which might have very different divergence rates at this level, see Wolfe *et al.* 1987). We used a uniform prior for the *ucld.mean* within values commonly observed in plant plastid markers (10^{-4}–10^{-1} substitutions/site/Ma, Wolfe *et al.* 1987) and a default exponential prior for the standard deviation (SD). As calibration points, we used secondary age constraints drawn from the fossil-rich, angiosperm-wide phylogenetic analysis of Bell *et al.* (2010). The split between *Lobelia* and Campanulaceae was calibrated using a normal distribution spanning the confidence interval in the aforementioned study (mean = 56 Ma, SD = 7.5, 95% high posterior density (HPD) = 41–67 Ma), whereas the split between Campanuleae (*Campanula*) and Platycodoneae was set to mean = 43 Ma (SD = 8, 95% HPD = 28–56 Ma). The ages estimated in this analysis were used to calibrate two nodes in the *Canarina* data set: the divergence between *Cyclocodon* and *Ostrowskia* (mean = 20.83 Ma, SD = 6.0) and the divergence between *Canarina* and *Ostrowskia* (mean = 13.7 Ma, SD = 3.5). The cpDNA + ITS data set was used for this analysis, because at this level differences in mutation rates are minor. Although BEAST selected the UCLD prior (Table S4, Supporting information), Tracer revealed poor mixing and low EES values for the *ucld.mean* and *ucld.stdev* parameters, which did not improve after increasing the run length. We thus used the model with the next lowest marginal likelihood, a Yule strict clock model, for the analysis. The mean clock rate was assigned a broad uniform distribution prior (10^{-6}–10^{-1}), with default prior settings for the rest of parameters.

Heterogeneous molecular data sets spanning both species- and population-level sampling such as the *Canarina* data set pose a set of problems in the estimation of lineage divergence times. First, there is the need to apply just one tree prior to the entire phylogeny, from the older deep-time branches to the younger infraspecific events towards the tips. A stochastic branching prior like Yule is likely to overestimate the date of the most recent divergence events, as for short time scales genetic divergence may precede species divergence (Ho *et al.* 2011), and the opposite effect is expected for coalescent demographic priors. Multispecies coalescent models such as those implemented in *BEAST (Heled & Drummond 2010) can address this problem but require infraspecific sampling for each species, whereas we only had one sequence for *C. abyssinica* and each outgroup genus. Second, Ho *et al.* (2005) demonstrated that when deep-time calibration points are used in heterogeneous species-population sampling schemes, extrapolation of molecular rates across the species-population boundary might yield biased estimates of the rate of molecular variation. In our case, the root and stem nodes in the *Canarina* data set are both constrained with deep-time calibration events (>10 Ma). One consequence of this is the need to use 'all-encompassing' priors for the mean clock rate (e.g. Blanco-Pastor *et al.* 2013; Nolasco-Soto *et al.* 2014) that accommodate the expected change as we move from the slow, long-term substitution rates at the base of the tree (above-species level) to the rapid mutation rates towards the tips (infraspecific sampling), which might result in uncertain time estimates with broad confidence intervals.

To solve this problem, we used here a *nested-dating* partitioned approach – first proposed by Pokorny *et al.* (2011) – in which a higher-level data set calibrated with external evidence (the *Platycodoneae* data set) is used to constrain the molecular clock rate of additional linked data sets containing population-level data. For this, we constructed two data sets containing all accessions of ITS and plastid markers (*pet*BD, *psb*J, *trn*LF and *trn*SG) for every sampled population within *Canarina eminii* (*n* = 7) and *C. canariensis* (*n* = 21). These two data sets were not constrained by any calibration point, but their molecular clock was drawn from the mutation rate of the higher-level *Platycodoneae* partition, that is the 'clock model' was linked across partitions and assigned a UCLD prior. The 'tree model' was unlinked to accommodate the fact that not all markers and taxa were represented equally across partitions, that is the *Platycodoneae* data set included only data for the plastid markers and one accession each within *C. canariensis* and *C. eminii*. This allowed us to assign a branching Yule tree prior to the above-species level (*Platycodoneae*) partition and a coalescent constant-size prior to the infraspecific *Canarina* partitions, the latter selected by Bayes factor PS and SS comparisons.

Ancestral area reconstruction

The Bayesian discrete phylogeographic approach of Lemey *et al.* (2009), implemented in BEAST v.1.7, was used to infer ancestral ranges and trace the history of migration events across space and time in *Canarina*. This is a continuous-time Markov chain (CTMC) model with the discrete states being the areas or geographic locations of the sequences and the transition rates between states and the migration rates between areas (Sanmartín *et al.* 2008). Bayesian MCMC inference is used to estimate simultaneously the posterior distribution of phylogenetic relationships, branch lengths and geographic ancestral states, while accounting for uncertainty in all of these parameters, including the estimation of ancestral frequencies for the root (Lemey *et al.* 2009). Migration rates between areas and the geodispersal rate scalar μ were modelled using default gamma prior distributions (Lemey *et al.* 2009). Two replicate searches of 20 million generations each, sampling every 1000th generation, were combined in TreeAnnotator, after removing the 10% burn-in, to produce a maximum clade credibility (MCC) tree. Bayesian stochastic variable selection (BSVS, Lemey *et al.* 2009) was used to infer the migration events that are best supported by the data. We run two different analyses. To reconstruct the biogeographic history of the genus, we used the *Canarina* data set with identical settings to the 'secondary calibration' dating analysis and four discrete areas: East Asia, central Asia, East Africa and Canary Islands. To reconstruct phylogeographic patterns within *C. eminii* and *C. canariensis*, we used the population-level data sets and a constant-size coalescent model, with the root node in each analysis calibrated with the divergence time estimates obtained from the nested analysis, and a finer-scale definition of areas (Fig. 1b). For *C. canariensis*, six discrete areas were defined corresponding to the islands in the Canarian Archipelago where the species is present: Gran Canaria (GC), La Gomera (GO), La Palma, and El Hierro (EH) and Tenerife, with the latter divided into two areas: eastern Tenerife (TFE) and western Tenerife (TFW), following previous phylogeographic studies pointing out to an east–west division within the island (Juan *et al.* 2000). For *C. eminii*, we divided the montane regions of eastern Africa following the criterion of Gehrke & Linder (2014), except that we subdivided the Ethiopian plateaus into north-west and south-east Ethiopia as several studies have shown phylogeographic disjunctions across the Ethiopian Rift (e.g. Assefa *et al.* 2007; Wondimu *et al.* 2014). In all, we have defined four areas, whose limits are shown in Fig. 1a,b): the Abyssinian plateau (the highlands located west of the Ethiopian Rift), Harar plateau (highlands east of the Ethiopian Rift), Imatong–Usambara (scattered 'sky islands' from south Sudan to Tanzania) and Kivu–Rwenzori (northern part of the Albertine Rift). We also ran an additional analysis in which each plateau and sky island has been considered as an independent region (areas = 5).

Ecological niche modelling

To understand whether the wide geographic disjunct distribution in *Canarina* might have been caused by environmental change, we constructed a species distribution model for the genus, using extant occurrence data from two species at the western and eastern side of the disjunction for which we had enough data. In all, we used 122 records: 67 for *C. canariensis* and 54 for *C. eminii* (Table S5, Supporting information), covering the entire distributional range of these two species. Data points were obtained from published monographs and inventories (Hedberg 1961; Fernández-López 2014), online databases (www.jardincanario.org/flora-de-gran-canaria; www.gbif.org, www.anthos.es), and data compiled through fieldtrips. Climatic data for current conditions were obtained from WorldClim (www.worldclim.org; Hijmans *et al.* 2005). For past climate scenarios, we used two global Hadley Centre general circulation models that incorporate the effect of changes in atmospheric CO_2 and that have been previously used to represent major changes in global climate (Meseguer *et al.* 2014): a 280-ppm CO_2 Late Miocene simulation (Bradshaw *et al.* 2012) and a 560-ppm CO_2 Mid-Pliocene simulation (Beerling *et al.* 2009). Simulations were cropped to include only Africa and surrounding areas. To model the distribution of *Canarina,* we combined the available 122 occurrences with a set of six bioclimatic variables that could be estimated for past scenarios: total annual precipitation, maximum and minimum monthly precipitation, annual mean temperature, and maximum and minimum monthly temperature. We ran the analyses considering two 4-month periods that cover the two seasons with more accentuated differences in precipitation: November to February and June to September; using as geographic boundaries, the grid included within 28°N to -10°S, for both paleoclimate and present-day simulations. Pseudoabsences were generated by selecting 5000 random points across an area that covers slightly further than the current latitudinal range of *Canarina* (latitude 40°N–20°S; longitude 30°W–50°E). We used ensemble modelling (a procedure integrating the results from multiple modelling techniques, Araújo & New 2007) to generate our predictions. Four modelling techniques – generalized linear models (GLM), generalized additive models (GAM), general boosting method (GBM) and random forests (RF) – were run and summarized using R packages *biomod2*, *foreign*, *raster*, *SDMTools*, *rms*, *gbm*, *gam*, *rJava*, *dismo* and *randomForest* (references for R packages are given in the Data S1).

Results

Phylogenetic relationships and molecular dating

Table 1 summarizes some statistics of the genomic regions studied. Figure 2 shows the results of the Bayesian analysis of the *Platycodoneae* data set. Most nodes received a clade support (PP) > 0.95, and the phylogeny was congruent between plastid and nuclear markers (Fig. 2a,b). *Ostrowskia* is recovered as the sister group of *Canarina*, with *Cyclocodon* and *Platycodon* diverging next. Genera *Cyananthus*, *Codonopsis* and *Campanumoea* form the sister clade (Fig. 2a). Analysis of the *Canarina* concatenated nuclear–plastid data set (Fig. 2c) recovered a monophyletic *Canarina* (PP = 1.0), with *Canarina abyssinica* as sister to a clade formed by *Canarina eminii* and *Canarina canariensis* with high support (PP = 1, ML bootstrap = 80). Geographically structured subclades were recovered within each species with varying levels of support. In general, sequence variation among populations was higher in *C. eminii* than in *C. canariensis* (Fig. 2c).

BEAST analysis of the Platycodoneae data set resulted in a phylogeny (Fig. 3a) that was congruent with the MrBayes MCC tree (Fig. 2). Divergence of Campanuleae and Platycodoneae is dated in the Late Eocene (41.9 Ma, 95% HPD = 28.6–54.7, Table S6, Supporting information), with the first divergence within the tribe dated as Oligocene 29.1 Ma (95% HPD = 18.2–42 Ma). *Canarina* and *Ostrowskia* diverged in the Mid-Miocene (13.8 Ma, 6.6–21.7), while the crown age of *Canarina* is dated as Late Miocene (8.2 Ma, 3.3–14.1). Within *Canarina*, the ‘standard’ and ‘nested’ BEAST approaches gave divergence time estimates with overlapping confidence intervals (Figs 3 and 4; Fig. S4, Table S6, Supporting information). Species divergences (stem ages) were dated in the Late Miocene (8.4–6.5 Ma), whereas crown ages in *C. eminii* and *C. canariensis* (the first population divergences) were dated much younger, in the Early–Mid-Pleistocene (1.76–0.76 Ma, Figs 3 and 4). Population ages were generally older in *C. eminii* (1.76–1.28 Ma) than in *C. canariensis* (0.81–0.76) (Fig. S4, Table S6, Supporting information). The nested approach (Fig. 4) resulted in generally younger age estimates for infraspecific events and older ages for the basal, backbone nodes compared to the standard approach (Fig. 3b); for example, the eastern subclade of *C. canariensis* is dated as 0.38 Ma (0.094–0.891) in the nested tree and 0.59 Ma (0.23–1.05) in the non-nested tree, whereas the opposite pattern is seen for the *Canarina–Ostrowskia* divergence (13.9 vs. 11.6 Ma) and the divergence between *Ostrowskia* and *Cyclocodon* (20.9 vs. 14.1 Ma, Figs 3 and 4). There was also a difference in the geographic structuring of the populations: the two populations in the Abyssinian plateau were grouped in a clade with Elgon and Rwenzori in the standard approach (Fig. 3b), but placed in a separate clade in the nested approach, although the latter with weak support (Fig. 3b).

Phylogeography and colonization history

Bayesian ancestral area reconstruction (Fig. 5) supports an origin of *Canarina* in East Africa, although there is considerable uncertainty due to the existence of long basal branches and the different geographic origin of the two outgroups (PP = 0.58). A prior migration event from East Asia to East Africa (PP = 0.41) is inferred along the branches separating *Canarina* from the most closely related genera *Cyclocodon* and *Ostrowskia*, although central Asia is another possibility (PP = 0.22, Fig. 5a). The ancestral area of *C. eminii* is reconstructed as East Africa (PP = 0.59), implying a migration event from East Africa along the long branch (7.9–1.0 Ma) leading to *C. canariensis* (Fig. 5a). Within each species, several migration events are inferred (Fig. 5b–c). In *C. eminii*, the Imatong–Usambara is inferred as the source area, although with low probability (PP = 0.3). Considering plateaus and each sky island as separate

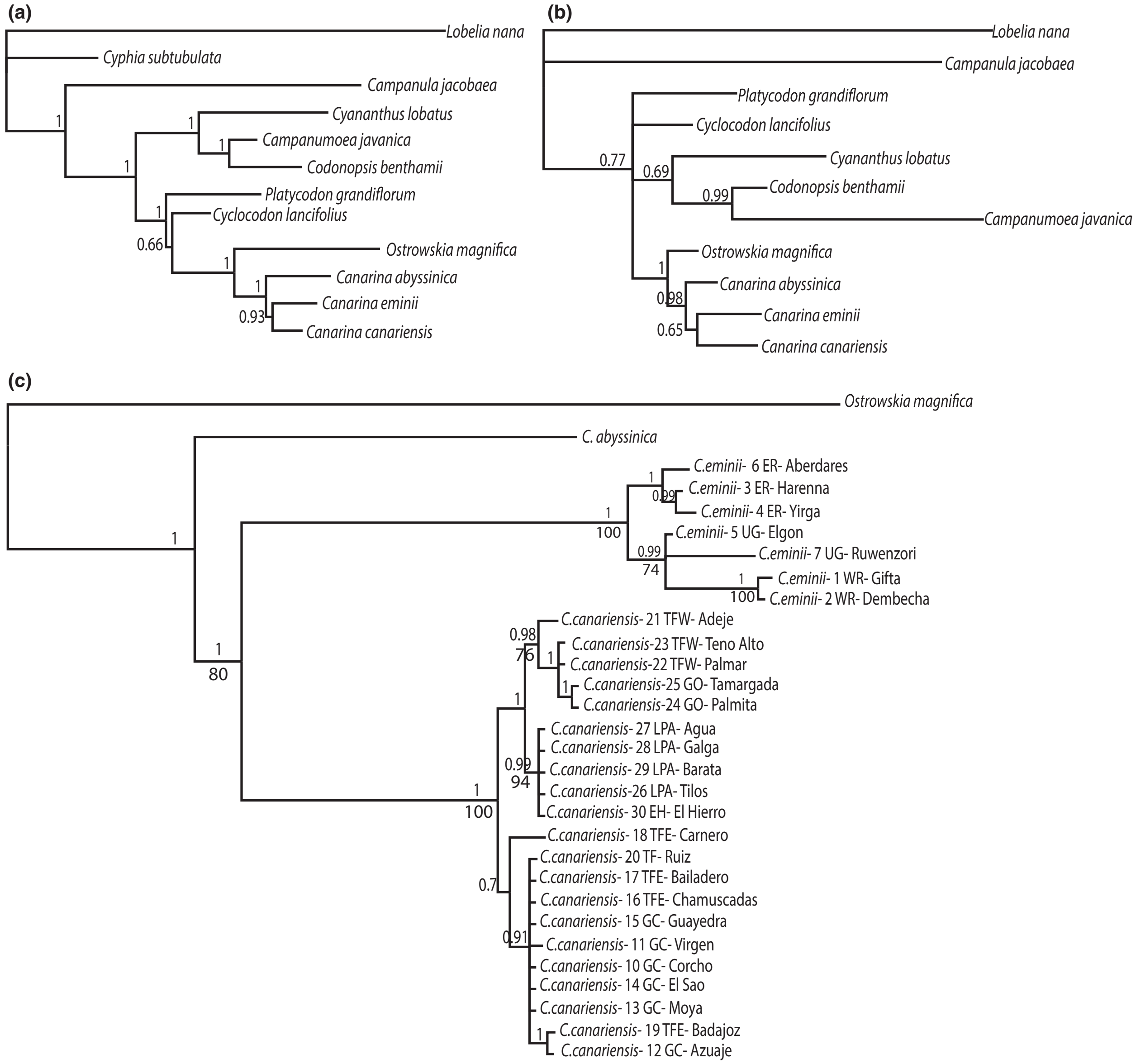


Fig. 2 Bayesian majority-rule consensus trees obtained by MrBayes from: (a) the *Platycodoneae* concatenated chloroplast data set (*psb*J-*pet*A, *trn*L-*trn*F, *pet*B-*pet*D); (b) the *Platycodoneae* nuclear ribosomal (ITS) data set; (c) the *Canarina* concatenated chloroplast and nuclear data set (ITS, *psb*J-*pet*A, *trn*L-*trn*F, *pet*B-*pet*D, *trn*S-*trn*G). Numbers above branches indicate Bayesian credibility values (PP); numbers below branches indicate maximum-likelihood bootstrap support values. Codes for *Canarina* populations correspond to those shown in Table S1.

areas (Fig. S5, Supporting information) resulted in the Abyssinian plateau being inferred as the source area (PP = 0.23), but marginal probabilities for ancestral areas were generally much lower (i.e. there was higher uncertainty because of a lower ratio area/data). In *C. canariensis*, colonization of East Tenerife is followed by an early separation between eastern and western Teneriffean clades (0.8 Ma), and several events of inter-island colonization to the east and west involving Tenerife. Migration from western Tenerife (Teno, Adeje) to La Gomera and to La Palma was inferred within the western subclade, with later migration from La Palma to El Hierro (Fig. 5c). Migration to the east from Tenerife (Tope del Carnero) to Gran Canaria is inferred within the eastern subclade, although colonization in the opposite direction is also possible. At least two other independent events of back colonization from Gran Canaria to Tenerife are inferred, involving the populations of Badajoz, Ruiz and Anaga (Fig. 5c). Constraining the dispersal rates according to geographic

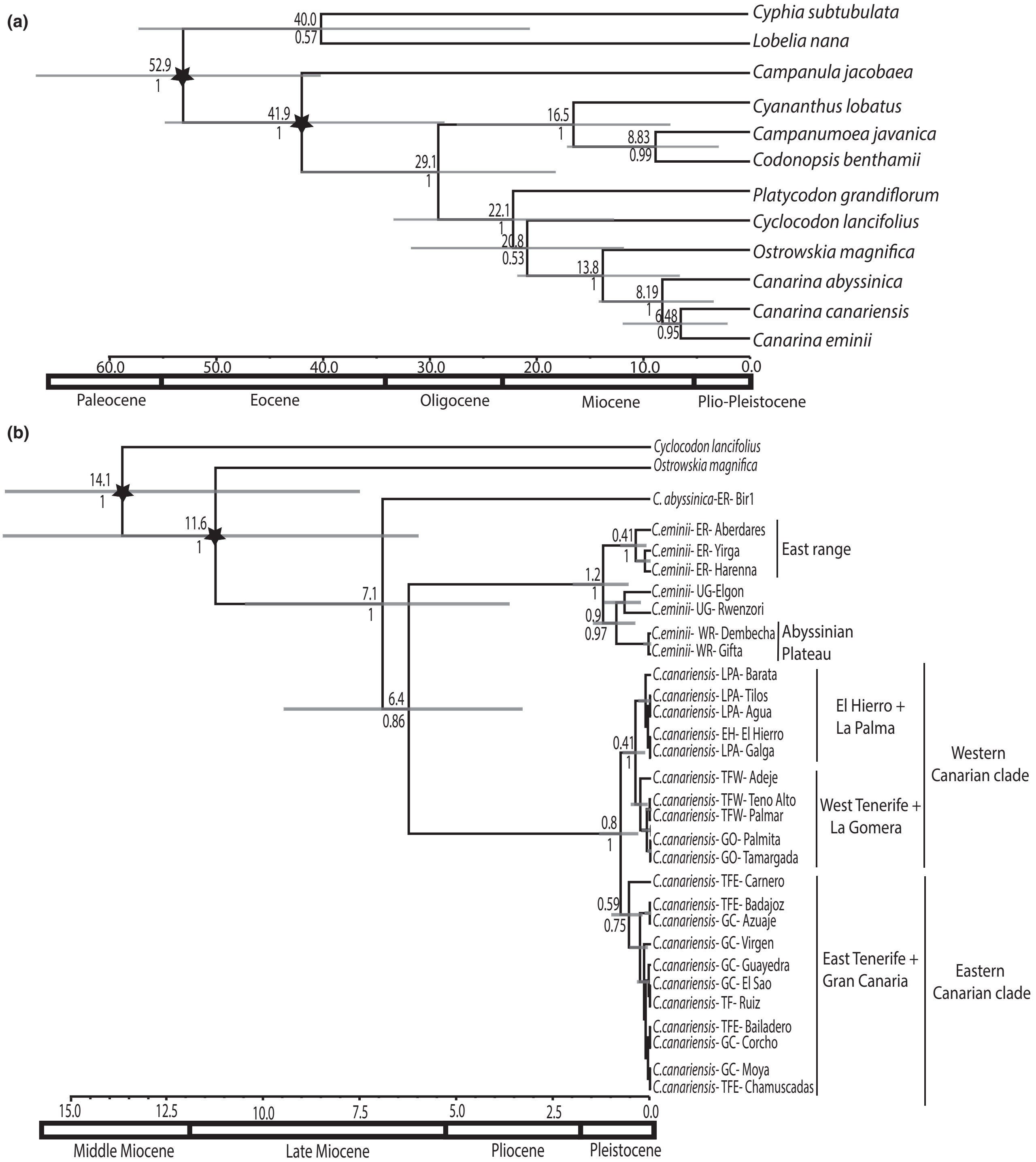


Fig. 3 MCC tree with 95% HPD confidence intervals for main phylogenetic relationships and lineage divergence times obtained in BEAST (stars indicate constrained nodes) for the: (a) Platycodoneae data set (*psb*J-*pet*A, *trn*L-*trn*F, *pet*B-*pet*D). (b) *Canarina* data set (*psb*J-*pet*A, *trn*L-*trn*F, *pet*B-*pet*D, *trn*S-*trn*G, ITS).

distance resulted in a very similar reconstruction, except that Gran Canaria rather than Tenerife was inferred as the ancestral area of the eastern clade of *C. canariensis*, albeit with very low support (TFE = 0.288, GC = 0.290). The geodispersal rate scalar μ (number of dispersal events per site per million year) was considerably higher in *C. canariensis* (3.6) than in *C. eminii* (1.8).

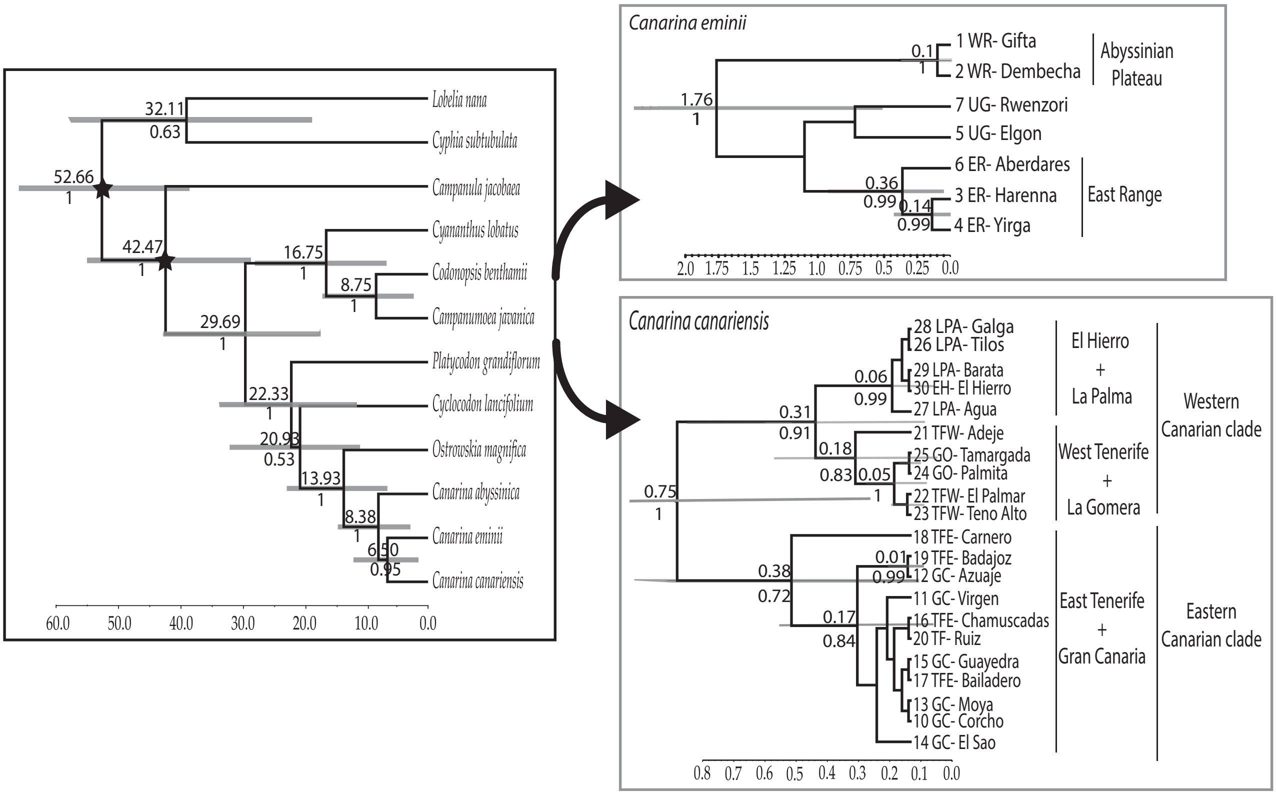


Fig. 4 Nested analyses of all three linked data sets: Platycodoneae (left) and *Canarina eminii* and *Canarina canariensis* (right) (see text for more details). Numbers above branches indicate mean ages and numbers below branches indicate Bayesian PP. Codes for *Canarina* populations correspond to those shown in Table S1. Mean ages and confidence intervals of all analyses are indicated in Fig. S4 and Table S6 (Supporting information).

Ecological niche modelling

Our climate niche projections predict that the geographical area with favourable climatic conditions for *Canarina* experienced a reduction from the Late Miocene to the present (Fig. 6). A climatic 'corridor' with suitable conditions can be observed in the Late Miocene projection, connecting east and western North Africa. This connection is interrupted in the Mid-Pliocene simulation, which shows fragmentation into isolated pockets of climatically favourable conditions. The inferred potential distribution for the present largely coincides with the extant distribution, showing an extreme reduction in range at both sides of the Sahara desert.

Discussion

Secondary calibration vs. nested-dating approach

A standard problem in plant phylogenetic dating studies is the lack of fossil calibration points. This is especially important in Rand Flora groups because of the limited number of macrofossils known from North Africa and the Canary Islands (Whittaker *et al.* 2008; but see Anderson *et al.* 2009). The most common solution to this problem has been to use a secondary calibration approach, in which age constraints derived from the analysis of a higher-level phylogeny including the group of interest (e.g. the Platycodoneae data set), itself calibrated with the fossil record or with other external evidence (e.g. Bell *et al.* 2010's analysis), is used to provide calibration points for the dating of a less inclusive data set, for example the *Canarina* data set. This often translates into a loss of precision in the age estimates due to the need to use an uninformative, broad mean rate prior. Second, if the data set used to estimate lineage divergence times spans both inter- and infraspecific divergences, this might result in biased age estimates, for example, when the phylogeny combines a dense population sampling for one species on one hand, embedded within a tree in which the rest of taxa, at species or above-species level, are represented by one sequence each, on the other (Nolasco-Soto *et al.* 2014). The change in the model of molecular evolution as we move from phylogenetic substitution rates at interspecific relationships to the coalescent dynamics

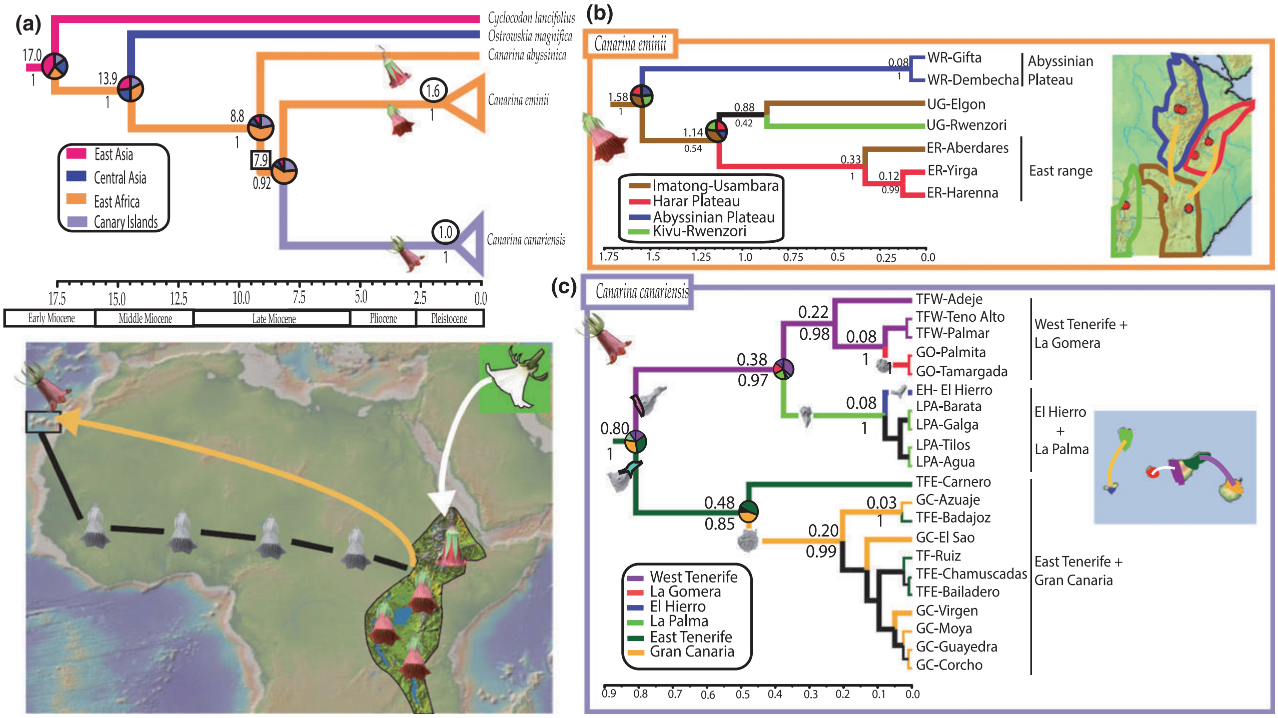


Fig. 5 Results from the BEAST Bayesian ancestral range reconstruction (Lemey *et al.* 2009). Coloured branch lengths (see legend) represent for each lineage the ancestral range with the highest posterior probability. Pie charts at nodes represent uncertainty in the estimation, with black colour representing ancestral areas receiving <0.1 posterior probabilities. (a) MCC tree from the analysis of the *Canarina* data set (standard secondary calibration approach; stem age highlighted inside a square and crown ages highlighted inside circles). (b) MCC tree of the *Canarina eminii* population-level data set (nested-dating approach). (c) MCC tree of the *Canarina canariensis* population-level data set (nested-dating approach). Numbers above branches indicate mean ages and numbers below branches indicate Bayesian PP. Lines in maps represent migration events that receive significant support from the data, as recovered by the BSVS procedure. Colour intensity and thickness of these lines proportional to relative strength (the thicker the line, the higher the dispersal rate) and support (the more intense the colour, the stronger the support: purple > yellow > white). Maps have been modified from satellite pictures in Google Earth.

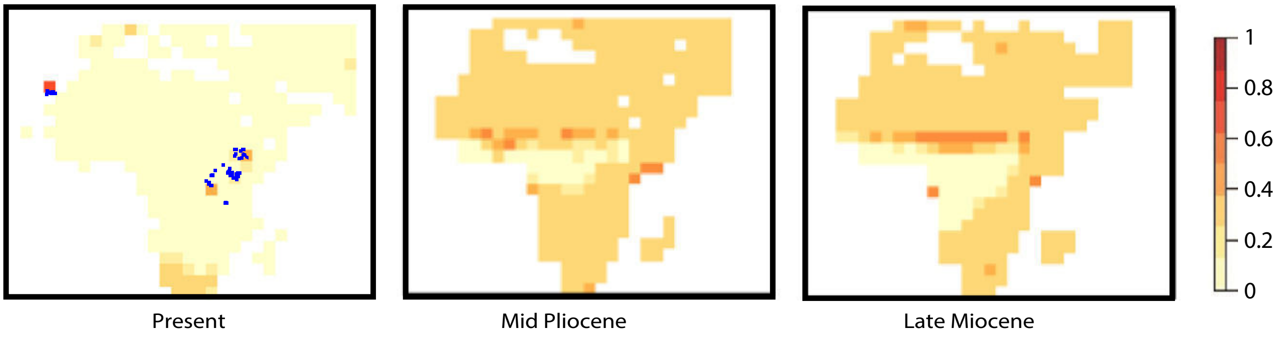


Fig. 6 Geographic projections of the climatic niche model of *Canarina* over three different time periods: present, Mid-Pliocene and Late Miocene. Blue circles indicate extant occurrences and represent the entire current distribution. Soft yellow-coloured regions indicate low climatic suitability values; conversely, dark red indicate high suitability areas.

characteristic of infraspecific evolution might overestimate the age of the most recent events, due to the time dependency in molecular rates and to the fact that gene coalescent events often precede species divergences at the population level (Ho *et al.* 2005, 2011). This is especially problematic if deep-time calibration points are used to date the basal nodes that require the inclusion of distantly related outgroup taxa.

To reconcile deep calibration and species demographic history, Ho *et al.* (2008) proposed an approach in which independent demographic (coalescent) priors were applied to each species, while the basal nodes connecting the clades in the tree are modelled according to a stochastic branching tree prior. The approach followed here, based on Pokorny *et al.* (2011), is slightly different as we do not have infraspecific sampling for all taxa in the phylogeny (e.g. the outgroup taxa are represented by one sequence each). Instead, we used different partitions, sharing some of the taxa and markers, in which the 'calibrated' higher-level partition informs the molecular clock from which the molecular rates of the lower-level partitions are drawn from. Our approach is also different to the 'multispecies coalescent' model in *BEAST (Heled & Drummond 2010) because the latter focuses on coestimating a species tree from multiple gene trees across closely related species, while accounting for coalescent-based phenomena that might cause discrepancy between species and gene trees, such as ILS. Heled & Drummond (2010)'s approach requires infraspecies sampling for each species (3–9 gene copies per lineage) to accurately estimate population parameters like effective population sizes (McCormack *et al.* 2011). In our analysis, only two species include population-level data (*Canarina canariensis*, *Canarina eminii*); whereas *C. abyssinica* and the outgroup taxa are represented by one sequence each. Also, ongoing gene flow is unlikely to be a problem for the deepest divergences in our phylogeny, such as the splits between *Canarina* and its closest relatives and between the outgroup taxa. The discussion below focuses on the results from this nested-dating analysis.

Early evolutionary history of canarina

Our phylogeny for Platycodoneae is congruent with previous studies, supporting a close relationship of *Platycodon*, and *Cyclocodon* with *Canarina* (the 'Platycodon clade', Wang *et al.* 2013) and confirming the monotypic genus *Ostrowskia* as the sister group of *Canarina* (Mansion *et al.* 2012). The origin of Platycodoneae is dated around the Late Eocene–Early Oligocene (29 Ma) in agreement with Roquet *et al.* (2009). *Canarina* is unique within Platycodoneae because of its African distribution. Our time estimates for the divergence with the central Asian *Ostrowskia* (14–11 Ma) suggest that *Canarina*'s ancestors could have taken advantage of the collision of the Arabian Plate with Eurasia (c. 16 Ma, Sanmartín 2003; see Allen & Armstrong 2008 for an earlier date) to migrate into eastern Africa from central–west Asia. This migration could also have been favoured by the uplift of the Red Sea margins (c. 14–13 Ma, Goudie 2005) and a dramatic change in climatic conditions around this period. Starting in the Mid-Miocene, a progressive aridification of the African continent – resulting from both global tectonic changes (e.g. the closing of the Tethys Seaway) and the uplift of eastern Africa (Trauth *et al.* 2009) – led to the gradual replacement of lowland rainforests by woodland savannah in the central and northern Sahara and in parts of South Africa, and later expansion of grasslands and open steppe habitats in south-west Asia and eastern Africa (Bonnefille *et al.* 1990; Coetzee 1993; Maley 1996; Plana 2004; Senut *et al.* 2009). It has been suggested that this created a dispersal route that was used by other non-tropical plant lineages – usually with adaptations to more continental conditions – to migrate from west Asia into Africa (Fiz *et al.* 2008; Popp *et al.* 2008; Roquet *et al.* 2009; Barres *et al.* 2013; Meseguer *et al.* 2013). A similar hypothesis has been argued for several East African 'sky island' species, which could have used the Arabian mountains as 'stepping stones' to reach East Africa (Assefa *et al.* 2007; Popp *et al.* 2008). Dispersal from central–west Asia to eastern Africa is also supported by the fact that the fruits of the sister genus of *Canarina* and *Ostrowskia*, are spherical capsules, which when dry are able to release multiple small slight seeds that can be easily dispersed by wind (Zhaparova 1996; Kamelina & Zhinkina 1998). The subsequent isolation of *Canarina* from its Asian ancestors could have been reinforced by the absence of post-Miocene Red Sea land bridges (Fernandes *et al.* 2006) and a global increase in aridification around 8–6 Ma, coincident with a new period of tectonic activity in eastern Africa and the expansion of grasslands in the Horn of Africa (Cerling *et al.* 1997; Sepulchre *et al.* 2006). This event could also account for the divergence of *C. abyssinica* from the ancestor of *C. eminii* and *C. canariensis*, which is estimated around this time in our analysis (8–7 Ma). *Canarina eminii* is commonly associated with well-preserved closed forests, while *C. abyssinica* occurs in open upland forests, so it is possible that habitat specialization driven by Late Miocene climate aridification explains the divergence between these two species.

An alternative topology, showing *C. eminii* and *C. abyssinica* as sister species to *C. canariensis*, was supported by chloroplast markers such as *rpl*32 and 3′*trn*V-*ndh*C. Although incongruence among genes might be attributed to several biological phenomena, in the case of *rpl*32 it is likely that homoplasy related to higher levels of molecular variation (i.e. saturation at deep phylogenetic levels) and difficulties in alignment due to a high indel/substitution ratio (Table S3, Supporting information) had misled the phylogenetic analysis. For 3′*trn*V-*ndh*C, the lack of a closely related outgroup could be the explanation, as when this marker is included in a concatenated cpDNA data set rooted with *Ostrowskia*, we recovered the 'right' topology grouping

C. eminii and *C. canariensis* with relatively high support (PP = 0.98, ML = 77; Fig. S6, Supporting information). In contrast, chloroplast intron regions like the *petD* II intron possess characteristics, such as high phylogenetic signal per informative character and a well-known secondary structure and molecular evolution, that make them an ideal choice for solving phylogenetic relationships at species level in Campanulaceae (Borsch *et al.* 2009; Mansion *et al.* 2012). This was also the marker for which we have sequences for all outgroup taxa. Moreover, *petBD* was, after ITS, the marker showing in our analyses the lowest levels of substitutional saturation and the largest number of potentially informative characters i.e., number of mutations per sequenced nucleotide (Korotkova *et al.* 2011). Therefore, although we recognize that inclusion of additional plastid and nuclear markers is desirable, we believe that the topology grouping *C. eminii* with *C. canariensis* as sister to *C. abyssinica* accurately reflects the evolutionary relationships among the species.

Long-distance dispersal vs. vicariance and climate-driven extinction

The vicariance–refugium hypothesis posits that the Rand Flora pattern was formed by the fragmentation of a once continuous flora by aridification events, leaving relicts at the eastern and western sides of the geographic disjunction. In *Canarina*, this hypothesis would predict a pattern of 'reciprocal monophyly' between the disjunct taxa, with eastern Africa and Canarian taxa recovered as sister groups (Couvreur *et al.* 2008; Thiv *et al.* 2010), and an age for the disjunction that must predate the barrier that caused the range division, that is the origin of the present Sahara desert. Conversely, the long-distance dispersal (LDD) hypothesis implies the expectation that the taxa at one extreme of the disjunction (i.e. the Canarian endemic) would be embedded within a clade formed by taxa from the other side (i.e. an eastern African clade) and that the disjunction should clearly postdate the formation of the barrier.

At first, the pattern found here, with *C. canariensis* nested within a clade of two East African endemics, agrees better with the LDD hypothesis. *Canarina* species are characterized by the presence of fleshy fruits, with passerine bird- and lizard-mediated zoochory reported for *C. canariensis* (Rodríguez *et al.* 2008). A dispersal event across the 7000 km of the Sahara probably requires other dispersal vectors, such as long-distance migratory birds. For example, Popp *et al.* (2011) argued that a recent (Holocene) single long-distance dispersal by a bird could explain the extreme bipolar distribution of crowberries (*Empetrum*), and similar LDD explanations have been proposed to explain wide range disjunctions between South Africa and North Africa/Canary Islands in *Senecio* (Coleman *et al.* 2003; Pelser *et al.* 2012). Nevertheless, the long temporal gap separating *C. canariensis* and *C. eminii*, with a stem age predating the formation of the Sahara, c. 6 Ma agrees better with a climate-driven vicariance explanation. Interestingly, the alternative topology recovered by *rpl*32, grouping *C. eminii* and *C. abyssinica* as sister to *C. canariensis*, would actually reinforce the vicariance explanation, as the divergence between *C. canariensis* and the East African endemics would probably be dated even earlier (>8–7 Ma), substantially predating the age of origin of the Sahara.

What could be the cause behind this vicariance (allopatric) event? Paleontological reconstructions show a wetter North Africa at least until the Late Miocene (Griffin 2002), which became increasingly more arid as a result of successive aridification events related to a variety of factors, including the opening of the Drake Passage, the closing of the Tethys Seaway and the uplift of eastern Africa (Sepulchre *et al.* 2006; Trauth *et al.* 2009). The first recorded signs of aridification in the Sahara date back to the end of the Miocene, ca. 7–6 Ma ago (Senut *et al.* 2009), which is roughly in agreement with the split between *C. eminii* and *C. canariensis* (6.5 Ma). Nevertheless, the rapid alternation of arid and humid periods starting in the Miocene–Pliocene boundary (Trauth *et al.* 2009; (Micheels *et al.* 2009) might have allowed repeated events of isolation and reconnection across both sides of the Sahara (Désamoré *et al.* 2011). We do not have evidence of any of these recent events of reconnection in the phylogeny of *Canarina*. Instead, the 6.4 Ma divergence estimated here between the Canarian and East African endemics is roughly in agreement with the age estimated for the disjunction of other Rand Flora lineages, for example, *Campylanthus* (Thiv *et al.* 2010) or *Plocama*.

In addition, our ecological niche models and paleoclimate projections support the hypothesis of a more widespread distribution of *Canarina* across north–central Africa in the past, which became fragmented by climate change. They show a more or less continuous 'climatic corridor' across North Africa during the Late Miocene period, which became interrupted during the more arid Mid-Pliocene period. The latter shows the presence of isolated patches of climatic suitability (Fig. 6), which could have acted as potential 'stepping stones' for dispersal across the Sahara, or as climatic refugia once aridification started. Worsening climate conditions, with increasing aridity at the Plio–Pleistocene boundary (Senut *et al.* 2009), might have caused the extinction of intermediate populations across central North Africa, leaving the current species as the only remnants (relicts) of a past widespread distribution. Similar scenarios have been hypothesized in other Rand Flora lineages for which supporting fossil evidence exists, such as *Dracaena* (Denk *et al.* 2014). Whether *Canarina* was ever continuously distributed across North Africa, with uninterrupted gene flow between both extremes of the disjunction, or whether, alternatively, the pattern is the result of gradual range expansion, westwards across the Sahara, is difficult to discern with the current evidence. The vicariance hypothesis, for example, predicts also range expansion across the Sahara prior to the allopatric (vicariant) event. Interestingly, the lower levels of genetic diversity found in *C. canariensis* compared to the East African *C. eminii* agree with a more recent dispersal event, perhaps from a now extinct and geographically closer, North African (Moroccan) population. What our evidence does suggest is that *Canarina* could have a wider distribution across north–central Africa in the past and that there has been a long history of isolation between the two extremes of the disjunction. The long stem between the stem divergence of *C. canariensis* and the start of infraspecies (population) divergence can be interpreted as evidence of extinction of the intermediate populations (Antonelli & Sanmartín 2011). Alternatively, it could be understood as the result of strong purifying selection with little population differentiation – driven perhaps by climatic change – and, followed by a recent demographic expansion. We favour extinction over purifying selection because the latter is expected to affect one gene but not to produce congruent patterns across genes (Williamson & Orive 2002). Although population-level studies are needed to test this hypothesis, an interesting corollary of our study is that the age of divergence of an island endemic from its continental sister species is not necessarily equivalent to the age of colonization of the island as it is often assumed in island studies (Kim *et al.* 2008), especially if extinction has been high in the continent.

Geographic oceanic islands vs. 'Ecological' mountain islands

Canarina, with its distribution in *true* oceanic islands and mountain 'sky islands', offers an interesting comparison on the role of geographic vs. ecological barriers in structuring plant genetic variation. It is well known that oceanic islands are able to cope with large climatic changes better than continental landmasses because of the tempering effect created by the ocean to which they are exposed. The sky islands of the Afromontane regions in East Africa (i.e. high plateaus and mountains in Ethiopia and subtropical East Africa) probably acted in a similar way, allowing species and communities to migrate altitudinally and thus avoid the thermal and hydric stress produced by aridification episodes (Fjeldsaå & Lovett 1997). Paleobotanical and phylogeographical evidence suggest that the slopes of these montane regions were covered by forests until recently (Bonnefille *et al.* 1990; Kuper & Kröpelin 2006). During the glacial arid periods of the Late Pliocene and Pleistocene, these forests probably became separated, and later reconnected during the humid, warmer interglacial periods (Coetzee 1964; Maley 1996; Kebede *et al.* 2007; Popp *et al.* 2008). In more recent times, land use and deforestation might have contributed to further isolation of these forest patches (EFAP 1994; FAO 2001). The relatively old infraspecific divergences estimated here for *C. eminii*, ranging from 700 000 years between Elgon and Rwenzori to a few thousand years between Gifta and Dembecha (Fig. 4), suggest that population divergence in this montane species was more likely driven by Pleistocene climatic events than by forest fragmentation after the expansion of agriculture. Moreover, our results support other phylogeographic studies in Afromontane taxa (Knox & Palmer 1998; Kebede *et al.* 2007) that pointed to the Ethiopian Rift Valley as an important geographic barrier, segregating populations to the east and west of this barrier. In contrast, the fact that the eastern subclade in *C. eminii* (0.4 Ma, PP = 1) groups together populations as far away as Harenna Forest and Yirga, in southern Ethiopia, and the Aberdare Range, in Kenya, suggests that the eastern range of the Rift has been less isolated than the west, probably due to the existence of better connections between forest patches on this side of Rift (Coetzee 1964; Hedberg 1969; Kebede *et al.* 2007).

The oldest extant Canary Islands emerged ca. 20 Ma (Fernández-Palacios *et al.* 2011), but our time estimates place population divergence in *C. canariensis* within the last 800 000 years, considerably younger than in *C. eminii*. The first recovered divergence event is one of within-island segregations between east and west Tenerife. This pattern has been reported in other endemic organisms (Juan *et al.* 2000) and attributed to the geological origin of Tenerife, which resulted from the merging of three paleoislands c.a. 1 Ma ago (Ancochea *et al.* 1990). Subsequent events, such as a central eruptive episode ca. 0.8 Ma and giant landslides on the northern flank of Tenerife (Krastel *et al.* 2001), might have later prevented reconnections between east and west *C. canariensis* populations. Interisland dispersal events from Tenerife to the east and west are also reconstructed, in agreement with the role of the central islands as a source of migration within the archipelago (Sanmartín *et al.* 2008), but these are all dated after the divergence within Tenerife, indicating that probably within-island catastrophic/geological events have been a more important barrier to dispersal for *C. canariensis* populations than the ocean waters separating the islands.

Conclusions

Continental-scale disjunct distribution patterns, such as the Rand Flora, are especially interesting in the context of the present biodiversity crisis because they are often attributed to climate-driven extinction that would have extirpated a once continuous biota from part of its distributional range (Axelrod & Raven 1978; Crisp & Cook 2007). Here, we show that in the case of genus *Canarina*, this disjunction predates the origin of the Sahara and might be explained by climate-driven vicariance and extinction. The potential ancient age of within-continent disjunctions (Crisp & Cook 2007) implies that we often do not have fossil taxa close to the group of interest. We benefit here from a nested-dating approach that implements two different tree models (birthdeath vs. coalescent) for simultaneous phylogenetic analysis of data at different levels of organization. Our study emphasizes the importance of climate-driven extinction in the assembly of regional biodiversity patterns, in particular in the context of the ongoing aridification of the Mediterranean Basin.

Acknowledgements

We are grateful to Richard Abbott and three anonymous reviewers, whose comments helped to significantly improve the manuscript. We thank Fatima Durán and Guillermo Sanjuanbenito for laboratory assistance. Field work could not have been conducted without the cooperation of Juan Ojeda, Oscar Saturno and the staff at the Jardín Botánico Canario Viera y Clavijo (Gran Canaria); Cabildo de Tenerife; Jacinto Leralta and Ángel Fernández (La Gomera); Félix Manuel Medina from the Cabildo of La Palma and the Cabildo of El Hierro are thanked for help with accommodation and sampling logistics during field expeditions. We thank to Kenya National Commission for Science, Technology and Innovation, the authorities of the Bale Mountains and the Harenna Forest National Park (Ethiopia) and Rwenzori Mountains National Park (Uganda) for its collaboration during fieldwork. We also thank Alejandro González, Juli Caujapé and Moisés Soto for providing fresh samples, David Beerling for providing climatic data and Loïc Pellissier, Javier Fuertes, Mike Thiv and José Luis Blanco Pastor for help during different stages of this work. This work was funded by the Spanish Ministerio de Economía y Competitividad (Projects CGL2006-09696, CGL2009-1332-C03-01, CGL2012-40129-C02-01) the JAE-Doc programme (CSIC/FSE) to MA, and a PhD research grant (BES-2010-037261) to MM. LP was funded by a research contract under CGL2012-40129-C02-01.

References

Allen MB, Armstrong HA (2008) Arabia-Eurasia collision and the forcing of Mid-Cenozoic global cooling. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **265**, 52–58.

Ancochea E, Fuster J, Ibarrola E *et al.* (1990) Volcanic evolution of the island of Tenerife (Canary Islands) in the light of new K-Ar data. *Journal of Volcanology and Geothermal Research*, **44**, 231–249.

Anderson CL, Channing A, Zamuner AB (2009) Life, death and fossilization on Gran Canaria–implications for Macaronesian biogeography and molecular dating. *Journal of Biogeography*, **36**, 2189–2201.

Andrus N, Trusty J, Santos-Guerra A, Jansen RK, Francisco-Ortega J (2004) Using molecular phylogenies to test phytogeographical links between East/South Africa-Southern Arabia and the Macaronesian islands—a review, and the case of *Vierea* and *Pulicaria* section Vieraeopsis (Asteraceae). *Taxon*, **53**, 333–346.

Ané C, Larget B, Baum DA, Smith SD, Rokas A (2007) Bayesian estimation of concordance among gene trees. *Molecular Biology and Evolution*, **24**, 412–426.

Antonelli A, Sanmartín I (2011) Mass extinction, gradual cooling, or rapid radiation? Reconstructing the spatiotemporal evolution of the ancient angiosperm genus *Hedyosmum* (Chloranthaceae) using empirical and simulated approaches. *Systematic Biology*, **60**, 596–615.

Araújo MB, New M (2007) Ensemble forecasting of species distributions. *Trends in Ecology and Evolution*, **22**, 42–47.

Assefa A, Ehrich D, Taberlet P, Nemomissa S, Brochmann C (2007) Pleistocene colonization of afro-alpine 'sky islands' by the arctic-alpine *Arabis alpina*. *Heredity*, **99**, 133–142.

Axelrod DI, Raven PH (1978) Late Cretaceous and Tertiary vegetation history of Africa. In: *Biogeography and Ecology of Southern Africa* (ed. Werger MJA), pp. 77–130. Springer, the Netherlands.

Baele G, Lemey P, Bedford T, Rambaut A, Suchard MA, Alekseyenko AV (2012) Improving the accuracy of demographic and molecular clock model comparison while accommodating phylogenetic uncertainty. *Molecular Biology and Evolution*, **29**, 2157–2167.

Barres L, Sanmartín I, Anderson CL *et al.* (2013) Reconstructing the evolution and biogeographic history of tribe Cardueae (Compositae). *American Journal of Botany*, **100**, 867–882.

Beerling D, Berner RA, Mackenzie FT, Harfoot MB, Pyle JA (2009) Methane and the CH_4 related greenhouse effect over the past 400 million years. *American Journal of Science*, **309**, 97–113.

Bell CD, Soltis DE, Soltis PS (2010) The age and diversification of the angiosperms re-revisited. *American Journal of Botany*, **97**, 1296–1303.

Blanco-Pastor JL, Fernández-Mazuecos M, Vargas P (2013) Past and future demographic dynamics of alpine species: limited genetic consequences despite dramatic range contraction in a plant from the Spanish Sierra Nevada. *Molecular Ecology*, **22**, 4177–4195.

Bonnefille R, Roeland JC, Gulot J (1990) Temperature and rainfall estimates for the past 40000 years in equatorial Africa. *Nature*, **346**, 347–349.

Borsch T, Korotkova N, Raus T, Lobin W, Lohne C (2009) The petD group II intron as a species level marker: utility for tree inference and species identification in the diverse genus *Campanula* (Campanulaceae). *Willdenowia*, **39**, 7–33.

Bradshaw CD, Lunt DJ, Flecker R *et al.* (2012) The relative roles of CO_2 and palaeogeography in determining Late Miocene climate: results from a terrestrial model-data comparison. *Climate of the Past Discussions*, **8**, 715–786.

Bramwell D (1985) Contribución a la biogeografía de las Islas Canarias. *Botánica Macaronésica*, **14**, 3–34.

Cerling TE, Harris JM, MacFadden BJ *et al.* (1997) Global vegetation change through the Miocene/Pliocene boundary. *Nature*, **389**, 153–158.

Christ H (1892) Exposé sur le rôle que joue dans le domaine de nos flores la flore dite *ancienne africaine*. *Archives des Sciences Physiques et Naturelles Genève*, **3**, 369–374.

Coetzee JA (1964) Evidence for a considerable depression of the vegetation belts during the Upper Pleistocene on the East African mountains.

Coetzee JA (1993) African flora since the terminal Jurassic. *Biological Relationships between Africa and South America*, **37**, 61.

Coleman M, Liston A, Kadereit JW, Abbott RJ (2003) Repeat intercontinental dispersal and Pleistocene speciation in disjunct Mediterranean and desert *Senecio* (Asteraceae). *American Journal of Botany*, **90**, 1446–1454.

Couvreur TL, Chatrou LW, Sosef MS, Richardson JE (2008) Molecular phylogenetics reveal multiple tertiary vicariance origins of the African rain forest trees. *BMC Biology*, **6**, 54.

Crisp MD, Cook LG (2007) A congruent molecular signature of vicariance across multiple plant lineages. *Molecular Phylogenetics and Evolution*, **43**, 1106–1117.

Denk T, Güner HT, Grimm GW (2014) From mesic to arid: leaf epidermal features suggest preadaptation in Miocene dragon trees (*Dracaena*). *Review of Palaeobotany and Palynology*, **200**, 211–228.

Désamoré A, Laenen B, Devos N *et al.* (2011) Out of Africa: north-westwards Pleistocene expansions of the heather Erica arborea. *Journal of Biogeography*, **38**, 164–176.

Donoghue MJ, Smith SA (2004) Patterns in the assembly of temperate forests around the Northern Hemisphere. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, **359**, 1633–1644.

Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, **7**, 214.

EFAP (1994) Ethiopia Forestry Action Program. Final Report, Ministry of Natural Resources Development and Environmental Protection, Addis Ababa.

Engler A (1910) Die Pflanzenwelt Afrikas inbesondere seiner tropischen Gebiete. Gründzuge der Pflanzenverbreitung in Afrika un die Characterpflanzen Afrikas. In: *Die Vegetation der Erde* (eds Engler A, Drude O), pp. 1–1030. Verlag vonc Wilhelm Engelmann, Leipzig.

FAO (2001) Food and Agricultural Organization of the United Nations. Forest Resource Assessment, 2000 (FAO, Rome). Forestry Paper 140.

Fernandes CA, Rohling EJ, Siddall M (2006) Absence of post-Miocene Red Sea land bridges: biogeographic implications. *Journal of Biogeography*, **33**, 961–966.

Fernández-López AB (2014) Mapa de *Canarina canariensis* para la flora vascular de la isla de la Gomera. Parque Nacional de Garajonay.

Fernández-Palacios JM, de Nascimento L, Otto R *et al.* (2011) A reconstruction of Palaeo-Macaronesia, with particular reference to the long-term biogeography of the Atlantic island laurel forests. *Journal of Biogeography*, **38**, 226–246.

Fiz O, Vargas P, Alarcón M, Aedo C, García JL, Aldasoro JJ (2008) Phylogeny and historical biogeography of *Geraniaceae* in relation to climate changes and pollination ecology. *Systematic Botany*, **33**, 326–342.

Fjeldsaå J, Lovett JC (1997) Geographical patterns of old and young species in African forest biota: the significance of specific montane areas as evolutionary centres. *Biodiversity and Conservation*, **6**, 325–346.

Gehrke B, Linder HP (2014) Species richness, endemism and species composition in the tropical Afroalpine flora. *Alpine Botany*, **124**, 165–177.

Givnish TJ, Renner SS (2004) Tropical intercontinental disjunctions: Gondwana breakup, immigration from the boreotropics, and transoceanic dispersal. *International Journal of Plant Sciences*, **165**, S1–S6.

Goudie AS (2005) The drainage of Africa since the Cretaceous. *Geomorphology*, **67**, 437–456.

Griffin DL (2002) Aridity and humidity: two aspects of the late Miocene climate of North Africa and the Mediterranean. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **182**, 65–91.

Hedberg O (1961) Monograph of the genus *Canarina* L. (Campanulaceae). *Svensk Botanisk Tidskrift*, **55**, 17–62.

Hedberg O (1969) Evolution and speciation in a tropical high mountain flora. *Biological Journal of the Linnaean Society*, **1**, 135–148.

Heled J, Drummond AJ (2010) Bayesian inference of species trees from multilocus data. *Molecular Biology and Evolution*, **27**, 570–580.

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**, 1965–1978.

Ho SY, Phillips MJ, Cooper A, Drummond AJ (2005) Time dependency of molecular rate estimates and systematic overestimation of recent divergence times. *Molecular Biology and Evolution*, **22**, 1561–1568.

Ho SY, Larson G, Edwards CJ *et al.* (2008) Correlating Bayesian date estimates with climatic events and domestication using a bovine case study. *Biology Letters*, **4**, 370–374.

Ho SY, Lanfear R, Bromham L *et al.* (2011) Time-dependent rates of molecular evolution. *Molecular Ecology*, **20**, 3087–3101.

Hooker JD (1853) *Introductory Essay to the Flora of New Zealand*. Lovell Reeve, London.

Hooker JD (1867) On insular floras: a lecture. *Journal of Botany*, **5**, 23–31.

von Humboldt A, Bonpland A (1805) *Essai sur la ge´Ographie des Plantes; Accompagne´ D'un Tableau Physique des re´Gions e´Quinoxiales, Fonde´ sur des Mesures exe´Cute´es, Depuis le Dixie'me Degre´ de Latitude Bore´ale Jusqu'au Dixie'me Degre´ de Latitude Australe, Pendant les Anne´es 1799, 1800, 1801, 1802 et 1803*. chez Levrault, Schoell et compagnie, libraries, Paris.

Juan C, Emerson BC, Oromí P, Hewitt GM (2000) Colonization and diversification: towards a phylogeographic synthesis for the Canary Islands. *Trends in Ecology & Evolution*, **15**, 104–109.

Kamelina OP, Zhinkina NA (1998) On the embryology of *Ostrowskia magnifica*. The ovule and seed. *Botanicheskii Zhurnal*, **83**, 9–20.

Kebede M, Ehrich D, Taberlet P, Nemomissa S, Brochmann C (2007) Phylogeography and conservation genetics of a giant lobelia (*Lobelia giberroa*) in Ethiopian and Tropical East African mountains. *Molecular Ecology*, **16**, 1233–1243.

Kim SC, McGowen MR, Lubinsky P, Barber JC, Mort ME, Santos-Guerra A (2008) Timing and tempo of early and successive adaptive radiations in Macaronesia. *PLoS ONE*, **3**, e2139.

Knox EB, Palmer JD (1998) Chloroplast DNA evidence on the origin and radiation of the giant lobelias in eastern Africa. *Systematic Botany*, **23**, 109–149.

Kornhall P, Heidari N, Bremer B (2001) Selagineae and Manuleeae, two tribes or one? Phylogenetic studies in the Scrophulariaceae. *Plant Systematics and Evolution*, **228**, 199–218.

Korotkova N, Borsch T, Quandt D, Taylor NP, Müller KF, Barthlott W (2011) What does it take to resolve relationships and to identify species with molecular markers? An example from the epiphytic Rhipsalideae (Cactaceae). *American Journal of Botany*, **98**, 1549–1572.

Krastel S, Schmincke HU, Jacobs CL, Rihm R, Le Bas TP, Alibés B (2001) Submarine landslides around the Canary Islands. *Journal of Geophysical Research: Solid Earth (1978–2012)*, **106**, 3977–3997.

Kuper R, Kröpelin S (2006) Climate-controlled Holocene occupation in the Sahara: motor of Africa's evolution. *Science*, **313**, 803–807.

Larget BR, Kotha SK, Dewey CN, Ané C (2010) BUCKy: gene tree/species tree reconciliation with Bayesian concordance analysis. *Bioinformatics*, **26**, 2910–2911.

Lebrun J (1961) Les deux flores d'Afrique tropicale. Royale de Belgique. Classe des Sciences. Memoires. Collection in-8. 2éme série 32, 1–82.

Lemey P, Rambaut A, Drummond AJ, Suchard MA (2009) Bayesian phylogeography finds its roots. *PLoS Computational Biology*, **5**, e1000520.

Linder HP (2014) The evolution of African plant diversity. *Frontiers in Ecology and Evolution*, **2**, 38.

Maley J (1996) The African rain forest–main characteristics of changes in vegetation and climate from the Upper Cretaceous to the Quaternary. *Proceedings of the Royal Society of Edinburgh. Section B. Biological Sciences*, **104**, 31–73.

Mansion G, Parolly G, Crowl AA *et al.* (2012) How to handle speciose clades? Mass taxon-sampling as a strategy towards illuminating the natural history of Campanula (Campanuloideae) *PLoS ONE*, **7**, e50076.

McCormack JE, Huang H, Knowles LL (2009) Sky Islands. In: *Encyclopedia of Islands* (eds Gillespie RG, Clague DA), pp. 841–843. University of California Press Ltd, Berkeley and Los Angeles.

McCormack JE, Heled J, Delaney KS, Peterson AT, Knowles LL (2011) Calibrating divergence times on species trees versus gene trees: implications for speciation history of *Aphelocoma* jays. *Evolution*, **65**(1), 184–202.

Meseguer AS, Aldasoro JJ, Sanmartín I (2013) Bayesian inference of phylogeny, morphology and range evolution reveals a complex evolutionary history in St. John's wort (*Hypericum*). *Molecular Phylogenetics and Evolution*, **67**, 379–403.

Meseguer AS, Lobo JM, Ree R, Beerling DJ, Sanmartín I (2014) Integrating Fossils, Phylogenies, and Niche Models into Biogeography to reveal ancient evolutionary history: the Case of *Hypericum* (Hypericaceae). *Systematic Biology*, syu**088**.

Micheels A, Eronen J, Mosbrugger V (2009) The Late Miocene climate response to a modern Sahara desert. *Global and Planetary Change*, **67**, 193–204.

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

Nolasco-Soto J, González-Astorga J, Nicolalde-Morejón F, Vergara-Silva F, De los Monteros AE, Medina-Villarreal A (2014) Phylogeography and demographic history of *Zamia paucijuga* Wieland (Zamiaceae), a cycad species from the Mexican Pacific slope. *Plant Systematics and Evolution*, **301**, 623–637.

Pelser PB, Abbott RJ, Comes HP *et al.* (2012) The genetic ghost of an invasion past: colonization and extinction revealed by historical hybridization in *Senecio*. *Molecular Ecology*, **21**, 369–387.

Plana V (2004) Mechanisms and tempo of evolution in the African Guineo-Congolian rainforest. *Philosophical Transactions of the Royal Society of London*, **359**, 1585–1594.

Pokorny L, Oliván G, Shaw A (2011) Phylogeographic patterns in two southern hemisphere species of *Calyptrochaeta* (Daltoniaceae, Bryophyta). *Systematic Botany*, **36**, 542–553.

Popp M, Gizaw A, Nemomissa S, Suda J, Brochmann C (2008) Colonization and diversification in the African 'sky islands' by Eurasian *Lychnis* L. (Caryophyllaceae). *Journal of Biogeography*, **35**, 1016–1029.

Popp M, Mirré V, Brochmann C (2011) A single Mid-Pleistocene long-distance dispersal by a bird can explain the extreme bipolar disjunction in crowberries (*Empetrum*). *Proceedings of the National Academy of Sciences, USA*, **108**, 6520–6525.

Quézel P (1978) Analysis of the Flora of Mediterranean and Saharan Africa. *Annals of the Missouri Botanical Garden*, **65**, 479–534.

Rambaut A, Drummond AJ (2013) TreeAnnotator v1. 7.0.

Rambaut A, Drummond AJ, Suchard M (2013) Tracer v1. 6.

Raven PH, Axelrod DI (1972) Plate tectonics and Australasian Paleobiogeography. *Science*, **176**, 1379–1386.

Renner S (2004) Plant dispersal across the tropical Atlantic by wind and sea currents. *International Journal of Plant Sciences*, **165**, S23–S33.

Riina R, Peirson JA, Geltman DV *et al.* (2013) A worldwide molecular phylogeny and classification of the leafy spurges, Euphorbia subgenus Esula (Euphorbiaceae). *Taxon*, **62**, 316–342.

Rodríguez A, Nogales M, Rumeu B, Rodríguez B (2008) Temporal and spatial variation in the diet of the endemic lizard *Gallotia galloti* in an insular Mediterranean scrubland. *Journal of Herpetology*, **42**, 213–222.

Ronquist F, Teslenko M, van der Mark P *et al.* (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology*, **61**, 539–542.

Roquet C, Sanmartín I, Garcia-Jacas N *et al.* (2009) Reconstructing the history of Campanulaceae with a Bayesian approach to molecular dating and dispersal–vicariance analyses. *Molecular Phylogenetics and Evolution*, **52**, 575–587.

Rothfels CJ, Larsson A, Kuo L-Y, Korall P, Chiou W-L, Pryer KM (2012) Overcoming deep roots, fast rates, and short internodes to resolve the ancient rapid radiation of eupolypod II ferns. *Systematic Biology*, **61**, 490–509.

Ryan WB, Carbotte SM, Coplan JO *et al.* (2009) Global multi-resolution topography synthesis. *Geochemistry, Geophysics, Geosystems*, **10**, 3.

Sanmartín I (2003) Dispersal vs. vicariance in the Mediterranean: historical biogeography of the Palearctic Pachydeminae (Coleoptera, Scarabaeoidea). *Journal of Biogeography*, **30**, 1883–1897.

Sanmartín I, Ronquist F (2004) Southern Hemisphere biogeography inferred by event-based models: plant versus animal patterns. *Systematic Biology*, **53**, 216–243.

Sanmartín I, Enghoff H, Ronquist F (2001) Patterns of animal dispersal, vicariance and diversification in the Holarctic. *Biological Journal of the Linnean Society*, **73**, 345–390.

Sanmartín I, Van Der Mark P, Ronquist F (2008) Inferring dispersal: a Bayesian approach to phylogeny-based island biogeography, with special reference to the Canary Islands. *Journal of Biogeography*, **35**, 428–449.

Sanmartín I, Anderson CL, Alarcon M, Ronquist F, Aldasoro JJ (2010) Bayesian island biogeography in a continental setting: the Rand Flora case. *Biology Letters*, **6**, 703–707.

Senut B, Pickford M, Ségalen L (2009) Neogene desertification of Africa. *Comptes Rendus Geoscience*, **341**, 591–602.

Sepulchre P, Ramstein G, Fluteau F, Schuster M, Tiercelin JJ, Brunet M (2006) Tectonic uplift and Eastern Africa aridification. *Science*, **313**, 1419–1423.

Shaw J, Lickey EB, Beck JT *et al.* (2005) The tortoise and the hare II: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. *American Journal of Botany*, **92**, 142–166.

Shaw J, Lickey EB, Schilling EE, Small RL (2007) Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the tortoise and the hare III. *American Journal of Botany*, **94**, 275–288.

Smith SA, Donoghue MJ (2010) Combining historical biogeography with niche modeling in the Caprifolium clade of *Lonicera* (Caprifoliaceae, Dipsacales). *Systematic Biology*, **59**, 322–341.

Stamatakis A, Hoover P, Rougemont J (2008) A rapid bootstrap algorithm for the RAxML web servers. *Systematic Biology*, **57**, 758–771.

Swofford D (2002) *PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods) v4*. Sinauer Associates, Sunderland, Massachusetts.

Thiv M, Thulin M, Hjertson M, Kropf M, Linder HP (2010) Evidence for a vicariant origin of Macaronesian–Eritreo/Arabian disjunctions in *Campylanthus* Roth (Plantaginaceae). *Molecular Phylogenetics and Evolution*, **54**, 607–616.

Trauth MH, Larrasoaña JC, Mudelsee M (2009) Trends, rhythms and events in Plio-Pleistocene African climate. *Quaternary Science Reviews*, **28**, 399–411.

Treviranus GR (1803) Biologie, oder Philosophie der lebenden. Natur, 2 volumes, Röwer, Göttingen.

Wang Q, Zhou SL, Hong DY (2013) Molecular phylogeny of the platycodonoid group (Campanulaceae s. str.) with special reference to the circumscription of *Codonopsis*. *Taxon*, **62**, 498–504.

White F (1983) *The Vegetation of Africa: A Descriptive Memoir to Accompany the UNESCO/AETFAT/UNSO Vegetation map of Africa by F White*. Natural Resources Research Report XX, UNESCO, Paris, France.

Whittaker RJ, Triantis KA, Ladle RJ (2008) A general dynamic theory of oceanic island biogeography. *Journal of Biogeography*, **35**, 977–994.

Wiens JJ, Donoghue MJ (2004) Historical biogeography, ecology and species richness. *Trends in Ecology and Evolution*, **19**, 639–644.

Williamson S, Orive ME (2002) The genealogy of a sequence subject to purifying selection at multiple sites. *Molecular Biology and Evolution*, **19**, 1376–1384.

Wolfe KH, Li W-H, Sharp PM (1987) Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proceedings of the National Academy of Sciences, USA*, **84**, 9054–9058.

Wondimu T, Gizaw A, Tusiime FM *et al.* (2014) Crossing barriers in an extremely fragmented system: two case studies in the afro-alpine sky island flora. *Plant Systematics and Evolution*, **300**, 415–430.

Yesson C, Culham A (2006) Phyloclimatic modeling: combining phylogenetics and bioclimatic modeling. *Systematic Biology*, **55**, 785–802.

Zhaparova NK (1996) Root development of *Ostrowskia magnifica* Regel, a rare species in Kazakhstan. *Acta Phytogeographica Suecica*, **81**, 88–91.

Zhou Z, Hong D, Niu Y *et al.* (2013) Phylogenetic and biogeographic analyses of the Sino-Himalayan endemic genus *Cyananthus* (Campanulaceae) and implications for the evolution of its sexual system. *Molecular Phylogenetics and Evolution*, **68**, 482–497.

I.S. and M.M. designed the study. M.M., J.J. and M.A. contributed samples. M.M. performed research under the supervision of I.S. M.M., L.P. and I.S. analysed the data. M.M. and I.S. wrote the manuscript with help from L.P.

Data accessibility

DNA sequences: GenBank Accession Nos KP761423 to KP761687. GenBank accessions, sampling locations and/or online-only appendices uploaded as online supplemental material. Original script input file used to perform the nested BEAST approach: Dryad doi: 10.5061/dryad.5jc73. NEXUS files for the single and concatenated data set: Dryad doi:10.5061/dryad.5jc73.

Supporting information

Additional supporting information may be found in the online version of this article.

Data S1 Supplementary text including: study group; expanded material and methods; on our sampling effort and the decreasing distribution of *Canarina abyssinica*; additional references.

Table S1 Voucher information and GenBank accession numbers for all taxa included in this study.

Table S2 Primers used for PCR amplification and sequencing.

Table S3 Summary of results from the congruence analysis among chloroplast markers (see text).

Table S4 Model likelihood estimators obtained using the Path sampling (PS) and Stepping-Stone (SS) sampling methods implemented in BEAST.

Table S5 Geographical coordinates used in the Ecological Niche Modelling of *Canarina*.

Table S6 Mean ages and 95% HPD confidence intervals for the different BEAST analyses represented in Fig. 3a–b–c: (a) *Platycodoneae* dataset; (b) *Canarina* dataset, standard dating; (c) nested dating of the *C. eminii* and *C. canariensis* datasets.

Fig. S1 Bayesian Majority-Rule consensus trees obtained by MrBayes from the single-gene analyses of the Platycodoneae dataset.

Fig. S2 Bayesian Majority-Rule consensus trees inferred from the single-gene analyses of the *Canarina* dataset.

Fig. S3 Saturation plots for the single-gene nuclear (ITS) and chloroplast markers, showing the uncorrected pairwise distances (*p* distance) against corrected maximum likelihood distances (ML distance) derived in PAUP using the appropriate model of substitution.

Fig. S4 Maximum clade credibility (MCC) trees obtained from different BEAST analyses, showing nodes with mean ages and 95% HPD confidence intervals (values specified in Table S4): (a) Platycodoneae dataset; (b) standard dating of the *Canarina* dataset; (c) nested dating analysis of all three linked datasets: Platycodoneae (left) and population-level *C. eminii* and *C. canariensis* (right).

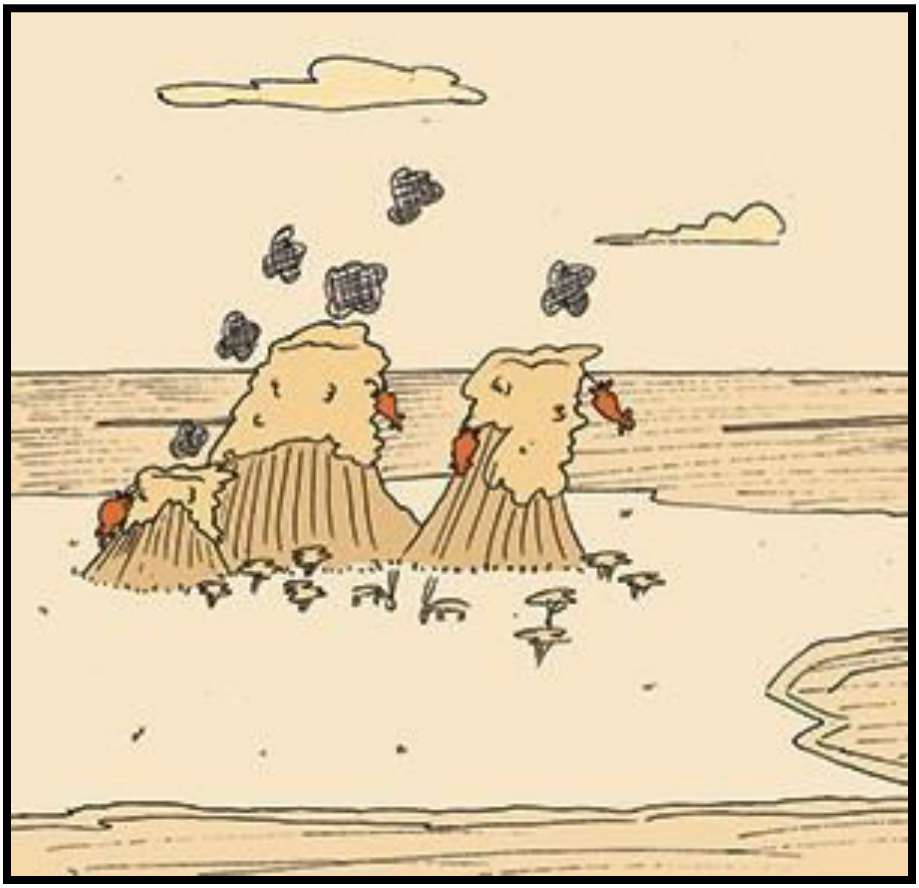
Fig. S5 Bayesian phylogeographic analysis of the *C. eminii* population-level dataset (nested dating approach) using an alternative coding of the geographic areas:(Elgon and Ruwenzori considered as different OTUs).

Fig. S6 Bayesian Majority-Rule consensus tree obtained by MrBayes from the *Canarina* concatenated chloroplast and nuclear dataset rooted with *Ostrowskia* (ITS, *psb*J-*pet*A, *trn*L-*trn*F, *pet*B-*pet*D, *trn*S-*trn*G, *trn*V-*ndh*C).

CHAPTER 2

Geographic barriers and Pleistocene climate change shaped patterns of genetic variation in the Eastern Afromontane Biodiversity Hotspot

Mario Mairal, Isabel Sanmartín, Alberto Herrero, Lisa Pokorny,
Pablo Vargas, Juan José Aldasoro, Marisa Alarcón



This chapter is in 2nd revision in the peer-review journal
Scientific Reports – Nature.

www.nature.com/scientificreports

SCIENTIFIC REPORTS

OPEN

Received: 18 October 2016
Accepted: 02 March 2017
Published: 11 April 2017

Geographic barriers and Pleistocene climate change shaped patterns of genetic variation in the Eastern Afromontane biodiversity hotspot

Mario Mairal[1], Isabel Sanmartín[1], Alberto Herrero[1], Lisa Pokorny[2], Pablo Vargas[1], Juan J. Aldasoro[3,4,*] & Marisa Alarcón[3,*]

The Eastern African Afromontane forest is getting increased attention in conservation studies because of its high endemicity levels and shrinking geographic distribution. Phylogeographic studies have found evidence of high levels of genetic variation structured across the Great Rift System. Here, we use the epiphytic plant species *Canarina eminii* to explore causal explanations for this pattern. Phylogeographic analyses were undertaken using plastid regions and AFLP fragments. Population genetic analyses, Statistical Parsimony, and Bayesian methods were used to infer genetic diversity, genealogical relationships, structure, gene flow barriers, and the spatiotemporal evolution of populations. A strong phylogeographic structure was found, with two reciprocally monophyletic lineages on each side of the Great Rift System, high genetic exclusivity, and restricted gene flow among mountain ranges. We explain this pattern by topographic and ecological changes driven by geological rifting in Eastern Africa. Subsequent genetic structure is attributed to Pleistocene climatic changes, in which sky-islands acted as long-term refuges and cradles of genetic diversity. Our study highlights the importance of climate change and geographic barriers associated with the African Rift System in shaping population genetic patterns, as well as the need to preserve the high levels of exclusive and critically endangered biodiversity harboured by current patches of the Afromontane forest.

The Afromontane Floristic Region[1,2] refers to a plant assemblage where a series of isolated highland forested areas are separated by lowland vegetation, and form an "archipelago-like" centre of endemism in the mountains of East and West Africa. The eastern part of this floristic region has been singled out as a biodiversity hotspot and is known as the Eastern Afromontane biodiversity hotspot[3] (EABH) (Fig. 1a). The Eastern African "sky islands" (including both high plateaus and isolated mountain peaks), are sometimes referred to as the "Galapagos of Africa"[4], and contain such flagship species as the mountain gorilla, the gelada, the Ethiopian wolf, or the coffee tree. These mountain ranges exhibit extraordinary levels of species endemism, for example, out of the > 10.000 plant species found in this region (7600 vascular plants), about one third are endemic[5]. This pattern of endemism has been attributed to long-term climatic stability, favouring lineage diversification and species accumulation over time[6]. The EABH is also one of the most threatened ecosystems in Africa[5]. Ethiopian montane forests, which two hundred years ago[7,8] covered up to 35% of the Afromontane areas, are estimated to have had their range dramatically reduced by human activity in the last century[9,10].

Geographically, the Eastern Afromontane region is divided longitudinally by the Great Rift System, a 4830 km north-south fissure whose terrestrial section extends from Djibouti to Mozambique. This orogenic fault marks the line along which the African Plate is splitting into two tectonic plates, the Somalian Plate and the Nubian Plate (Fig. 1a). The formation of the Rift commenced 31 Ma, as a result of the collision of the African and Arabian Plates

[1]Real Jardín Botánico (RJB-CSIC), 28014 Madrid, Spain. [2]Royal Botanic Gardens, Kew (RBGK), Richmond, Surrey, TW9 3DS, UK. [3]Instituto Botánico de Barcelona (IBB-CSIC-ICUB), 08038 Barcelona, Spain. [4]Universidad Rey Juan Carlos, Móstoles, Spain. *These authors contributed equally to this work. Correspondence and requests for materials should be addressed to M.M. (email: mariomairal@gmail.com)

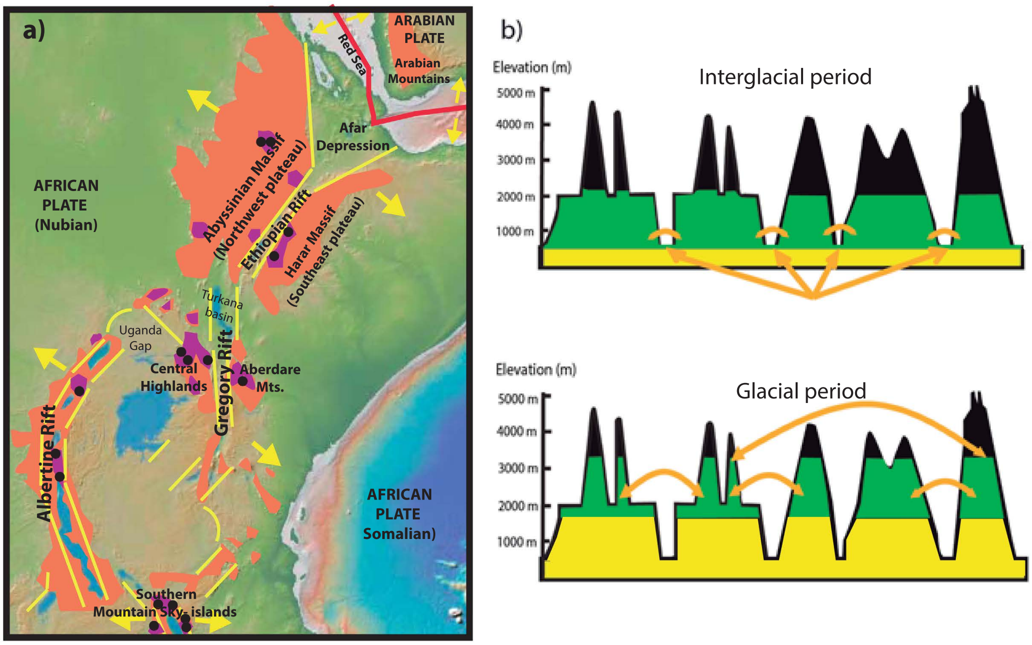


Figure 1. Geographic map of the Afromontane forests and main phylogeographic hypotheses. (**a**) Geographic map of Eastern Africa showing the main geographic features named in the text. The Eastern Afromontane biodiversity hotspot is shaded in orange. The distribution of *Canarina eminii* is shaded in purple. The Rift System is shown in yellow; lines represent major tectonic faults and arrows indicate the direction of relative divergent Rift movement. The red line shows the tectonic plate boundaries. The map was generated using the software GeoMapApp (v. 2.3) (http://www.geomapapp.org/)[85]. Black dots represent the sampling for *C. eminii.* (**b**) Sky-islands housing Afromontane forests (in green) separated by savannahs and open forests (in yellow), showing the two main phylogeographic hypotheses postulated to explain patterns of genetic variation in Afromontane organisms. Upper) the "mountain forest bridge" hypothesis (arrows indicate forest reconnections); Lower) the "long-distance dispersal" hypothesis (arrows indicate migration events).

with Eurasia; as extension continues, with ongoing volcanic activity[11], complete breakup of the Somalian Plate is predicted to occur within 10 million years, with a new ocean basin originating where shallow lakes can now be found[12]. The northern part of the Rift runs south from the Afar Triangle and is surrounded by the Ethiopian Highland Plateaus (Fig. 1a): the Abyssinian massif (northwest Plateau) and the Harar massif (southeast Plateau). In the south tropical East Africa, the Rift System is divided into two arcs: the Eastern Gregory Rift and the Western Albertine Rift, which join at the northern end of Lake Malawi, where the Southern Mountain Sky-Islands (SMSI) lie (Fig. 1a). Due to its elevation and extent, this latter region is sometimes known as the "Roof of Africa". Because of its geographic position, the Rift System is considered one of the most isolated mountain ranges in the world, compared to other tropical floras[13].

The Afromontane vegetation is divided into three altitudinal belts, each with unique floristic composition[14–16]: the lowermost Afromontane forest (1300–3000 m), the ericaceous belt (3000–4100 m), and the uppermost Afroalpine zone ($>$3550 m). These vegetation belts are not continuous in their distribution but appear divided into forest patches isolated from one another by environmental barriers (Fig. 1b), including extensive savannah areas, semi-deserts, lowland forests (in the past), or (at present) an agricultural matrix of cultivated lands and forestry plantations[1,17–20]. When compared to other mountain systems, the Afromontane vegetation is considered to be rather uniform in composition and is formed by a large number of broad-range species[13]; though given the fragmented nature of the Afromontane region, with numerous forest patches, some of these widespread species are likely to include cryptic diversity[21].

Interest on geographic patterns of genetic variance in Eastern African species has increased in recent years. Many of these phylogeographic studies (focusing on a diverse array of organisms, Table 1) have reported high levels of genetic variation structured with respect to the Great Rift Valley[8,19,22], suggesting it has acted as a dispersal barrier for an extended period of time.

The complex phylogeographic structure found in African highland species has also been explained by two non-mutually-exclusive dispersal-migration scenarios (Fig. 1b.): I) the "mountain-forest bridge" hypothesis postulates that current patterns might be explained by short-range or stepping-stone dispersal (SSD) between adjacent mountain ranges on each side of the Rift Valley, which would have been promoted by a more extensive forest coverage during Plio-Pleistocene warm and humid interglacial periods[8,17,23–25]; II) the "long-distance dispersal" hypothesis postulates that long-distance dispersal (LDD) events among isolated mountain populations during Pleistocene glacial periods are responsible for the observed phylogeographic patterns[18,19,26]. These two hypotheses are not mutually exclusive. For example, proponents of the SSD forest bridge hypothesis[17] suggested LDD migration among some isolated mountain ranges[18]. A requisite to discriminate between these two scenarios

Organism	**Phylogeographic disjunctions in the Ethiopian Rift**	**Phylogeographic disjunctions in the Volcanic arcs**	**Method (Reference)**
***Angiosperms* (habitat)**			
Coffea arabica (Afromontane)	Two distinct groups across the Ethiopian Rift	—	ISSR; microsatellites[50,86]
Lobelia giberroa (Afromontane)	Two distinct groups across the Ethiopian Rift: 1. Simien- Choke; 2. Chilallo-Bale-Gara Muleta	Two distinct groups across the Gregory Rift	AFLP[8]
Hagenia abyssinica (Afromontane)	Structure both sides of the Rift with rare long-distance dispersal events crossing the Rift	—	microsatellites[51]
Cordia africana (Afromontane)	Three different groups across the Ethiopian Rift: 1. Northwest plateau; 2. South-west Ethiopia; 3. Southeast plateau	—	AFLP and microsatellites[87]
Juniperus procera (Afromontane)	Two relatively different groups across the Ethiopian Rift: 1.Goba-Yabelo; 2. Chilimo-Suba-Ziquala-Washa	—	AFLP[88]
Prunus africana (Afromontane)	—	Two distinct groups across the Uganda Gap: 1. Albertine Rift and western Gregory Rift; 2. Eastern Gregory Rift	Plastid haplotypes, SSR[52,89]
Warburgia ugandensis (Afromontane transitional forest)	—	Two distinct groups across the Uganda Gap: 1. Albertine Rift and western Gregory Rift; 2. Eastern Gregory Rift	AFLP[90]
Erica trimera (Ericaceous)	No difference for AFLP, but different haplotypes	A complex pattern in the Gregory Rift	AFLP and plastid haplotypes[42]
Arabis alpina (Afroalpine)	Two distinct groups across the Ethiopian Rift: 1. Simien, Gara M.; 2. Choke, Elgon, Meru, Kilimanjaro	A complex pattern in the Gregory Rift	Plastid haplotypes[19]
Cardus schimperi (Afroalpine)	Two distinct groups across the Ethiopian Rift: 1.Simien; 2. Elgon-Aberdare-Bale	Mt. Kenya population is a different subspecies; while populations from Mt. Elgon and Aberdare are closely related	AFLP[26]
Trifolium cryptopodium (Afroalpine)	Two distinct groups across the Ethiopian Rift: 1. Simien Choke; 2. Bale-Aberdare–Elgon-Kilimanjaro	Populations related across the Rift Valley	AFLP[26]
Deschampsia cespitosa (Afroalpine)	—	Two distinct groups across the Rift Valley: 1. Rwenzori; 2. Kilimanjaro-Bale	AFLP[27]
Senegalia senegal (savannah, semi-desert)	—	Two distinct groups across the Gregory Rift with mixed population (Marigat): 1. Western side; 2. Eastern side.	microsatellite[91]
Senegalia mellifera (savannah)	—	Two distinct groups across the Gregory Rift: 1. Western side; 2. Eastern side; 3. Central Gregory Rift	microsatellite[49]
Vertebrates			
Ostrich *Struthio camelus*	Two subspecies both sides of the Rift (*molybdophanes*/*camelus*)	Two subspecies both sides of the Rift (*Molybdophanes*/*camelus*)	mtDNA[46]
Olive sunbird *Nectarinia*	—	Two distinct species across the Rift Valleys. 1. Western Rift Valley *Nectarinia obscura*. 2. Eastern Rift Valley *Nectarinia olivacea*	mtDNA[44]
Springhare *Pedetes capensis*	—	Two distinct groups across the Rift Valleys: eastern populations/southern populations	mtDNA[92]
Wildebeest *Connochaetes taurinus*	—	Two subspecies both sides of the Gregory Rift. 1. Western Gregory Rift: sbsp. *mearnsi* (Loliondo, Masai-Mara) 2. Eastern Gregory Rift: sbsp. *albojubatus* (Nairobi and Amboseli)	mtDNA[93]
African wild dog *Lycaon pictus*	—	Two distinct groups across the Gregory Rift. 1. Eastern clade. 2. Southern clade	mtDNA and microsatellites[94]
Sable antelope *Hippotragus niger*	—	Two distinct groups across the Gregory Rift: 1. West Tanzania and Kenya; 2. East Tanzania)	mitochondrial DNA and cyt. *b*[95]
Ethiopian wolf *Canis simensis*	Three distinct groups across the Ethiopian Rift: Northwest plateau (1. Wollo/Shoa; 2. Simien/Mt. Guna) and southeast plateau (3. Arsi/Bale)	—	mtDNA[22]
Lion *Panthera leo*	—	Two distinct groups across the Gregory Rift: 1. eastern (Tsavo-Transvaal); 2. western (Aberdare)	cytochrome *b* and NADH desh. subunit 5[96]
Continued			

Organism	Phylogeographic disjunctions in the Ethiopian Rift	Phylogeographic disjunctions in the Volcanic arcs	Method (Reference)
Baboon *Theropithecus gelada*	Two distinct groups across the Ethiopian Rift (northwest and southeast plateau)	—	RFLPs[97]
Grass mouse *Lemniscomys striatus*	—	Two distinct groups across the Rift Valleys: Gregory Rift/Albertine Rift	cytochrome *b*[98]
Rodent *Mastomys natalensis*	—	Two distinct groups across the Rift Valleys: eastern populations/southern populations	cytochrome *b*[99]
African clawed frogs (*Xenopus clivii* and *X. largeni*)	Two distinct groups across the Ethiopian Rift (northwest and southeast plateau)	—	mtDNA, autosomal loci[100]
Ethiopian anurans (*Tomopterna, Amietia, Leptopelis, Ptychadena*)	Two distinct groups across the Ethiopian Rift	—	Several mitochondrial and nuclear genes[101]
Insects			
Mosquito *Anopheles gambiae*	—	Two distinct groups across the Rift Valleys: eastern populations (Kimili, Asembo bay, Kisian, Awendo)/western populations (Malindi, Jego)	microsatellites[102–105]
Tsetse fly *Glossina pallidipes*	—	Two distinct groups across the Rift Valley: East Rift Valley (Dakabuko, Alangoshira, Shimba Hills, Kibwezi); West Rift Valley (Nguruman, Shompole, Marech)	Allozimes, microsatellites, mitochondrial loci[106]
Mosquito *Anopheles funestus*	—	Two distinct groups across the Rift Valley: western Kenya pop. (Mbita, Udhoro)/Coastal populations (Majajani, Magaoni)	microsatellites[107]
Termite *Scledorhinotermes lamanianus*	—	Two distinct groups across the Rift Valley: eastern populations/western populations	AFLPs[57]

Table 1. Phylogeographic studies of organisms showing genetic variation structured around the Rift System.

is a temporal framework. Most Eastern African phylogeographic studies in plants do not include time estimates for colonization events and, therefore, do not allow discriminating between the mountain-forest bridge and long-distance dispersal hypotheses. This is further aggravated by the difficulties intrinsic to collecting material of widespread Afromontane species, due to the complex geographic landscape and the recent political instability[27] (see below).

Here, we examine support for these two hypotheses, and the role of the Rift Valley as a long-term dispersal barrier, by reconstructing patterns of genetic variation in *Canarina eminii* Asch. & Schweinf, a strictly Afromontane species with a geographic distribution extending from the Ethiopian massifs to the SMSI in the northern end of Lake Malawi[28], along the Rift Valley. This species grows mostly as a twig epiphyte on trees, mainly of genera *Hagenia* J.F.Gmel., *Conopharyngia* G. Don, and *Afrocarpus* (J. Buchholz & E.G. Gray) C.N. Page[28], all characteristic elements of the Afromontane forest belt[28]. Little is known about the reproductive biology of *C. eminii*, except that the species is pollinated by sunbirds (unpublished field observations) and exhibits floral traits that have been associated with this type of ornitophilous pollination[29] (e.g., large orange-red coloured flowers producing abundant diluted nectar). Fruits are fleshy, with the seeds enclosed in a sticky sweet jelly, suggesting endozoochorous dispersal (e.g., by monkeys[30]). Mairal *et al.*[31] recently reconstructed the biogeographic history of *Canarina* L., a small genus of three species within tribe Platycodoneae (family Campanulaceae). This genus exhibits a wide disjunct distribution across North Africa and probably originated from a Central Asian ancestor, which arrived to Eastern Africa in the Mid Miocene (13.7 Ma) through the Arabian Plate[31]. The lineage leading to the Eastern African species *C. abyssinica* diverged first, followed by the split between *C. eminii* and the endemic Canarian species *C. canariensis*[31] in the Late Miocene (c. 7 Ma). Divergence within *C. eminii* was traced back to 1.76 Ma[31] (southern range of the Great Rift Valley and followed by subsequent colonisation of the northern range). However, the intraspecific sampling in the study by Mairal *et al.*[31] was limited, which prevented analysis of patterns of genetic variation or examination of the processes population divergence in *C. eminii*.

In this study, we reconstruct the phylogeographic history of *C. eminii* among and within populations. Our aims were to: i) describe the geographic distribution of genetic variation within this species; ii) examine the role of the Rift Valley as a long-term dispersal barrier; iii) understand how Pleistocene climatic fluctuations might have affected population ranges, especially in relation to the "mountain-forest bridge" hypothesis and the role of isolated "sky islands"; and iv) provide insights into the phylogeography of an Afromontane epiphytic species. Since the epiphytic growth of this species is tied to dominant tree species of the Afromontane forests, it makes *C. eminii* a good case study for tracing the history of fragmentation and expansion of forest coverage in this region.

Results

Haplotype analyses. In order to examine patterns of genetic variation within *C. eminii*, we sequenced three highly variable plastid (pDNA) regions covering the entire distributional range of this species (see Materials and Methods). We generated 237 new sequences from 79 individuals: *rpl*32-*trn*LUAG (79 sequences), *trn*SGCU–*trn*GUCC (79 sequences), and *pet*B1365–*pet*D738 (79 sequences). The number of nucleotide sites ranged from 648

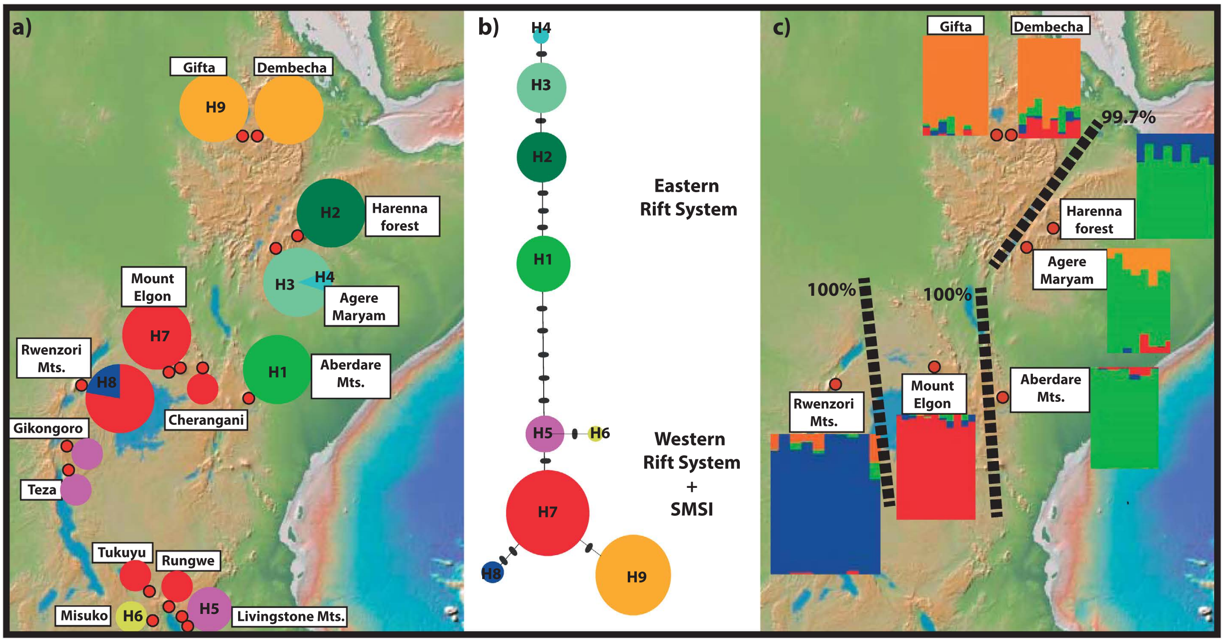


Figure 2. Plastid and nuclear datasets analysed for *Canarina eminii*. (**a**) Haplotype distribution. Red dots represent the geographic location of populations and pie charts show the frequency of occurrence of each haplotype. (**b**) Statistical Parsimony network inferred using the DNA plastid sequences by TCS. Black dashes on long connecting lines indicate nucleotide changes. Circle size is proportional to the frequency of haplotypes. Each haplotype is shown in a different colour, where codes (H1 to H9) correspond to the haplotypes shown in Fig. 2a. (**c**) Phylogroups using AFLP markers. Histograms showing the Bayesian clustering of individuals within populations (STRUCTURE); colours represent the proportion of individual membership to each inferred Bayesian group. Dashed lines indicate barriers to gene flow and their percentage, as inferred by BARRIER. The map was generated using the software GeoMapApp (v. 2.3) (http://www.geomapapp.org/)[85].

in *trn*SGCU–*trn*GUCC, 834 in *pet* B$_{1365}$–*pet*D$_{738}$, to 963 in *rpl*32-*trn*L$_{UAG}$. The concatenated matrix consisted of 79 sequences of *C. eminii* with 2445 nucleotide sites. The haplotype network recovered by TCS identified nine haplotypes (H1 to H9; Fig. 2a), divided into two main groups separated by five substitutions (Fig. 2b): one group east of the Rift Valley (Harenna Forest, Agere Maryam, and Aberdare Mts.), and a second group including populations of the Abyssinian massif, west of the Gregory Rifts, the Albertine Rift, and the SMSI. All haplotypes, save H5 and H7, were restricted to one population. Lineage divergence time estimation in the Bayesian software BEAST (Fig. 3, Supplementary Fig. S2.1) supported a similar geographic structure. The first population divergence event (1.92 Ma, 95% high posterior density (HPD) credibility interval (CI): 0.63–3.82 Ma) separated the populations situated at each side of the Rift Valley, with subsequent divergence events (starting 0.4 Ma) dividing populations located within each side of the Rift System (Fig. 3). We performed a phylogeographic analysis using Bayesian ancestral state reconstruction methods in BEAST (Fig. 4a,b); these were not conclusive, supporting a geographic origin of *C. eminii* in the Harar Massif or the SMSI, on each side of the main Rift System, with nearly equal probability. Nine migration routes were inferred by Bayesian Stochastic Search Variable Selection (BSSVS, see M&M) (Table S2.5); these seem to be arranged into parallel routes, following migration from north to south and vice versa on each side of the Rift System, and connecting also the intermediate mountain ranges (Fig. 4b).

AFLP polymorphism, genetic diversity and structure. To complement the plastid signal (above), we carried out an analysis of fragment length polymorphism (AFLP) in the nuclear compartment. The final data set after scoring comprised 773 loci. The Bayesian software STRUCTURE assigned individuals to four geographic clusters based on patterns of genetic variance ($K = 4$, Fig. 2c; Fig. S2.2). These clusters were broadly coincident with those recovered by the haplotype network: i) a cluster of populations on the eastern side of the main Rift System (Harenna Forest, Agere Maryam and Aberdare Mts.), ii) the Abyssinian Plateau, iii) the Central Highlands, and iv) the Albertine Rift (Rwenzori Mts.). Neighbour-joining and neighbour-net diagrams (Fig. S2.3a,b) identified six well-supported groups coincident with main mountain ranges within the Rift, except for a group (Agere Maryam) that had a low bootstrap value (72.6%). The two populations of the Abyssinian Plateau were also genetically very close. Principal Component Analysis (PCO) of genetic variance also differentiated six groups: those of the eastern Rift occupied the centre of the plot, with the remaining located on the periphery (Fig. S2.4). Genetic differentiation among populations (Fst values) was higher among the western populations (Debre Markos, Mt. Elgon and Rwenzori, 0.28–0.38) than among those in the eastern side (Harenna Forest, Agere Maryam and Aberdare Mts., 0.17–0.21; Table S2.6), suggesting greater isolation among the populations west of the Rift than between those in the east. East of the Rift Valley System, a rarity index (frequency-down-weighted marker value, DW) was highest for the Harar Massif population, which also presented the highest number of

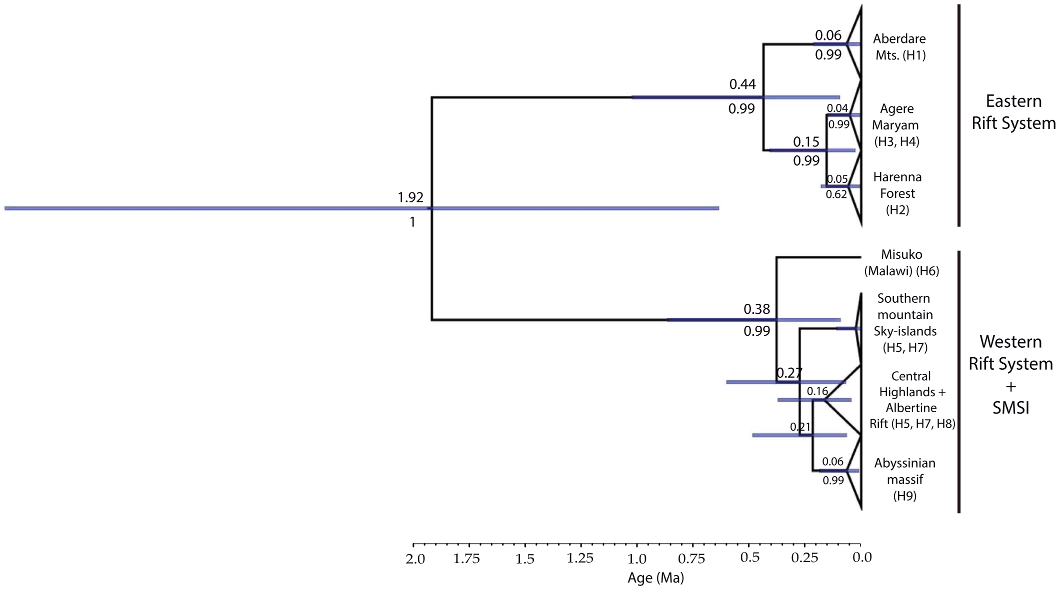


Figure 3. Maximum clade credibility (MCC) tree obtained from the BEAST analysis of pDNA haplotypes of *Canarina eminii*. Blue bars show 95% HPD credibility intervals. Numbers above branches show mean ages and numbers below branches indicate Bayesian posterior clade support values.

polymorphic and private fingerprints (Table 2). Populations in the Central Highlands (Mt. Elgon and Cherangani Hills) had the highest DW values and a large number of polymorphic fragments, plus two fixed fragments. Populations west of the Rift Valley showed generally lower values of genetic diversity and no fixed fragments. No linear relationship was found between pairwise F_{ST} and geographic distance (Mantel analysis, $r = 0.1884$, $p = 0.5012$; Fig. S2.5). However, the correlation increased after excluding Mt. Elgon's population from the analysis ($r = 0.6198$, $p = 0.0560$; see black stars in Supplementary Fig. S2.5). This population showed the largest genetic distance in pairwise-comparisons in relation to all other populations (Table 2.6, see Supplementary Fig. S2.5). In contrast, the population of the Abyssinian massif, which is separated by 500–1300 km from the southernmost populations (Mt. Elgon, Rwenzori), showed much lower pairwise F_{ST} values (Table S2.6). All intervals analysed with SPAGeDi (see Supplementary Fig. S2.6) gave a significant negative result, suggesting that populations are not more different with increasing distance. In agreement with this, BARRIER detected three major boundaries separating the four STRUCTURE clusters: i) Rwenzori Mountains from Mt. Elgon (100%), ii) Mt. Elgon from all areas in the east (100%), and iii) Abyssinian massif from the remaining areas (99.7%) (Fig. 2c). Hierarchical AMOVA showed the greatest genetic variance among the same four groups identified by STRUCTURE and BARRIER (Table S2.4). The relatively low among-group genetic variance (Table S2.4) might be explained by the limited number of populations sampled at each side of the Rift, and therefore should be taken with caution. Yet, the geographic structure recovered by AMOVA is fully congruent with the results from the Bayesian clustering (STRUCTURE and BARRIER) and with the Fst and DW estimates, lending support to this analysis.

Discussion

The BEAST "nested dating" approach used here provided divergence time estimates similar to those in Mairal *et al.*[31] for the origin of *Canarina* (8.98 My.) and the stem-age of *C. eminii* (7.14 My.) (Fig. S2.1). The first divergence event within *C. eminii* (ca. 1.92 Ma, 95% HPD 0.63–3.82 Ma, Fig. 3) postdates the age of establishment of the Afromontane forest[16] (Early Pleistocene, ca. 2.4 Ma[32]), and is roughly coincidental with a period of aridification in East Africa, peaking at 1.7 Ma[33,34]. Our results also support that divergence in *C. eminii* followed the opening of the Rift Valley in the Early Pliocene, which divided the Rift System into two landstrips[35].

A strong phylogeographic structure with the Great Rift Valley at its central axis was recovered for *C. eminii* populations. This agrees well with phylogeographic studies in other Afromontane groups (i.e, angiosperms, insects, and vertebrates, Table 1), all supporting the Rift System as an effective barrier to gene flow among populations. The depth of the phylogeographic split among STRUCTURE groups within *C. eminii* (Fig. 2c), the distribution of haplotypes (Fig. 2a,b), and the pattern of reciprocal monophyly and deep divergence times (1.92 Ma, HPD: 0.63–3.82 Ma) inferred with BEAST (Fig. 3), suggest a history of long-term isolation among populations east and west of the Rift Valley. In contrast, the large geographic distances and comparatively lower genetic distances between populations located on either side of the Rift Valley, suggest that geographic distance played a minor role in structuring patterns of genetic variation within *C. eminii*, confirmed by the lack of correlation the Mantel test establishes.

Patterns of colonization and biotic assemblage in the Eastern African Mountains have been argued to follow a northern (an "African pan-temperate element"[28,36,37]) and a southern (an "Afromontane track"[38–40]) component. Migrants coming from the north would have needed to overcome the widest point of the Rift Valley, the Afar Triangle (Fig. 1a), which acts as a major division before reaching the Ethiopian Plateaus, while southern

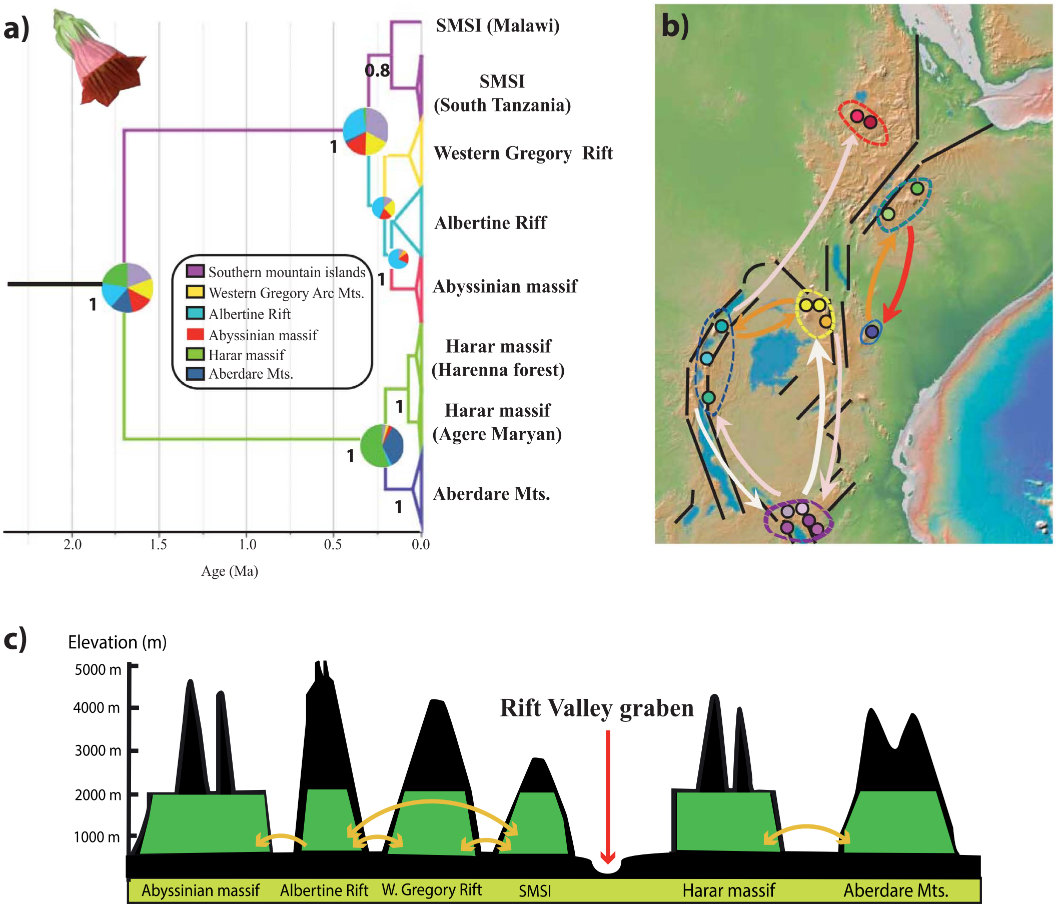


Figure 4. Phylogeographic analysis and reconstruction of the colonization of *Canarina eminii*. (**a**) BEAST MCC tree showing the Bayesian ancestral range reconstruction analysis[63]. Coloured branches (see legend) represent the ancestral range with the highest posterior probability for each population; node pie charts show marginal probabilities for alternative ancestral ranges. Numbers below branches represent Bayesian posterior probabilities. (**b**) Map representing migration events that receive a BF support > 3, as recovered by BSSVS; colour tint is proportional to the support (dark red > orange > pale pink > white). The map was generated using the software GeoMapApp (v. 2.3) (http://www.geomapapp.org/)[85] (**c**) A hypothetical reconstruction of the colonization of the Afromontane forest by *Canarina eminii* during interglacial periods based on our data and literature of other Afromontane groups. Orange arrows show possible short to medium distance dispersal events between isolated mountain ranges or across forest galleries. Expansion of the Afromontane forest (in green) during interglacial periods may have facilitated migration across Stepping Stone dispersal or forest bridges.

migrants probably needed to split into two routes when reaching the two parallel volcanic arcs of the Albertine and the Gregory Rift[39] (Fig. 1a). Thus, dispersal into the Eastern African Mountains from the north or from the south would have been forced by topography to follow two parallel routes, one on each side of the Rift Valley[19,22]. This scenario is supported by the Discrete Phylogeographic Approach (DPA) analysis of *C. eminii*, showing two colonization routes that split along the topography of the Rift System (Fig. 4b). Though this method has been recently criticized as being too decisive[41], other studies in Afromontane taxa have found similar phylogeographic connections through the east[26,27,42] and west[8,43,44] (Table 1).

Which mechanisms are behind this pattern of geographic isolation between the two sides of the Rift Valley for *C. eminii* and other Afromontane taxa? As mentioned above, the origin of *C. eminii* considerably postdates the formation of the Rift, so strict vicariance is not feasible (nor for any of the groups in Table 1). However, the major faults and flooding of graben associated to the Rift activity, may have acted as a corridor filtering migration, spatially restricting dispersal to two parallel paths on either side of the Rift Valley. Indeed, the volcanic desert of Afar in Ethiopia (the hottest place on Earth), and the drier floor and deep lakes at the bottom of Rift Valley, are ecologically very different from the surrounding higher elevation areas[45,46], and constitute ecological barriers to dispersal (i.e., hostile environments that are beyond the tolerance limits of montane organisms). It would be interesting to test this hypothesis using genetic and ecological evidence in other lineages exhibiting a similar pattern of genetic isolation along the Rift Valley (Table 1).

Eastern Africa has experienced several climate fluctuations over the last 800,000 years, resulting in mountain forests extending and retracting successively, with further isolation and bottlenecking events for some species[47,48]. The Afromontane forest probably did not descend far enough to reach the bottom of the Rift Valley

Haplotypes							**AFLPs**					
Population	**Mountain range**	**N° samples**	**Haplotypes**	**H (n)**	**H (d)**	**π**	**N° samples**	**N° of polymorphic fragments (% in brackets)**	**Hj (se)**	**N° of private fragments (% in brackets)**	**N° of fixed private fragments**	**DW index**
West of the Rift												
Gifta, Debre Markos (Ethiopia)	Abyssinian massif (Northwest Plateau)	8	H9	1	0	0	8	477 (61.7)	0.244 (0.0071)	8 (1.68)	0	93,187
Dembecha, Debre Markos (Ethiopia)	Abyssinian massif (Northwest Plateau)	7	H9	1	0	0	8	443 (57.3)	0.208 (0.0069)	6 (1.35)	0	99,213
Rwenzori Mts. (Uganda)	Albertine Rift	12	H7, H8	2	0.303	0,00026	10	417 (53.9)	0.193 (0.007)	6 (1.439)	0	111,177
Gikongoro-Teza (Rwanda-Burundi)	Albertine Rift	2	H5	1	0	0	–	—	—	—	—	—
Central highlands												
Mt. Elgon 1 & 2-Cherangani Hills (Kenya)	Central sky-islands	19	H7	1	0	0	11	477 (61.7)	0.183 (0.0063)	5 (1.048)	2	140,694
East of the Rift												
Agere Maryam, Yirga (Ethiopia)	Harar massif (Southeast Plateau)	9	H3, H4	2	0.222	0.0001	8	495 (64.0)	0.241 (0.0067)	13 (2.626)	1	108,327
Harenna Forest, (Ethiopia)	Harar massif (Southeast Plateau)	8	H2	1	0	0	9	491 (63.5)	0.224 (0.0067)	10 (2.037)	2	128,342
Aberdare Mts. (Kenya)	Aberdare Mts. (sky-island)	9	H1	1	0	0	7	435 (56.3)	0.208 (0.0067)	8 (1.839)	2	91,057
Southern Mountain Sky-Islands												
Misuku-Tukuyu-Rungwe-Livingstone Mts. (Tanzania-Malawi)	Southern Mountain Sky-islands	5	H5, H6, H7	3	0,4	0,00032	—	—	—	—	—	—

Table 2. Descriptors of within-population genetic diversity in the cpDNA haplotypes and AFLP markers for each population studied of *Canarina eminii*. Abbreviations: H(n): number of haplotypes; H(d): haplotype diversity; π: nucleotide diversity; Hj (se): Nei's gene diversity (standard error); DW: frequency-down-weighted value.

graben[25] during the humid-climate expansion periods, and a widespread montane forest only seems likely on the Ethiopian plateaus, on either side of the Rift Valley[25]. However, some species might have had opportunities for gradual SSD migration through grasslands and open forest corridors between mountains[25]. This would have resulted in sister phylogroups occupying separate mountain ranges on either side of the Rift Valley, which could have been followed by *in situ* diversification during the arid glacial periods[25,49]. Patterns of genetic variation in *C. eminii* agree with this hypothesis, suggesting long-term isolation of populations in mountain enclaves on each side of the African Rift. High genetic differentiation and exclusiveness (Fig. 2a–c, Table S2.4, see Supplementary Fig. S2.3 and S2.4) was observed among populations on each sky-island, in agreement with other Afromontane phylogeographic studies[8,50–52] and those in the Ericaceous[42] and Afroalpine[26,27,53] vegetation belts. In contrast to the glaciation-induced latitudinal range shifts in European taxa, which apparently resulted in a loss of genetic diversity (i.e., extinction and bottlenecks[54]), repeated events of fragmentation and reconnection during Pleistocene glacial/interglacial cycles may have contributed to the accumulation and maintenance of genetic diversity in the Eastern African mountains, and could explain their current status as centres of endemicity[52,53]. This agrees well with the hypothesis of a more moderate effect of glacial cycles at equatorial latitudes, in contrast with regions at higher latitudes[54].

Another explanation for the east/west connectivity pattern and the observed low genetic variance between populations on each side of the Rift in *C. eminii* and other Afromontane endemics, is long-distance dispersal of seed and pollen through vectors such as wind, insects, or birds; this in turn could have been reinforced by wider forest coverage in the past, i.e., greater area availability for seedling establishment[25]. Long-range seed dispersal is likely the best explanation for connections among populations of *C. eminii* across the Turkana Basin and the Uganda Gap, which probably constitute ecological barriers (climatically unfavourable land) for gradual SSD migration through forest connections. Results from the plastid DNA suggest seed dispersal on each side of the

Rift Valley, with more recent connectivity among the eastern side populations (Figs 2a and 3). On the other hand, the strong interpopulation structure found in the nuclear AFLP data suggests limited pollen dispersal (Fig. 2c). This could be explained by the strong territoriality and reduced mobility of Nectariniidae birds, the presumed pollinators of *C. eminii*. Some recent studies on these birds have revealed cryptic species within the same lineage occupying different mountain ranges, probably due to low connectivity between populations[44,55]. This pattern might have later been reinforced by a landscape strongly transformed by agriculture.

Both plastid and nuclear genetic data suggest a pattern of greater historical isolation among populations located in the western side of the Rift Valley than among those in the east (Fig. 2, Table 2). Whether this is associated to greater past connectivity through forest patches (mountain-bridge hypothesis) in the east, or whether it reflects a more recent historical pattern associated to agricultural activity, is not clear. Similar to other studies[8,50,51], the highest level of genetic diversity was found in the Harar Massif, an area known for harbouring the best-preserved fragments of Afromontane forest (Table 2: Agere Maryam and Harenna Forest). This area may have acted as an important refuge for Afromontane forest organisms during glacial maxima. In the DPA analysis of *C. eminii*, the Harar Massif was inferred as the source of migration events on the eastern side of the Rift (Fig. 4b). For the western Rift, the SMSI was suggested as the ancestral area: these mountains harbour the highest plastid genetic diversity (Table 2), though these results could not be corroborated with the nuclear data (no AFLP sampling). The two arcs of the Rift Valley (Albertine and Gregory Rift) meet geographically in the north on the Central Highlands (Mt. Elgon and Cherangani Hills) and in the South in the SMSI (Fig. 1). These two regions (Mt. Elgon- Cherangani and the SMSI) apparently act as secondary contact points where the two parallel migration routes mentioned above interconnect[19,44], and have been described as cradles of new genetic diversity[53]. This is congruent with our results in *C. eminii*, showing Mt. Elgon and the SMSI as a crossroad linking both sides of the west Rift System in the DPA analysis (Fig. 4), and both exhibiting high levels of genetic diversity (Table 2). Mount Elgon is also the oldest and most species-rich volcano of Eastern Africa[13,56], with endemic lineages that are related to groups in either the eastern[43] or the western side of the Rift Valley[8,43]. In contrast, the Afar Depression and the arid zone around the Turkana Basin seem to have acted as major barriers against gene flow (Fig. 1a and 4b). The Afar Depression separates the eastern and western Ethiopian populations, north of the two volcanic arcs, while the aridity around the Turkana Basin and the Uganda Gap, might have constituted an important obstacle to gene flow between the eastern and western populations[57]. The low haplotype diversity found in the Abyssinian massif population could reflect long-term isolation due to the influence of glaciations, though a low DW value suggests that recent dispersal could be a more likely explanation.

In summary, our study suggests that the topography and land formations of the Great Rift Valley underlie the phylogeographic pattern of east/west vicariance observed in *C. eminii* and other Afromontane endemic species. Though population sampling was limited to discriminate between the "mountain-forest" bridge and the LDD hypotheses, Pleistocene climatic oscillations and cyclical expansion and contraction of the Afromontane forest appear to have played a role in structuring levels of genetic variation in these species. Short-range dispersal among forest patches on each side of the Rift probably was complemented by LDD across some regions such as the Turkana Basin or the Uganda Gap, historically devoid of montane forests. Isolated sky-islands such as Mount Elgon or the Harar Massif acted as both refuges and cradles of genetic diversity. Given that the forested areas of Eastern Africa are currently in serious decline[9] —especially the northern Ethiopian highlands subject to extensive logging— and that many of these harbour high levels of genetic variability (see Table 1, our study), designing new measures for the protection of these regions (e.g., new natural parks) may be crucial to preserve the endemic and highly-threatened Afromontane forest biota.

Methods

Sampling and DNA sequencing.

The geographic distribution of *C. eminii* extends from the Ethiopian massifs to southern Tanganyika[28]. Despite our sampling effort (two field expeditions in 2009, 2015 to Ethiopia and one in 2010 to Kenya-Uganda), we failed to find *C. eminii* in many of the localities recorded by Hedberg in his classic monography of this species[28]. Much of the forest in these localities is highly degraded, threatened by the advancement of extensive agriculture, cattle rising, forestry, and other human activities (see Table S2.1 for a list of the visited localities for this study and a detailed description of the sampling). Despite this, we were able to sample fresh plant tissue from seven populations of *C. eminii* (Table S2.2): two populations in the Abyssinian massif (Gifta and Dembecha), two in the Harar Massif (Harenna forest and Agere Maryam), one in Rwenzori Mts. (Albertine Rift), one in Mt. Elgon (west of the Gregory Rift), and one in Aberdare Mts. (east of the Gregory Rift) (Fig. 2). Whenever possible leaves from 8–17 plants, 7–15 m apart were collected from each population. In addition, we added nine localities, obtained from old herbarium specimens (Royal Botanic Gardens, Kew): Cherangani (Kenya), Sasa trail in Mt. Elgon (Uganda), Gikongoro (Rwanda), Teza (Burundi), Tukuyu, Mwakelele, Livingstone Mountains, Mporoto Mountains (Tanzania) and Misuku (Malawi). Altogether, we sampled 16 localities of *C. eminii* (Table S2.2), representing all major geographic blocks in the distribution of the species (Fig. 1a). Three plastid (pDNA) regions were sequenced for a total of 79 individuals. For details on PCR amplification and sequence alignment see Appendix S1.1 and Table S2.3 in the Supplementary Material. Sources of the material examined, location of vouchers, GenBank accession numbers, and full references are listed in Table S2.2.

Phylogeographic analysis.

Haplotype analyses were done on a concatenated dataset comprising all three sequenced plastid regions (see above). Genealogical relationships among haplotypes were inferred via the statistical parsimony algorithm[58] implemented in TCS 1.2.1[59]. The number of mutational steps resulting from single substitutions among haplotypes was calculated with 95% confidence limits, with gaps coded as missing data. Summary statistics for genetic diversity calculated for each population were: number of haplotypes H (n), haplotype diversity (Hd), and nucleotide diversity (π); using DnaSP 5.1 with the option "not consider gaps" selected[60].

Haplotype divergence times were estimated using Bayesian relaxed clocks implemented in the software package BEAST v.1.7.5[61] and using the 'nested dating approach' described in Mairal *et al.*[31]; in this approach, a high phylogenetic-level dataset, including representatives of all three species of *Canarina* and nine outgroup taxa from Platycodoneae and Campanulaceae, was used to inform the clock rate of a linked population-level dataset of *C. eminii*, under a mixed Yule-coalescent model[62]. Congruence among results from the three plastid markers was tested by comparing clade support values for individual clades. See Appendix S1.2 for more details on these analyses.

The Discrete Phylogeographic Approach (DPA) of Lemey *et al.*[63], implemented in BEAST, was used to infer ancestral ranges and to trace the history of migration events in *C. eminii*. It is based on a continuous-time Markov Chain process where the discrete states correspond to the geographic locations and the state transition rates to the migration rates between areas (Ronquist & Sanmartín, 2011). Bayesian Stochastic Search Variable Selection (BSSVS, Lemey *et al.*[63]) was used to identify the rates (colonization routes) that are best supported by the data, using a cut off value of three for the Bayes Factor comparison. We defined six discrete areas, assigning each plateau and Rift mountain ranges to a different area[15,31]: 1) Harar massif, 2) Abyssinian massif, 3) East Gregory Rift (Aberdare Mountains), 4) Central Highlands (Mount Elgon and Cherangani Hills), 5) Albertine Rift (Rwenzori, Rwanda and Burundi), and 6) Southern Mountain Sky-Islands (SMSI: South Tanzania and Malawi) (see Appendix S1.3 for more details).

AFLP fingerprinting. Laboratory molecular protocols for the AFLP analysis[64] were implemented using the AFLP plant mapping kit (Applied Biosystems). We could not use the individuals sampled from herbarium collections for AFLP fingerprinting, as this approach is quite sensitive to the quality of the starting DNA and thus requires well-preserved material[65]. Therefore, sampling of populations in the nuclear dataset was more reduced than in the pDNA sequence dataset. Genomic DNA was digested with the enzymes *Eco*RI and *Mse*I and linked to the adaptors *Eco*RI (5'-CTCGTAGACTGCGTACC-3'/5'-AATTGGTACGCAGTCTAC-3') and *Mse*I (5'- GACGATGAGTCCTGAC-3'/5'-ATCTCAGGACTCAT-3'). We tested 32 combinations of selective primers and chose the four pairs that produced the most polymorphic and clear profiles: 1-*Eco*RI$_{6\text{-FAM}}$-ACT/*Mse*I-CAA; 2-*Eco*RI$_{6\text{-FAM}}$-ACT/*Mse*I-CAT; 3-*Eco*RI$_{\text{VIC}}$-AGG/*Mse*I-CTA, and 4-*Eco*RI$_{\text{VIC}}$-AGG/*Mse*I-CTT.

Amplified fragments were analysed using GeneMapper 3.7 software (Applied Biosystems), and peaks ranging between 100 and 500 bp recorded. AFLP Scorer software[66] was employed to run a reproducibility test, with the maximum acceptable error for each primer combination fixed at $< 5\%$[67]. The AFLPdat R package[68] was used to determine the numbers of private fragments per population. Only unambiguous fragments shared among duplicates were scored. Data reliability was assessed by comparison of duplicates, using one or two individuals per population (21 tests). The reproducibility value was 88–100%, with a mean of 97.5%.

AFLP data analysis. The resulting AFLP presence/absence matrix was analysed using AFLPSURV v.1.0[69] to estimate Nei's gene diversity (Hj), pairwise differentiation among subpopulations (F_{ST}), the percentage of polymorphic fragments per population (P), and the bootstrapped Nei's genetic distance matrix between individuals and populations[70,71]. The inbreeding coefficient (F_{IS}) was set to 0.1 as suggested by Hardy[72]. The permutation test involved 20,000 permutations. In addition, a Bayesian method was employed to estimate allelic frequencies, using a non-uniform prior distribution[73]. Ten thousand permutations were performed to calculate F_{ST} values. Genetic distances between individuals, populations and geographic groups were also calculated. We used AFLPdat[74] to calculate DW value per population, equivalent to the weighted endemism value[75,76]; this value is expected to be high in long-term isolated populations where rare markers should accumulate due to mutations, whereas newly established populations are expected to exhibit low values, thus helping to differentiate recent dispersal from more ancient isolation.

To distinguish genetic groups of individuals in the AFLP dataset, a comparison was made by constructing a pairwise similarity matrix for all individuals: Dice's coefficient was calculated and the resulting matrix was transformed using principal coordinates analysis (PCO) with Ntsys v.2.1[77]. Neighbour-nets of AFLP data were also calculated, both for individuals and populations, using the SplitsTree v.4.10 software[78]. To quantify the amount of genetic differentiation attributable to geographic and population subdivision, a hierarchical analysis of molecular variance was performed[79] in ARLEQUIN v.3.0 (Table S2.4). For assessing the structure of populations, we used the Bayesian method implemented in STRUCTURE 2.2[80,81], assuming admixture conditions and uncorrelated allele frequencies between groups, 500,000 generations (plus a burn-in of 100,000) were run for *K* values of 1–10, with ten repetitions each. For each *K* value, only the run with the highest maximum likelihood value was considered. The LnP (D) for the successive decomposition of groups was used in all STRUCTURE analyses[82]. To test the effect of the spatial distance on the genetic structure of the *C. eminii* populations, correlations between genetic (measured as F_{ST}) and spatial distances between pairs of populations were determined using the Mantel permutation procedure in Ntsys. The genetic distance matrix used was based on the presence/absence matrix; the geographic distance matrix was based on absolute distances between the geographic coordinates for each collected population. In addition, the kinship multilocus coefficient (F_{IJ}) was estimated using SPAGeDi[83] to determine the spatial structure of the examined populations, taking spatial distances into account in the analyses. BARRIER v.2.2[84] was used to identify possible geographic locations acting as major genetic barriers among *C. eminii* populations, based on genetic distances. The significance of these was examined with 1000 bootstrapped distance matrices obtained using AFLPsurv.

Data Accessibility. DNA sequences: Genbank Accession nos KF028817–KM189329. GenBank accessions, sampling locations and/or online-only appendices uploaded as online Supplemental material.

References

1. White, F. The history of the Afromontane archipelago and the scientific need for its conservation. *African Journal of Ecology*. **19,** 33–54 (1981).
2. White, F. The Vegetation of Africa. A Descriptive Memoir to Accompany The Unesco/AETFAT/UNSO Vegetation Map of Africa (3 Plates, Northwestern Africa, Northeastern Africa, and Southern Africa, 1: 5,000,000) (1983).
3. Mittermeier, R. A. *et al.* Hotspots revisited: earth's biologically richest and most endangered ecoregions. CEMEX, Mexico City, Mexico (2004).
4. Kingdon, J. *Island Africa: The Evolution Of Africa's Rare Animals And Plants* (ed. Collins, London, 1990).
5. C.E.P.F., *Eastern Afromontane Biodiversity Hotspot - Ecosystem Profile*, http://www.cepf.net/Documents/Eastern_Afromontane_Ecosystem_Profile_FINAL.pdf, (2012) (11/10/16).
6. Fjeldså, J. & Lovett, J. C. Geographical patterns of old and young species in African forest biota: the significance specific montane areas as evolutionary centres. *Biodiversity & Conservation*. **6,** 325–346 (1997).
7. EFAP. *Ethiopian forestry action program*: *Final Report - Volume II: The Challenges For Development*, Ministry of Natural Resources Development and Environmental Protection, Addis Ababa, Ethiopia (1994).
8. Kebede, M., Ehrich, D., Taberlet, P., Nemomissa, S. & Brochmann, C. Phylogeography and conservation genetics of a giant lobelia (*Lobelia giberroa*) in Ethiopian and Tropical East African mountains. *Molecular Ecology*. **16,** 1233–1243 (2007).
9. F.A.O., Food and Agricultural Organization of the United Nations. Forest Resource Assessment, Page 4, http://www.fao.org/3/a-i4808e.pdf (2015) (11/10/16).
10. Kloos, H. & Legesse, W. Water Resources Management In Ethiopia: Implications For The Nile Basin. Cambria Press (2010).
11. Ring, U. The east African Rift System. *Austrian Journal of Earth Sciences*. **107,** 132–146 (2014).
12. Fernandes, R. M. S. *et al.* Angular velocities of Nubia and Somalia from continuous GPS data: implications on present-day relative kinematics. *Earth and Planetary Science Letters*. **222,** 197–208 (2004).
13. Sklenář P., Hedberg, I. & Cleef, A. M. Island biogeography of tropical alpine floras. *Journal of Biogeography*. **41,** 287–297 (2014).
14. Hedberg, O. Vegetation belts of the East African mountains. *Svensk Botanisk Tidskrift Utgifven af Svenska Botaniska Foreningen, Stockholm*. **45,** 140–202 (1951).
15. Gehrke, B. & Linder, H. P. Species richness, endemism and species composition in the tropical Afroalpine flora. *Alpine Botany*. **124,** 165–177 (2014).
16. Linder, H. P. The evolution of African plant diversity. *Frontiers in Ecology and Evolution*. **2,** 1–14 (2014).
17. Hedberg, O. Evolution and speciation in a tropical high mountain flora. *Biological Journal of the Linnean Society*. **1,** 135–148 (1969).
18. Hedberg, O. Evolution of the Afroalpine Flora. *Biotropica*. **2,** 16–23 (1970).
19. Assefa, A., Ehrich, D., Taberlet, P., Nemomissa, S. & Brochmann, C. Pleistocene colonization of afro-alpine 'sky islands' by the arctic-alpine *Arabis alpina*. *Heredity*. **99,** 133–142 (2007).
20. Popp, M., Gizaw, A., Nemomissa, S., Suda, J. & Brochmann, C. Colonization and diversification in the African 'sky islands' by Eurasian *Lychnis* L. (Caryophyllaceae). *Journal of Biogeography*. **35,** 1016–1029 (2008).
21. Demos, T. C., Peterhans, J. C. K., Agwanda, B. & Hickerson, M. J. Uncovering cryptic diversity and refugial persistence among small mammal lineages across the Eastern Afromontane biodiversity hotspot. *Molecular Phylogenetics and Evolution*. **71,** 41–54 (2014).
22. Gottelli, D., Marino, J., Sillero-Zubiri, C. & Funk, S. M. The effect of the last glacial age on speciation and population genetic structure of the endangered Ethiopian wolf (*Canis simensis*). *Molecular Ecology*. **13,** 2275–2286 (2004).
23. Bonnefille, R., Roeland, J. & Guiot, J. Temperature and rainfall estimates for the past 40,000 years in equatorial Africa. *Nature*. **346,** 347–349 (1990).
24. Mairal, M., Sanmartín, I. & Pellissier, L. Lineage-specific climatic niche drives the tempo of vicariance in the Rand Flora. *Journal of Biogeography*, doi: 10.1111/jbi.12930 (2017).
25. Chala, D., Zimmermann N. E., Brochmann, C. & Vakkestuen, V. Migration corridors for alpine plants among the 'sky islands' of eastern Africa: do they, or did they exist? *Alpine Botany* (in press).
26. Wondimu, T. *et al.* Crossing barriers in an extremely fragmented system: two case studies in the afro-alpine sky island flora. *Plant Systematics and Evolution*. **300,** 415–430 (2014).
27. Masao, C. A. *et al.* Phylogeographic history and taxonomy of some afro-alpine grasses assessed based on AFLPs and morphometry: *Deschampsia cespitosa*, *D. angusta* and *Koeleria capensis*. *Alpine Botany*. **123,** 107–122 (2013).
28. Hedberg, O. Monograph of the genus Canarina L. (Campanulaceae). *Svensk Botanisk Tidskrift Utgifven af Svenska Botaniska Foreningen. Stockholm*. **55,** 17–62 (1961).
29. Olesen, J. M., Alarcon, M., Ehlers, B. K. & Aldasoro, J. J. & Roquet, C. Pollination, biogeography and phylogeny of oceanic island bellflowers (Campanulaceae). *Perspectives in Plant Ecology, Evolution and Systematics*. **14,** 169–182 (2012).
30. Yamagiwa, J. & Basabose, A. K. Diet and seasonal changes in sympatric gorillas and chimpanzees at Kahuzi–Biega Nation. *Primates*. **47,** 74–90. (2006).
31. Mairal, M., Pokorny, L., Aldasoro, J. J., Alarcon, M. & Sanmartin, I. Ancient vicariance and climate-driven extinction explain continental-wide disjunctions in Africa: the case of the Rand Flora genus *Canarina* (Campanulaceae). *Molecular Ecology*. **24,** 1335–1354 (2015).
32. Yalden, D. & Largen. The endemic mammals of Ethiopia. *Mammal Review*. **22,** 115–150 (1992).
33. Bobe, R. The evolution of arid ecosystems in eastern Africa. *Journal of Arid Environments*. **66,** 564–584 (2006).
34. Sepulchre, P. *et al.* Tectonic uplift and eastern Africa aridification. *Science*. **313,** 1419–1423 (2006).
35. Davidson, A. & Rex, D. Age of volcanism and rifting in southwestern Ethiopia. *Nature*. **283,** 657–658 (1980).
36. Engler, A. Plants of the northern temperate zone in their transition to the high mountains of Tropical Africa. *Annals of Botany*. **4,** 523–540 (1904).
37. Gizaw, A. *et al.* Colonization and diversification in the African 'sky islands': insights from fossil-calibrated molecular dating of *Lychnis* (Caryophyllaceae). *New Phytologist*. **35,** 1016–1029 (2016).
38. White, F. The Afromontane Region. Biogeography And Ecology Of Southern Africa (ed. Springer) 463–513 (Netherlands, 1978).
39. Galley, C., Bytebier, B., Bellstedt, D. & Linder H. P. The Cape element in the Afrotemperate flora: from Cape to Cairo? *Proceedings of the Royal Society B: Biological Sciences*. **274,** 535–543 (2007).
40. Gehrke, B. & Linder, H. P. The scramble for Africa: pan-temperate elements on the African high mountains. *Proceedings of the Royal Society of London B: Biological Sciences*. **276,** 2657–2665 (2009).
41. de Maio, N., Wu, C. H., O'Reilly, K. M. & Wilson, D. New routes to phylogeography: A Bayesian structured coalescent approximation. *PLoS Genetics*. **11,** e1005421 (2015).
42. Gizaw, A. *et al.* Phylogeography of the heathers *Erica arborea* and *E. trimera* in the afro-alpine 'sky islands' inferred from AFLPs and plastid DNA sequences. *Flora*. **208,** 453–463 (2013).
43. Knox, E. & Palmer, J. Chloroplast DNA evidence on the origin and radiation of the giant *Lobelias* in Eastern Africa. *Systematic Botany*. **23,** 109–149 (1998).
44. Bowie, R. C., Fjeldsa, J., Hackett, S. J. & Crowe, T. M. Molecular evolution in space and through time: mtDNA phylogeography of the Olive Sunbird (*Nectarinia olivacea/obscura*) throughout continental Africa. *Molecular Phylogenetics and Evolution*. **33,** 56–74 (2004).

45. Benson, C. W., Irwin, M. S. & White, C. M. N. The significance of valleys as avian zoogeographical barriers. *Annals of Cape Province Museum of Natural History*. **2,** 155–189 (1962).
46. Freitag, S. & Robinson, T. J. Phylogeographic Patterns in Mitochondrial DNA of the Ostrich (*Struthio camelus*). *The Auk*. **110,** 614–622 (1993).
47. DeMenocal, P. B. Plio-Pleistocene African climate. *Science*. **270,** 53–59 (1995).
48. Trauth, M. H., Maslin, M. A., Deino, A. & Strecker, M. R. Late Cenozoic moisture history of East Africa. *Science*. **309,** 2051–2053 (2005).
49. Ruiz Guajardo, J. C. *et al.* Landscape genetics of the key African acacia species *Senegalia mellifera* (Vahl)- the importance of the Kenyan Rift Valley. *Molecular Ecology*. **19,** 5126–5139 (2010).
50. Aga, E., Bekele, E. & Bryngelsson, T. Inter-simple sequence repeat (ISSR) variation in forest coffee trees (*Coffea arabica* L.) populations from Ethiopia. *Genetica*. **124,** 213–221 (2005).
51. Ayele, T. B., Gailing, O., Umer, M. & Finkeldey, R. Chloroplast DNA haplotype diversity and postglacial recolonization of *Hagenia abyssinica* (Bruce) J.F. Gmel. in Ethiopia. *Plant Systematics and Evolution*. **280,** 175–185 (2009).
52. Kadu, C. A. C. *et al.* Phylogeography of the Afromontane *Prunus africana* reveals a former migration corridor between East and West African highlands. *Molecular Ecology*. **20,** 165–178 (2011).
53. Ehrich, D. *et al.* Genetic consequences of Pleistocene range shifts: contrast between the Arctic, the Alps and the East African mountains. *Molecular Ecology*. **16,** 2542–2559 (2007).
54. Hewitt, G. Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society*. **58,** 247–276 (1996).
55. Kahindo C., Bowie R. C. & Bates J. M. The relevance of data on genetic diversity for the conservation of Afromontane regions. *Biological Conservation*. **134,** 262–270 (2007).
56. Hedberg, O. Afroalpine vegetation compared to paramo: convergent adaptations and divergent differentiation. *Páramo:* An Andean Ecosystem Under Human Influence (ed. Academic Press, London, 15–30, 1992).
57. Wilfert, L., Kaib, M., Durka, W. & Brandl, R. Differentiation between populations of a termite in eastern Africa: implications for biogeography. *Journal of Biogeography*. **33,** 1993–2000 (2006).
58. Templeton, A. R., Crandall, K. A. & Sing, C. F. A. Cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics*. **132,** 619–633 (1992).
59. Clement, M., Posada, D. C. & Crandall, K. A. TCS: a computer program to estimate gene genealogies. *Molecular Ecology*. **9,** 1657–1659 (2000).
60. Librado, P. & Rozas, J. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*. **25,** 1451–1452 (2009).
61. Drummond, A. J. & Rambaut, A. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*. **7,** 214 (2007).
62. Pokorny, L., Oliván, G. & Shaw, A. J. Phylogeographic patterns in two southern hemisphere species of *Calyptrochaeta* (Daltoniaceae, Bryophyta). *Systematic Botany*. **36,** 542–553 (2011).
63. Lemey, P., Rambaut, A., Drummond, A. J. & Suchard, M. A. Bayesian phylogeography finds its roots. *PLoS Computational Biology*. **5,** e1000520 (2009).
64. Vos, P. *et al.* AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research*. **23,** 4407–4414 (1995).
65. Meudt, H. M. & Clarke, A. C. Almost forgotten or latest practice? AFLP applications, analyses and advances. *Trends in plant science*. **12,** 106–117 (2007).
66. Whitlock, R., Hipperson, H., Mannarelli, M., Butlin, R. K. & Burke, T. An objective, rapid and reproducible method for scoring AFLP peak-height data that minimizes genotyping error. *Molecular Ecology Resources*. **8,** 725–735 (2008).
67. Bonin, A. *et al.* How to track and assess genotyping errors in population genetics studies. *Molecular Ecology*. **13,** 3261–3273 (2004).
68. Ehrich, D. Aflpdat: a collection of R functions for convenient handling of AFLP data. *Molecular Ecology Notes*. **6,** 603–604 (2006).
69. Vekemans, X., Beauwens, T., Lemaire, M. & Roldan-Ruiz, I. Data from amplified fragment length polymorphism (AFLP) markers show indication of size homoplasy and of a relationship between degree of homoplasy and fragment size. *Molecular Ecology*. **11,** 139–151 (2002).
70. Nei, M. & Li, W. H. Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences*. **76,** 5269–5273 (1979).
71. Lynch, M. & Milligan, B. G. Analysis of population genetic structure with RAPD markers. *Molecular Ecology*. **3,** 91–99 (1994).
72. Hardy, O. J. Estimation of pairwise relatedness between individuals and characterization of isolation-by-distance processes using dominant genetic markers. *Molecular Ecology*. **12,** 1577–1588 (2003).
73. Zhivotovsky, L. A. Estimating population structure in diploids with multilocus dominant DNA markers. *Molecular Ecology*. **8,** 907–913.
74. Ehrich, D. AFLPdat: a collection of R functions for convenient handling of AFLP data. *Molecular Ecology Notes*. **6,** 603–604 (2006).
75. Crisp, M. D., Laffan, S., Linder, H. P. & Monro, A. Endemism in the Australian flora. *Journal of Biogeography*. **28,** 183–198 (2001).
76. Schönswetter, P. & Tribsch, A. Vicariance and dispersal in the alpine perennial *Bupleurum stellatum* L.(Apiaceae). *Taxon*. **54,** 725–732 (2005).
77. Rohlf, F. J. NTSYS-pc version 2.0. *Numerical taxonomy and multivariate analysis system*. Exeter software, Setauket, New York (1998).
78. Huson, D. H. & Bryant, D. Application of phylogenetic networks in evolutionary studies. *Molecular biology and evolution*. **23,** 254–267 (2006).
79. Excoffier, L., Laval, G. & Schneider, S. Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evolutionary bioinformatics online*. **1,** 47 (2005).
80. Pritchard, J. K., Stephens, M. & Donnelly, P. Inference of population structure using multilocus genotype data. *Genetics* **155,** 945–959 (2000).
81. Falush, D., Stephens, M. & Pritchard, J. K. Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Molecular Ecology Notes*. **7,** 574–578 (2007).
82. Evanno, G., Regnaut, S. & Goudet, J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*. **14,** 2611–2620 (2005).
83. Hardy, O. J. & Vekemans, X. Spagedi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology Notes*. **2,** 618–620 (2002).
84. Manni, F., Guerard, E. & Heyer, E. Geographic patterns of (genetic, morphologic, linguistic) variation: how barriers can be detected by using Monmonier's algorithm. *Human biology*. **76,** 173–190 (2004).
85. Ryan, W. B. *et al.* Global multi-resolution topography synthesis. *Geochemistry, Geophysics, Geosystems*. **10,** Q03014, doi: 10.1029/2008GC002332 (2009).
86. Silvestrini, M. *et al.* Genetic diversity and structure of Ethiopian, Yemen and Brazilian Coffea arabica L. accessions using microsatellites markers. *Genetic Resources and Crop Evolution*. **54,** 1367–1379 (2007).
87. Derero, A., Gailing, O. & Finkeldey. R. Maintenance of genetic diversity in Cordia africana Lam., a declining forest tree species in Ethiopia. *Tree Genetics & Genomes*. **7,** 1–9 (2010).

88. Sertse, D., Gailing, O., Eliades, N. G. & Finkeldey, R. Anthropogenic and natural causes influencing population genetic structure of Juniperus procera Hochst. ex Endl. in the Ethiopian highlands. *Genetic Resources and Crop Evolution*. **58,** 849–859 (2011).
89. Kadu, C. A. C. *et al.* Divergent pattern of nuclear genetic diversity across the range of the Afromontane Prunus africana mirrors variable climate of African highlands. *Annals of Botany*. **111,** 47–60 (2013).
90. Muchugi, A. *et al.* Genetic structuring of important medicinal species of genus Warburgia as revealed by AFLP analysis. *Tree Genetics & Genomes*. **4,** 787–795 (2008).
91. Omondi, S. F. *et al.* Genetic Diversity and Population Structure of Acacia senegal (L) Willd. in Kenya. *Tropical Plant Biology.* **3,** 59–70 (2010).
92. Matthee, C. A. & Robinson, T. J. Molecular phylogeny of the springhare, Pedetes capensis, based on mitochondrial DNA sequences. *Molecular Biology and Evolution* **14,** 20–29 (1997).
93. Arctander, P., Johansen, C. & Coutellec-Vreto, M. A. Phylogeography of three closely related African bovids (tribe Alcelaphini). *Molecular Biology and Evolution*. **16,** 1724–1739 (1999).
94. Girman, D. J. *et al.* Patterns of population subdivision, gene flow and genetic variability in the African wild dog (Lycaon pictus). *Molecular Ecology.* **10,** 1703–1723 (2001).
95. Pitra, C., Hansen, A. J., Lieckfeldt, D. & Arctander, P. An exceptional case of historical outbreeding in African sable antelope populations. *Molecular Ecology.* **11,** 1197–1208 (2002).
96. Dubach, J. *et al.* Molecular genetic variation across the southern and eastern geographic ranges of the African lion, Panthera leo. *Conservation Genetics*. **6,** 15–24 (2005).
97. Belay, G. & Mori, A. Intraspecific phylogeographic mitochondrial DNA (D-loop) variation of Gelada baboon, Theropithecus gelada, in Ethiopia. *Biochemical Systematics and Ecology*. **34,** 554–561 (2006).
98. Nicolas, V. *et al.* Phylogeographic structure and regional history of Lemniscomys striatus (Rodentia: Muridae) in tropical Africa. *Journal of Biogeography*. **35,** 2074–2089 (2008).
99. Colangelo, P. *et al.* A mitochondrial phylogeographic scenario for the most widespread African rodent, Mastomys natalensis. *Biological Journal of the Linnean Society*. **108,** 901–916 (2013).
100. Evans, B. J., Bliss, S. M., Mendel, S. A. & Tinsley, R. C. The Rift Valley is a major barrier to dispersal of African clawed frogs (Xenopus) in Ethiopia. *Molecular Ecology.* **20,** 4216–4230 (2011).
101. Freilich, X., Anadón, J. D., Bukala, J., Calderon, O., Chakraborty, R. & Boissinot, S. Comparative Phylogeography of Ethiopian anurans: impact of the Great Rift Valley and Pleistocene climate change. *BMC evolutionary biology* **16,** 206 (2016).
102. Field, L. M. *et al.* Analysis of genetic variability in Anopheles arabiensis and Anopheles gambiae using microsatellite loci. *Insect Molecular Biology*. **8,** 287–297 (1999).
103. Lehmann, T. *et al.* The Rift Valley complex as a barrier to gene flow for Anopheles gambiae in Kenya. *Journal of Heredity*. **90,** 613–621 (1999).
104. Lehmann, T. *et al.* Brief communication. The Rift Valley complex as a barrier to gene flow for Anopheles gambiae in Kenya: the mtDNA perspective. *Journal of Heredity*. **91,** 165–168 (2000).
105. Lehmann, T. *et al.* Population Structure of Anopheles gambiae in Africa. *Journal of Heredity*. **94,** 133–147 (2003).
106. Krafsur, E. S. Population structure of the tsetse fly Glossina pallidipes estimated by allozyme, microsatellite and mitochondrial gene diversities. *Insect Molecular Biology*. **11,** 37–45 (2002).
107. Braginets, O. P., Minakawa, N., Mbogo, C. M. & Ya, G. Population genetic structure of the African malaria mosquito Anopheles funestus in Kenya. *The American journal of tropical medicine and hygiene*. **69,** 303–308 (2003).

Acknowledgements

We thank E. Cano and M. I. Garcia (*Parque Científico de Madrid*) for technical assistance; J. M. Olesen, M.-X. Ren, E. Habiba and A. Sánchez-Meseguer for help in field collection and herbarium searches, and V. Culshaw for her useful comments on the language. We thank the authorities of the National Parks of the Bale Mountains and Harenna Forest (Ethiopia), the Aberdare Mountains Park (Kenya) and Rwenzori Mountains Park (Uganda) for help with fieldwork. Special thanks are given to P. Kamau from the National Museums of Kenya, Botany Department, for her help in herbarium and field searches. This study was supported by the Spanish government (grants CGL2006-09696 and CGL2010-18631/BOS to J.J.A., and CGL2012-40129-C02-01 and CGL2015-67849-P to I.S.), and from the *Consejería de Educación de la Comunidad de Madrid*, and EU FP7 Synthesys Programme award to L.P. (GB-TAF-5153). M.A. was funded by the *Junta para la Ampliación de Estudios Doctorales* programme (CSIC/FSE). M.M. was supported by a MINECO PhD fellowship (BES-2010-037261), under project CGL2009-13322-C03-01 to I.S.

Author Contributions

M.M., M.A., I.S. and J.J.A. designed the study. J.J.A., M.A. and L.P. contributed samples. M.M., A.H., J.J.A., and M.A. performed the research. M.M., M.A., J.J.A. and I.S. wrote the manuscript. All authors commented and approved the final version.

Additional Information

Competing Interests: The authors declare no competing financial interests.

How to cite this article: Mairal, M. *et al.* Geographic barriers and Pleistocene climate change shaped patterns of genetic variation in the Eastern Afromontane biodiversity hotspot. *Sci. Rep.* **7**, 45749; doi: 10.1038/srep45749 (2017).

CHAPTER 3

Paleo-islands as refugia and sources of genetic diversity within volcanic archipelagos: The case of the widespread endemic *Canarina canariensis* (Campanulaceae)

Mario Mairal, Isabel Sanmartín, Juan José Aldasoro, Victoria Culshaw, Ioanna Manolopoulou, Marisa Alarcón

This chapter has been published in the peer-review journal *Molecular Ecology*, 24: 3944–3963. Doi: 10.1111/mec.13282

MOLECULAR ECOLOGY

Molecular Ecology (2015) **24**, 3944–3963 doi: 10.1111/mec.13282

Palaeo-islands as refugia and sources of genetic diversity within volcanic archipelagos: the case of the widespread endemic *Canarina canariensis* (Campanulaceae)

M. MAIRAL,* I. SANMARTÍN,*[1] J. J. ALDASORO,† V. CULSHAW,* I. MANOLOPOULOU‡ and M. ALARCÓN†[1]

*Real Jardín Botánico (RJB-CSIC), 28014 Madrid, Spain, †Institut Botànic de Barcelona (IBB-CSIC-ICUB), 08038 Barcelona, Spain, ‡University College London, WC1E 6BT London, UK

Abstract

Geographical isolation by oceanic barriers and climatic stability has been postulated as some of the main factors driving diversification within volcanic archipelagos. However, few studies have focused on the effect that catastrophic volcanic events have had on patterns of within-island differentiation in geological time. This study employed data from the chloroplast (cpDNA haplotypes) and the nuclear (AFLPs) genomes to examine the patterns of genetic variation in *Canarina canariensis*, an iconic plant species associated with the endemic laurel forest of the Canary Islands. We found a strong geographical population structure, with a first divergence around 0.8 Ma that has Tenerife as its central axis and divides Canarian populations into eastern and western clades. Genetic diversity was greatest in the geologically stable 'palaeo-islands' of Anaga, Teno and Roque del Conde; these areas were also inferred as the ancestral location of migrant alleles towards other disturbed areas within Tenerife or the nearby islands using a Bayesian approach to phylogeographical clustering. Oceanic barriers, in contrast, appear to have played a lesser role in structuring genetic variation, with intra-island levels of genetic diversity larger than those between-islands. We argue that volcanic eruptions and landslides after the merging of the palaeo-islands 3.5 Ma played key roles in generating genetic boundaries within Tenerife, with the palaeo-islands acting as refugia against extinction, and as cradles and sources of genetic diversity to other areas within the archipelago.

Keywords: ancestral areas, extinction, oceanic islands, palaeo-islands, volcanic refugia

Received 5 March 2015; revision received 16 June 2015; accepted 17 June 2015

Introduction

Due to their small size, discrete geographical boundaries, substantial environmental heterogeneity and buffered climates, oceanic islands represent excellent natural laboratories to examine the role of ecological adaptations vs. geographical isolation as drivers of diversification (Carson & Templeton 1984; Gillespie 2004). Patterns of genetic variation are expected to be hierarchical in islands, with between-island genetic differentiation stronger than within-islands (Bottin *et al.* 2005; García-Verdugo *et al.* 2010), because oceanic barriers are generally more effective than topographic barriers at promoting isolation in insular systems (Gillespie & Clague 2009). However, the birth and development of volcanic islands is usually followed by a large number of destructive events in the form of secondary eruptions, landslides, merging of palaeo-islands, etc. (Carracedo 1994; Fernández-Palacios *et al.* 2011). These events promote habitat fragmentation and the subsequent genetic isolation of populations, and they ultimately drive differentiation and speciation within-islands (Carson *et al.* 1990; Gillespie & Roderick 2002, 2014; Macías-Hernández *et al.* 2013). Therefore, volcanic archipelagos represent an ideal framework for studying patterns of diversification at dif-

---

Correspondence: Mario Mairal, Fax: +34914200157; E-mail: mariomairal@gmail.com

[1] Equal contributions.

ferent spatial scales (between- and within-islands) and also over different temporal levels (allopatric speciation, extirpations and recolonizations).

The Canary Islands are a volcanic archipelago formed by a chain of seven islands, located 110 km from the north-western coast of Africa. Their geological history and highly endemic biota have interested scientists since the early 19th century (von Humboldt 1814; Lyell 1855). The islands were formed in the last 21 Ma with an east to west pattern of chronological emergence due to a mantle plume (Carracedo *et al.* 1998; Zaczek *et al.* 2015) (Fig. 1a). They are separated by deep oceanic trenches and have never been connected to the mainland. The Canary Islands have long been considered refugia for continental lineages that have survived the climatic changes of the Late Cenozoic, but also as cradles of biodiversity where multiple in situ diversification events have taken place (Francisco-Ortega *et al.* 2000; Juan *et al.* 2000). In addition, these islands are regarded as a hotspot for plant diversity (Médail & Quézel 1997): approximately 40% of Canarian vascular plants are endemics (Santos-Guerra 2001) and the flora is generally characterized by high levels of interpopulation differentiation in comparison with other archipelagos (Francisco-Ortega *et al.* 2000; de Paz & Caujapé-Castells 2013). Recently, they have been proposed as reservoirs and sources of genetic diversity (García-Verdugo *et al.* 2015; Patiño *et al.* 2015).

Most studies on the Canarian flora have focused on the pattern of interisland diversification, particularly at the species level (Francisco-Ortega *et al.* 2002; Kim *et al.* 2008; Vitales *et al.* 2014a,b). Nevertheless, the complex topographies and long histories of avalanches and secondary eruptions of many of these islands are likely to have favoured within-island diversification (Juan *et al.* 2000; Brown *et al.* 2006). Tenerife has the most complex history of all the islands. It existed at first as three separate islands, dating back to the Late Miocene: Roque del Conde (11.9–8.9 Ma) in the southwest, Teno (6.2–5.6 Ma) in the northwest, and Anaga (4.9–3.9 Ma) in the northeast. Eruptive central volcanic episodes fused these islands 3.5 Ma and gave rise to the present island of Tenerife (Ancochea *et al.* 1990; Fig. 1b). The palaeo-islands remained thereafter relatively stable, whereas the central part of Tenerife continued to be active until 0.13 Ma (Ancochea *et al.* 1990, 1999; Cantagrel *et al.* 1999; Guillou *et al.* 2004; Carracedo 2014; see Fig. 1b). Interestingly, the three palaeo-islands of Tenerife, together with La Gomera—which has also remained geologically stable since the Pliocene—are presently home to the highest phylogenetic diversity and endemic richness of the Canarian Archipelago (Reyes-Betancort *et al.* 2008). These areas also harbour the best-preserved laurel forests, considered as an ancient, unique flora restricted to Macaronesia, and they share several endemic and restricted species (see Table S1, Supporting information). Besides their geological

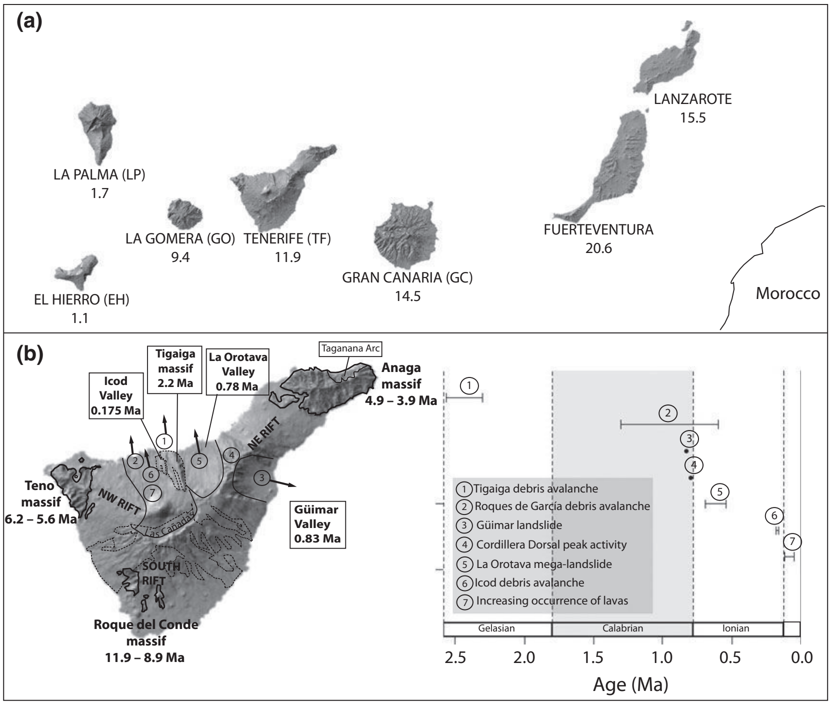


Fig. 1 (a) Geographical map of the Canary Island archipelago, with the code for each island used in the text and their age of emergence (million years, Ma). (b) Map of Tenerife showing main Pleistocene landslides and other geological events cited in the text. The old basaltic series corresponding to the palaeo-islands are shown in the three contoured massifs at the edges of the island. Lines show valleys created after landslides. Dashed lines indicate the post-erosional Las Cañadas volcano. Arrows indicate the directions of the landslides. The graph on the right shows the main Tenerife catastrophic events that are commented in the text; the numbers in the legend refer to those in the map. Ages and maps adapted from: Carracedo *et al.* 1998, 2007 ; Ancochea *et al.* (1999); Cantagrel *et al.* (1999); Guillou *et al.* (2004); Boulesteix *et al.* (2012); Zaczek *et al.* (2015).

(volcanic) stability, the three palaeo-islands of Tenerife also exhibit a topographic complexity and variety of micro-climates that might have favoured their role as micro-refugia against climate- or human-induced extinction (Harter *et al.* 2015).

In reviewing the role of Tenerife palaeo-islands as refugia across several plant lineages, Trusty *et al.* (2005) found that species endemic to the palaeo-islands often occupied a derived position in the lineage's phylogeny. This position argued against the idea of these massifs as ancient refugia. However, other studies, especially in animals, have reported ages for divergence events between taxa endemic to these palaeo-islands, either at the species or at the intraspecies level (Juan *et al.* 1996, 2000; Dimitrov *et al.* 2008; Macías-Hernández *et al.* 2013; Puppo *et al.* 2014) that are contemporaneous or predate the age of merging of the precursor palaeo-islands 3.5 Ma (Ancochea *et al.* 1990). Few plant studies (Gómez *et al.* 2003; García-Verdugo *et al.* 2010) have focused on patterns of within-island genetic variation for widespread Canarian endemics, and none of them have provided estimates of lineage divergence times, which is necessary to relate within-island patterns to the island geological history. Species that are present in multiple islands ('multiple island endemics' or MIEs) are especially relevant to understand the role of palaeo-islands as undisturbed areas that have acted as reservoirs and sources of genetic diversity not only within- but between-islands.

Here, we study patterns of genetic diversity and the demographic and spatial history of a multiple island endemic, *Canarina canariensis* (L.) Vatke. This 'flagship' species of the Canary Islands, elected as its 'national flower' (Kunkel 1991), is a diploid (2n = 34) herbaceous plant that grows mostly in cleared areas surrounding the endemic laurisilva forest. It presently occurs in the central and western Canary Islands: Gran Canaria, Tenerife, La Gomera, La Palma and El Hierro. *Canarina canariensis* is a herbaceous plant that occasionally climbs on nearby plants; it is pollinated by generalist birds (Rodríguez-Rodríguez & Valido 2011) and its fleshy fruits are dispersed by vertebrates (Valido *et al.* 2003; Rodríguez *et al.* 2008). Genus *Canarina* belongs to tribe Platycodoneae, a basal group within family Campanulaceae (Mansion *et al.* 2012; Olesen *et al.* 2012; Wang *et al.* 2013; Mairal *et al.* 2015). In addition to *C. canariensis*, the genus comprises two other species inhabiting the Afromontane forests of Eastern Africa, *C. eminii* and *C. abyssinica*, being this an example of a wide, continental-scale disjunction of 7000 km spanning across the Sahara. Mairal *et al.* (2015) recently reconstructed the phylogeny and spatiotemporal evolution of *Canarina*. They inferred that *C. canariensis* diverged from its African relatives at the end of the Miocene (*c.* 7 Ma). This extraordinary temporal and spatial disjunction was explained as the result of vicariance and climate-driven extinction resulting in the fragmentation of an ancient widespread distribution. The colonization of the Canary Islands by the ancestors of *C. canariensis* apparently occurred much later, in the Pleistocene, probably from a now extinct and geographically closer North African population: the earliest event of population divergence is dated at only around 1 Ma (Late Pleistocene) and involved an east–west vicariance within Tenerife (Mairal *et al.* 2015).

This age and the presence of *C. canariensis* in several islands makes it an ideal candidate to evaluate patterns of within-island diversification in relation to the recent geological history of the archipelago. Our main aims were to: (i) determine the geographical distribution of genetic variation within *C. canariensis*, (ii) find evidence of extinction and diversification processes that may be related to geological events, (iii) find ancestral areas and reconstruct interisland migration events and (iv) examine the putative role of the palaeo-islands of Tenerife as refugia of genetic diversity, both relictual and recent.

Haplotype networks are commonly used in population-level studies because they provide a clearer picture of the reticulate relationships between genetic pools than a branching tree, especially when gene flow is present. These networks are often inferred using Statistical Parsimony (SP, Templeton *et al.* 1992) implemented in the software TCS (Clement *et al.* 2000), which allows estimation of the haplotype network while minimizing the number of mutation events differing among haplotypes. However, this method fails to incorporate the uncertainty associated with the network inference and therefore does not allow for statistical evaluation of alternative phylogeographical scenarios (Bloomquist *et al.* 2010). Moreover, unobserved events such as local population growth or past extinction of haplotypes may mislead inference in parsimony-based methods. Here, we compare results from TCS with those obtained from a model-based, Bayesian statistical method, Bayesian phylogeographic and ecological clustering (BPEC, Manolopoulou *et al.* 2011), which allows estimating the posterior probabilities for haplotype tree networks under a coalescent-based migration–mutation model (Manolopoulou & Emerson 2012). To our knowledge, this is the first study to use this method for island phylogeography.

Materials and methods

Population sampling and DNA extraction

Canarina canariensis has a significantly greater presence on Gran Canaria (GC) and Tenerife (TF), in comparison with La Gomera (LG), La Palma (LP) and El Hierro (EH). Seventeen populations of *C. canariensis* were sampled in several fieldtrips between 2009 and 2012: four in GC, eight in TF, two in LG, two in LP and one in EH. Where possible, we collected a minimum of 10 samples per population. To reduce inflation in gene descriptors due to biased sampling (Caujapé-Castells 2010), samples were collected from individuals scattered across the whole occupancy area of each population. DNA from 160 individuals and preserved in silica gel was extracted using the DNeasy Plant Mini Kit (QIAGEN Inc., Valencia, CA, USA), from 20 to 25 mg of silica-gel-dried leaves obtained from the fresh plant tissue collected from the field expeditions.

Chloroplast DNA sequencing

We selected three cpDNA intergenic spacers regions for sequencing; these markers have proven to be useful for intraspecific analyses of population structure (Mairal *et al.* 2015). We generated 432 new sequences: *rpl*32-*trn*L^{UAG} (144 sequences), *trn*S^{GCU}–*trn*G^{UCC} (144 sequences) and *pet*B^{1365}–*pet*D^{738} (144 sequences). PCR and sequencing protocols followed those of Mairal *et al.* (2015). The sources of the material examined, the GenBank Accession nos and full references are detailed in Table S2 (Supporting information).

Haplotype analyses

Sequences for each region were aligned using MAFFT 6.814b (Katoh *et al.* 2002), implemented in the software GENEIOUS PRO 5.4.4. (Biomatters Ltd., Auckland, New Zealand). Sequences were checked and manually adjusted where necessary by following alignment rules described in Kelchner (2000). We analysed the three sequenced regions as three data partitions to perform phylogenetic analyses. MRMODELTEST v.2.2 software (Nylander 2004) was used to determine the best fitting model of sequence evolution of each data partition.

Summary statistics for within-population genetic diversity were calculated as follows: the number of haplotypes H(n), haplotype diversity (Hd), nucleotide diversity π, nucleotide heterozygosity θ, and G_{ST} and the number of migrants per generation (Nm) were estimated for each population using DnaSP (version 5.10; Librado & Rozas 2009).

The relationships among lineages were investigated through haplotype network analysis, using 6–12 individuals from different populations and examining the three sequenced regions. Genealogical relationships among haplotypes were inferred via the statistical parsimony algorithm (Templeton *et al.* 1992) implemented in TCS 1.21 (Clement *et al.* 2000). The number of mutational steps resulting from single substitutions among haplotypes was calculated with 95% confidence limits, and gaps were represented as missing data.

The BPEC method (Manolopoulou *et al.* 2011; Manolopoulou & Emerson 2012) was implemented in the R package to identify genetically distinct geographical population clusters and ancestral locations. Like TCS, BPEC relies on parsimony in order to reduce the number of candidate trees to a manageable set. The BPEC method, unlike Standard Parsimony, fits a prior over all possible trees in order to identify trees with high posterior probability in a fully model-based framework, thus accommodating for uncertainty in haplotype relationships, which is one of the main criticisms of TCS (Knowles 2008). Each possible tree defines a set of possible migration events that may have led to the observed population substructure. Different scenarios of trees and migration events are explored through Markov chain Monte Carlo (MCMC), similar to the method proposed by Sanmartín *et al.* (2008) for estimating rates of interisland dispersal. Migration events were assumed to occur when a haplotype (with or without a mutation from its parent haplotype) migrates to a new geographical cluster. MCMC simultaneously estimates high probability trees, number of migration events and corresponding clusters. The method assumes that the migration rate and the population growth are constant. BPEC requires two main user-defined inputs: the maximum number of migrations (denoted as 'MaxMig' in the software) and the parsimony relaxation parameter used to reconstruct the set of possible trees (denoted as 'ds' in the software). MAXMIG allows the user to set the upper bound for the number of migration events and hence the maximum number of clusters (MaxMig + 1). Larger values include more models but require much greater computation time. As the ds value is increased, the parsimony assumption is relaxed: if two observed sets of sequences have an unobserved intermediate missing sequence (an unobserved mutation), then any pair of sequences of distance $\{1, \ldots, ds\}$ nucleotides will be considered as the 'missing path'. Two MCMC chains were run for 3 million iterations. The results were stable, with ds > 3 having no effect on the inferred tree, and similarly, any number of migrations above four converging to a 5–cluster model. The phylogeographical clustering obtained was superimposed upon a haplotype tree and used to estimate ancestral locations for migration events. As a further exploration, we divided the data set into two groups (eastern and western clades—for details, see *coalescent dating* section in results). Haplotypes sampled from Roque del Conde were quite divergent; thus, separate analyses were run in which we included and excluded this location from the two groups. We ran BPEC analysis on each of these four data sets (eastern and western groups with and without Roque del Conde).

Haplotype divergence times were estimated in BEAST v.1.7 (Drummond & Rambaut 2007). We carried out a first analysis under a strict clock model and a coalescent constant population tree prior, using a secondary age estimate (Mairal *et al.* 2015; normal prior: mean = 0.76 Ma, standard deviation (SD) = 0.327 Ma) to calibrate the root node of the '*C. canariensis* data set'; this included all haplotypes detected in our sampling ($N = 10$). This analysis gave us very large 95%HPD (high posterior density) credibility intervals and poor ESS for posterior age estimates, probably due to the low information content at the population level (see Results) and the presence of a single calibration point. We carried out a second analysis applying the 'nested dating approach' described in Mairal *et al.* (2015), in which a higher-level data set including representatives of all three species of *Canarina* and nine outgroup taxa was used to inform the clock rate of a linked population-level data set (*C. canariensis*) under a *mixed Yule-coalescent model* (Ho *et al.* 2005; Pokorny *et al.* 2011). The higher-level data set was calibrated with fossil-derived secondary age estimates (see Mairal *et al.* 2015), while the tree prior was unlinked to apply a coalescent constant size model to the population-level data set and a stochastic birth–death (Yule) prior to the species-level one (Mairal *et al.* 2015). The clock model was set to an uncorrelated log-normal prior to accommodate the change in mutation rate from species to populations, with a uniform distribution for the *ucld.mean* (10^{-4}–10^{-1}) and a default exponential distribution for the *ucld.stdev*; the substitution model was set to GTR+G; choice of priors was based on Bayes Factor comparisons using the path sampling method in BEAST (Baele *et al.* 2012); see Table S3 (Supporting information) for results from exploratory analyses to assess the reliability of our date estimates with reference to these settings. Two MCMC chains were run for 50 million generations, sampling parameters every 1000 generation. We used Tracer v1.6 (Rambaut *et al.* 2007) to verify the following: whether a stationary distribution was attained, whether there was convergence among chains and whether effective sample sizes (ESS values) were >200 for all parameters. A 10% burn-in of the sampled populations was discarded (5 million). Post-burn-in trees were summarized into a maximum clade credibility tree using TREEANNOTATOR v.1.6.1, with mean values and 95% credible intervals for nodal ages, and were visualized in FIGTREE 1.3.1 (Rambaut & Drummond 2009). The resulting age estimates from this second analysis exhibited considerably larger ESS values and narrower 95%HPD intervals than in the first and are the ones reported here.

Demographic history

Statistics used to describe demographic patterns may be biased by a strong genetic structure or lack of panmixia among populations. As we detected strong genetic structure in our data set (see results), we performed demographic analyses in subsetted data sets that were less genetically structured (including approximate panmictic populations). Three groups previously recognized by the haplotype network and BEAST analyses were used: ET-GC (including populations in east Tenerife and Gran Canaria), TENO-GO (including populations in Teno and La Gomera) and LP-EH (including populations in La Palma and El Hierro). For the same reasons as described above, the Roque del Conde population was included and discluded from the ET-GC and TENO-GO groups. Overall, we ran the analyses on five groups.

We used three different approaches to infer the demographic processes shaping the genetic structure of *C. canariensis*. First of all, to test for evidence of population expansion, we carried out a neutrality test—Fu and Li's tests (Fu & Li 1993; Fu 1996) and Tajima's *D* test (Tajima 1989)—for each population group. We used the DNAsp program, version 5.0 (Librado & Rozas 2009), and assessed the significance of each test with 10 000 coalescent simulations. Second, we plotted the mismatch distribution for each group using the observed number of differences between all pairs of sequences with the ARLEQUIN v.3.0 software (Excoffier *et al.* 2005). The goodness of fit of the observed mismatch distribution to the theoretical distribution under a constant population size model was tested with the raggedness index (*HRag*) (Harpending 1994). Third, (i) we created the extended Bayesian skyline plot (EBSP), implemented in BEAST, for each population group, and (ii) we performed EBSP analyses for each of the 16 population groups. For each group from (i) and (ii): two independent chains were run simultaneously for 150 million generations, sampling every 1000 generations; a strict clock model was used, whereas all other parameters were set identical to those described above for the nested dating analysis; the root node was calibrated using a normal prior with a mean age estimate and 95% high posterior density (HPD) credible intervals obtained from this analysis.

AFLP fingerprinting

For the AFLP analysis, we used a total of 97 individuals from 10 populations, which covered all of the islands: one population from Gran Canaria, five from Tenerife, one from La Gomera, two from La Palma and one from El Hierro. Laboratory molecular protocols for the AFLP analysis (Vos *et al.* 1995) were implemented using the AFLP plant mapping kit (Applied Biosystems®, Foster City, CA, USA). To select the appropriate primers, we first carried out a pilot study combining fluorophores and restriction enzymes for five geographically distant individuals (one per island), using in total 32 primer combinations. One sample from each individual was duplicated as a blind sample to test for reproducibility and contamination. Reproducibility and the number of alleles per sample were calculated by choosing three combinations of primers: 1-*Eco*RI$_{6\text{-FAM}}$-ACT/*Mse*I-CAA, 2-*Eco*RI$_{\text{VIC}}$-AGG/*Mse*I-CTA and 3-*Eco*RI$_{\text{VIC}}$-AGG/*Mse*I-CTT and using the GENEMAPPER v3.7 software (Applied Biosystems). These three primer pairs showed high reproducibility and homogeneously scattered bands and produced polymorphic AFLP profiles and clear fragments. In the digestion phase, samples of DNA were digested with the restriction enzymes EcoRI and MseI and linked to the primers EcoRI 5′-CTCGTAGAC TGCGTACC-3′/5′AATTGGTACGCAGTCTAC-3′) and MseI (5′- GACGATGAGTCCTGAC-3′/5′-ATCTCAGGA CTCAT-3′). The three different AFLP reactions were as follows: (i) restriction and ligation in a single reaction; (ii) and (iii) consecutive PCR amplifications (preselective and selective). PCR products were checked on 1% agarose gels.

AFLP data analyses

The resulting AFLP fragments were analysed using the GENEMAPPER 3.7 software. Peaks were recorded in 100–500 base pairs ranges. Shorter fragments were discarded because the majority of this size class have a high chance of being nonhomologous fragments (Vekemans *et al.* 2002). For each primer combination, an automated size detection and peak binning was employed followed by manual editing of bins to exclude shoulder peaks and unreliable loci (variation between replicates). Peak height data were then exported and loaded into the R package AFLPSCORE version 1.4a (Whitlock *et al.* 2008), and the AFLP profiles were scored and the error rates were estimated. These rates were below the critical bound of 5% indicated in previous reports (Bonin *et al.* 2004) for each primer combination. Data reliability was assessed through comparison of duplicates, from one or two individuals per population. Data were converted into binary presence⁄absence scores for each locus. The resulting AFLP presence/absence matrix was analysed using a selection of different analyses. The AFLPSURV v.1.0 software (Vekemans *et al.* 2002) was used to estimate demographic statistics such as Nei's gene diversity (Hj), pairwise differentiation among subpopulations (F_{ST}) and the percentage of polymorphic fragments per population (P) (Nei & Li 1979; Lynch & Milligan 1994). This was done under the assumptions of either the Hardy–Weinberg equilibrium or partial self-fertilization, based on a previous study on the reproductive biology of *C. canariensis* (Rodríguez-Rodríguez & Valido 2011). A Bayesian method in AFLPSURV was also used to estimate allelic frequencies through employment of a nonuniform prior distribution (Zhivotovsky 1999). Ten thousand permutations were run to calculate the F_{ST} parameter from which genetic distances between individuals, populations and geographical groups were calculated. To locate genetic clustering of individuals within the AFLP data set, a pairwise similarity matrix for all individuals using the Dice's coefficient as similarity distance was constructed, and the resulting matrix subjected to a principal coordinates analysis (PCA) implemented in NTSYS v.2.1 (Rohlf 1998). Next, genetic relationships among samples were visualized in SPLITSTREE v.4.10 (Huson & Bryant 2006) using neighbour-net analysis through the use of the split decomposition method. Finally, to quantify the amount of genetic differentiation attributable to geographical and population subdivision, a hierarchical analysis of molecular variance was performed using ARLEQUIN v.3.0 software (Excoffier *et al.* 2005). Exploratory analyses were performed considering, alternatively, islands and palaeo-islands as geographical units in order to investigate the distribution of genetic variance attributable to oceanic barriers.

Bayesian clustering methods implemented in STRUCTURE v.2.3 (Pritchard *et al.* 2000; Falush *et al.* 2007) were used to assess the genetic structure of populations. This model-based approach assumes that loci are in Hardy–Weinberg equilibrium and linkage equilibrium within populations. Analyses were performed under admixture conditions and correlated allele frequencies between groups. 500 000 MCMC generations (plus a burn-in of 100 000) were run for K values of 1–10, with 10 repetitions for each. The most likely K value was determined by the method from Evanno *et al.* (2005), which is implemented in STRUCTURE HARVESTER (Earl 2012). We explored other values of K to detect further genetic substructure of populations, especially in Tenerife. To test the effect of the spatial distance on the genetic structure of the populations, correlations between genetic (measured as F_{ST}) and spatial distances between pairs of populations were determined using the Mantel permutation procedure implemented in NTSYS v. 2.1. The genetic distance matrix used was based on the presence/absence matrix; the geographical distance matrix was based on the absolute distances between the geographical coordinates for each collected population. In addition, to identify possible geographical locations acting as major genetic barriers among *C. canariensis* populations, we computed barriers on a Delaunay triangulation using Monmonier's algorithm in BARRIER v.2.2 (Manni *et al.* 2004). The significance was examined by the mean of 1000 bootstrapped distance matrices obtained using AFLPSURV. Only barriers with support >96% were considered as significant.

Results

Haplotype network analysis and coalescent dating

The *pet* B^{1365}–*pet* D^{738} region consisted of 937 sites, *rpl*32–*trn*L^{UAG} of 654 sites and *trn*S^{GCU} –*trn*G^{UCC} of 658 sites. The final concatenated data matrix consisted of 144 sequences and 2249 nucleotide sites, of which nine were polymorphic (Table S4, Supporting information). Among the 144 individuals sampled from the 16 populations, we observed 10 different plastid DNA haplotypes (H1 to H10, Fig. 2a) with haplotype diversity of 0.6632. There were three haplotypes that dominated and were clearly geographically delimited: H1, H9 and H3. The most frequently observed haplotype was H1 (52.08%), which was dominant in GC and east Tenerife (ET). Haplotype H9 was restricted to West Tenerife (WT) and LG. Haplotype H3 was present in the western islands of LP and EH. Populations with the commonest and rarest haplotypes (frequencies $\leq 2\%$) were in the most rugged areas of Tenerife: the palaeo-islands of Anaga (An-W, An-C, An-E—exclusive haplotypes H2 and H5), Teno (Ten-A, Ten-Ep—exclusive allele H8 and allele H10 shared with LG) and Roque del Conde (Conde—exclusive H7). The small western islands presented one unique haplotype, each: LP (LP-Lt—H4) and EH (H6).

Summary statistics for within-population genetic diversity are shown in Table 1. Populations situated in the three palaeo-islands [Anaga (An-W), Teno (Ten-A) and Roque del Conde (Conde)] showed the highest values for the number of haplotypes H(*n*), haplotype diversity H(d), nucleotide diversity π and nucleotide heterozygosity θ. The least diverse population areas were found in Gran Canaria and outside the palaeo-islands regions of Tenerife. G_{ST} and N_m values indicated high genetic cohesion between some islands: east Tenerife populations were highly connected with Gran Canaria populations; west Tenerife with La Gomera; and La Palma with El Hierro (Table S5, Supporting information). Interestingly, the highest G_{ST} values and lowest Nm were found between populations situated east and west of Tenerife, with the exception of Anaga An-W. This is the only population in east Tenerife that showed some genetic cohesion with populations in west Tenerife, in particular with the Teno and Adeje massifs.

The haplotype network constructed with TCS (Fig. 2b) shows a loop involving three dominant haplotypes, with haplotype H1 as the ancestral root haplotype according to coalescent criteria on haplotype frequency (Templeton *et al.* 1992). Dominant haplotypes at each side of Tenerife (East H1 and West H9) were separated by five mutational steps, with haplotype H10 as intermediate. Haplotype H3, present in the western islands, was separated by four mutational steps from H1 and three from H9.

BPEC results are summarized in Table 2. With mutational step limit equal to 10 and the maximum number of migrations equal to 8, we allowed for high gene flow, genetic divergence and numerous hidden mutations, which separate the haplotype clusters. The clustering with high posterior probabilities (pp = 1) showed the existence of five clusters (contour regions in Fig. 3a). The total number of haplotypes was 10 plus one missing haplotype (Fig. 3b). Location of the ancestral populations for migration events and inference of the root node were consistent across exploratory analyses of the

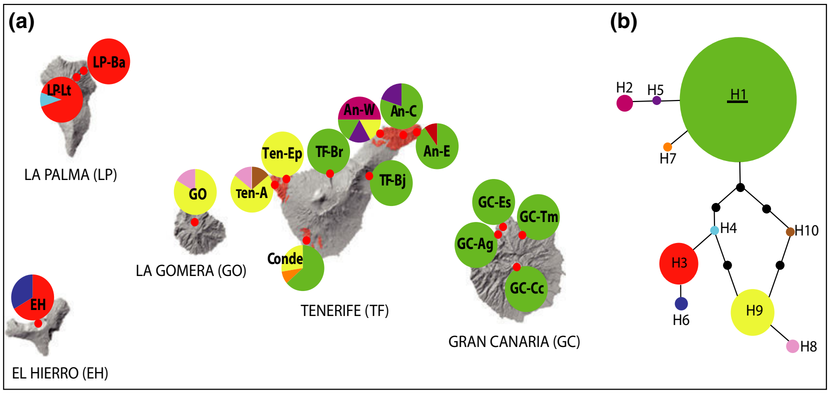


Fig. 2 Haplotype distribution and network inferred for the chloroplast markers by TCS. (a) Pie charts show the geographical location of populations and the frequency of occurrence of each haplotype, and circle size is proportional to population size. Population codes are given in Table S2 (Supporting information). (b) Statistical Parsimony network inferred by TCS. Each haplotype is shown in a different colour; circle size is proportional to its frequency among populations.

Table 1 Descriptors of within-population genetic diversity in the cpDNA haplotypes and AFLP markers for each population studied of *C. canariensis*

	Haplotypes							AFLPs			
Island (Palaeo-island)	Population	No samples	Haplotype	H (*n*)	H (d)	π	θ	No samples	No of polymorphic fragments (% in brackets)	Hj (SE)	Number of private fragments
Gran Canaria	GC-Cc	8	H1	1	0	0	0	8	409 (68.7)	0.2344 (0.0085)	0
Gran Canaria	GC-Tm	11	H1	1	0	0	0	NA	NA	NA	NA
Gran Canaria	GC-Es	10	H1	1	0	0	0	NA	NA	NA	NA
Gran Canaria	GC-Ag	11	H1	1	0	0	0	NA	NA	NA	NA
Tenerife (Anaga)	An-W	6	H1, H2, H5, H9	4	0.8	0.00112	0.00136	7	401 (67.4)	0.2677 (0.0088)	4
Tenerife (Anaga)	An-E	10	H1, H2	2	0.2	0.00018	0.00031	10	345 (58.0)	0.2303 (0.0080)	0
Tenerife (Anaga)	An-C	11	H1, H5	2	0.327	0.00015	0.00015	11	375 (63.0)	0.2453 (0.0078)	1
Tenerife	TF-Bj	4	H1	1	0	0	0	NA	NA	NA	NA
Tenerife	TF-Br	6	H1	1	0	0	0	NA	NA	NA	NA
Tenerife (Roque del Conde)	Conde	11	H1, H9, H7	3	0.618	0.00121	0.00091	NA	NA	NA	NA
Tenerife (Teno)	Ten-Ep	8	H9	1	0	0	0	9	351 (59.0)	0.2449 (0.0082)	0
Tenerife (Teno)	Ten-A	7	H8, H9, H10	3	0.524	0.00038	0.00054	7	415 (69.7)	0.2997 (0.0088)	2
Gomera	GO	12	H8, H9	2	0.303	0.00013	0.00015	12	473 (79.5)	0.3012 (0.0077)	1
La Palma	LP-Ba	10	H3	1	0	0	0	9	285 (47.9)	0.1923 (0.0080)	0
La Palma	LP-Lt	10	H3, H4	2	0.2	0.00009	0.00016	10	348 (58.5)	0.2309 (0.0079)	0
El Hierro	EH	9	H3, H6	2	0.5	0.00022	0.00016	9	303 (50.9)	0.2115 (0.0082)	1

H(*n*), number of haplotypes; H(d), haplotype diversity; π, nucleotide diversity; θ, nucleotide heterozygosity; Hj (se), Nei's gene diversity (standard error). Geographical locations for population codes are shown in Fig. 2. The most diverse population in each palaeo-island is shown in bold.

complete data set. The root node was inferred as missing (H11—Fig. 3b) and carried a high uncertainty. Haplotypes H1, H2 and H8 also carried significant posterior mass probabilities (Table 2). The most likely source for ancestral migration events within- and between-islands were populations located on the following: Conde (pp = 0.13), An-W (pp = 0.10) and LP-Lt (pp = 0.086). The subsetting analysis, including Conde, always recovered this population as the ancestral area with the highest probability. Without Conde, other palaeo-island populations were recovered as ancestral areas: Ten-A for the Western clade; and the three populations of Anaga for the Eastern clade, with An-W the most probable (Table 2; Fig. S1, Supporting information).

The Bayesian chronogram of haplotypes showed a geographical pattern of divergence congruent with the groups detected by TCS and BPEC (Fig. 4). The crown age or first divergence event among *Canarina* haplotypes was estimated to be 0.878 Ma (95% HPD = 0.452–1.365), corresponding to the Mid-Pleistocene period; it divided haplotypes into an eastern and western clade. The first divergence event in the eastern clade was estimated at 0.255 Ma (95% HPD = 0.041–0.633), while that in the western clade was dated older: 0.522 Ma (95% HPD = 0.161–0.967), separating La Palma and El Hierro from Teno and La Gomera. Mean age values and 95% HPD intervals for the BEAST analysis are shown in Table S6 (Supporting information).

Demographic analyses

Negative values of Fu's and Tajima's indicated a recent population expansion, although these values were only significant for the Fu's test of the East group (Table S7, Supporting information). The frequencies of pairwise

Table 2 Summary of results from the Bayesian phylogeographic and ecological clustering (BPEC) analyses. Selected values for parameters MaxMig (maximum number of migrations) and ds (relaxed parsimony assumption) are shown. Ancestral locations (Anc. Loc.) for migration events are shown with their associated posterior probabilities

	Maxmig	ds	Anc. Loc. 1	Anc. Loc. 2	Anc. Loc. 3	Root node	Root node probability	No of clusters
Total	8	10	Conde (0.13)	An-W (0.10)	LP-Lt (0.086)	H11 (missing)	0.17	5
Western Clade (WT + LP + GO + EH)	4	3	Ten-A (0.40)	Lp-Lt (0.31)	EH (0.18)	H11 (missing)	0.18	3
Eastern Clade (ET + GC)	1	3	An-W (0.212)	An-C (0.122)	An-E (0.116)	H11 (missing)	0.16	2
Western Clade + Conde	4	3	Conde (0.416)	Ten-A (0.197)	GO (0.190)	H11 (missing)	0.18	3
Eastern Clade + Conde	1	3	Conde (0.23)	An-W (0.17)	An-C (0.086)	H11 (missing)	0.16	3

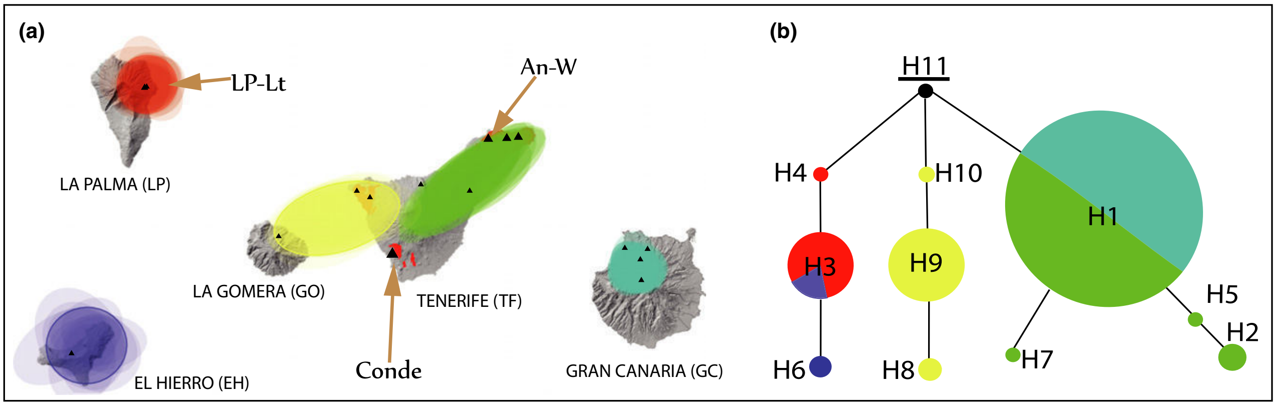


Fig. 3 Results from the Bayesian phylogeographic and ecological clustering (BPEC) analyses for the cpDNA markers. (a) Phylogeographical clusters (coloured blobs) and ancestral location for migration events (denoted by arrows). The contour regions are centred at the 'centre' of each population cluster, and the shaded areas show the radius of 50% concentration contours around it. Locations situated beyond the clusters could also belong to these clusters, but with low probability; in the case of Conde, it suggests a mixed composition, with this population as source of migrant alleles to the east and western clusters. (b) Haplotype network receiving the highest posterior probability. The small black circle H11 indicates an unobserved (missing) ancestral haplotype.

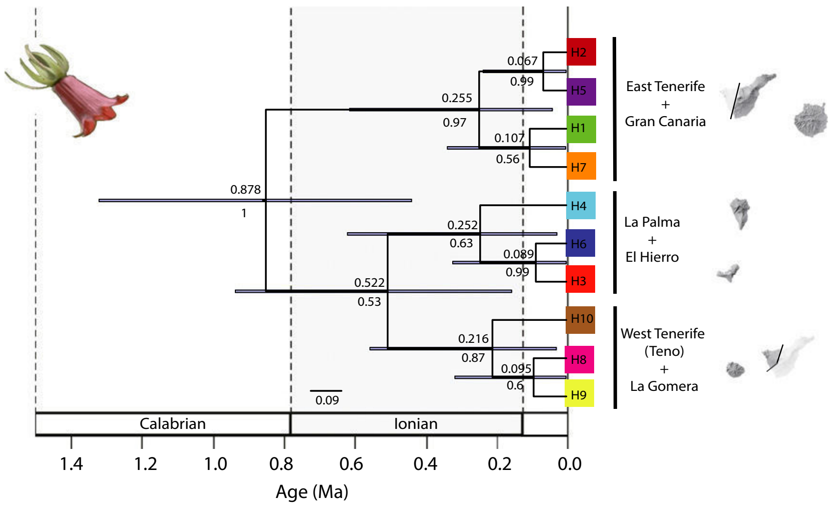


Fig. 4 Maximum clade credibility (MCC) tree obtained from the BEAST analysis of cpDNA haplotypes, showing mean ages (above branches) and 95% HPD credible intervals. Numbers below branches indicate Bayesian posterior clade support values. Codes for tips (H1 to H10) correspond to the haplotypes shown in Fig. 2.

differences in the mismatch distribution analysis resulted in unimodal distributions that were consistent with an expansion model. When we include the population of Conde in the East group, we obtained a second small peak. The raggedness statistics derived from the mismatch distribution were not significant and thus failed to reject the null hypothesis of recent population expansion. EBSP indicated a constant population size (Fig. S2, Supporting information). EBSP for each of the individual populations also obtained a constant population size (results not shown).

AFLP polymorphism, genetic diversity and structure

These results are summarized in Table 1. The final data set, after scoring, comprised 572 loci from 97 individuals and 10 populations. For AFLP analyses, we removed populations from Gran Canaria and Roque del Conde, which had been extracted with a different (less rigorous) method (CTAB) to standardize the quality of extracted DNA. Hj and P were higher in the La Gomera (GO) population, followed by the Teno (Ten-A) population in west Tenerife. The largest number of private fragments was detected in Tenerife (seven fragments). In La Gomera and El Hierro islands, a single fragment was detected. In Tenerife, five private fragments were detected in the Anaga massif (specifically in the An-W population—four fragments) and two in the Teno massif (Table 1). F_{ST} values with the nuclear data (Table S8, Supporting information) were consistent with the results obtained from the chloroplast markers (see above G_{ST} and Nm) and showed the same genetic cohesions between-islands and palaeo-islands.

A PCA differentiated three groups (Fig. S4, Supporting information): (i) east Tenerife populations; (ii) west Tenerife populations; and (iii) the An-W population, which demonstrate an intermediate position between (i) and (ii). The split network analysis outlined the same divergences between the two well separated groups, including the West group An-W population (Fig. 5b).

Hierarchical AMOVA analyses showed the largest proportion of genetic variation to be found among groups 4 and 5 (Table 3; Table S9, Supporting information). By analysing each island separately, we observed that differentiation within-islands (24.27%) is greater than between-islands (3.62%). A further analysis considering islands and palaeo-islands as separate units showed that differentiation within-islands and within-palaeo-islands is lower (11.30%) than between-islands and between-palaeo-islands (16.63%). According to the method of Evanno *et al.* (2005), STRUCTURE indicated that the most likely number of genetic clusters $K = 2$ ($\Delta K = 460$) represented the optimal number of Bayesian groups within *C. canariensis*, separating the east populations from the west, but also detected the presence of admixture in the intermediate populations. K levels $K = 4$ and $K = 5$ (the latter identified as $K = 6$, with 5 defined clusters plus one 'ghost' cluster with no individuals assigned so it was ignored, see Guillot *et al.* 2005) revealed a more complex genetic structure in these intermediate populations, resulting in additional clusters (Fig. 5a; Fig. S5, Supporting information). The STRUCTURE analysis using only Tenerife populations also delimited two groups (inset $K = 2$ in Fig. 5a; Fig. S5b, Supporting information). Overall, the results revealed a strong interisland genetic structure (Gran Canaria – Anaga massif and La Palma—El Hierro), admixture phenomena (among An-W, Teno populations and La Gomera) and an even more complex substructure within the island of Tenerife. The STRUCTURE sublevels ($K = 4$ and $K = 5$) were consistent with the AMOVA analyses. Only La Gomera differed between analyses; this is explained due to its high genetic variability and mixed composition. These groups are also congruent

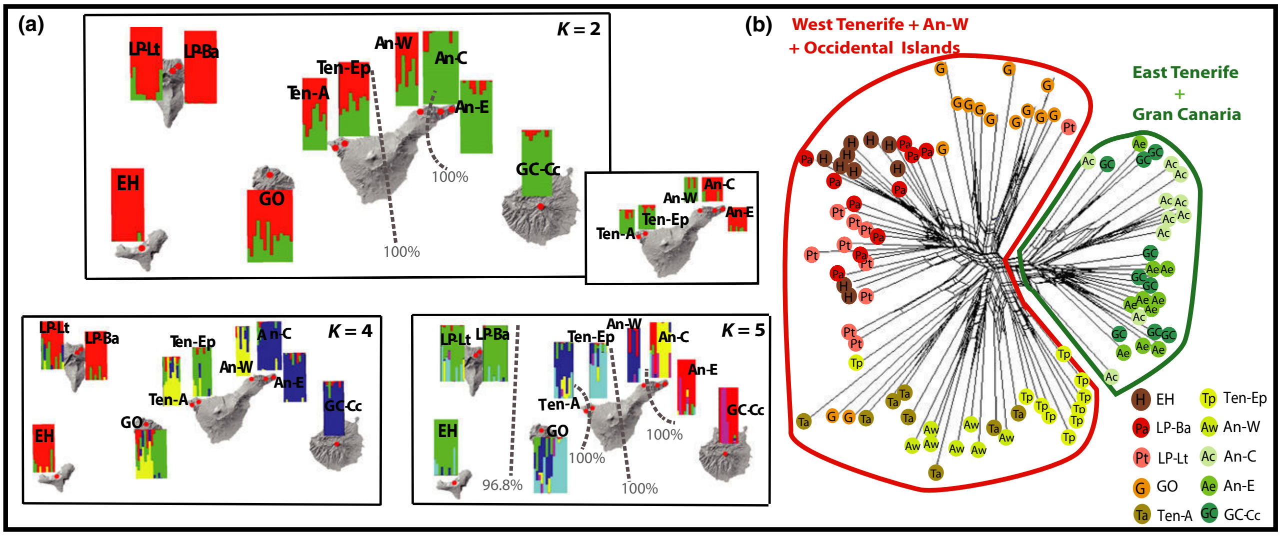


Fig. 5 Results from the analysis of AFLP markers. (a) Histograms showing the Bayesian clustering of individuals within populations (STRUCTURE); colours represent the proportion of individual membership to each inferred Bayesian group. Dotted lines indicate barriers to gene flow and their percentage, as inferred by BARRIER. (b) Split network with points coloured according to location, as indicated in the legend. Codes for populations found in Table S2 (Supporting information).

with the BARRIER results, which revealed two or four major boundaries (Fig. 5a for $K = 2$ and $K = 5$). All barriers showed values of 100% except LP-EH (96.8%). No linear relationship was found between pairwise F_{ST} and geographical distance with the Mantel analysis ($r = 0.42$, $P = 1$).

Discussion

Palaeo-islands as refugia against volcanic (catastrophic) extinctions

The strong genetic structure often detected in island endemics has been postulated to respond to several interconnected physical and biotic factors (Stuessy *et al.* 2014); foremost among them are geographical isolation and extinction-recolonization processes (Carson *et al.* 1990; Gübitz *et al.* 2005; Macías-Hernández *et al.* 2013). The strongest genetic structure and levels of genetic variation in *C. canariensis* were detected between two of the ancient palaeo-islands that currently form Tenerife: Teno in the west and Anaga in the east. Populations in other islands were genetically associated with these two lineages: La Palma, La Gomera and El Hierro to Teno; and Gran Canaria to Anaga. The fact that these patterns are shared between chloroplast (cpDNA, Fig. 2) and nuclear (AFLPs, Fig. 5) markers supports a long history of isolation among populations (Zink & Barrowclough 2008). Furthermore, the hierarchical AMOVA analysis based on the AFLP data showed that genetic variation was notably higher among populations within Tenerife than those located in different islands (Table 3). This pattern has been found in a few other Canarian endemics (*Olea europaea guanchica* and *Pinus canariensis*), but

Table 3 Hierarchical analysis of molecular variance (AMOVA) for *C. canariensis* based on allelic variation at different levels: (A) among groups, (B) among populations within groups and (C) within populations

		Levels			*F*-statistics		
AMOVA groups	No of groups (*K*)	A	B	C	Fsc	Fst	Fct
[GC + Anaga] [An-W + Ten-A] [Ten-Ep] [GO + LP + EH]	4	17.44	13.06	69.5	0.15820	0.305	0.17439
[GC+ An-E] [An-C] [Ten-Ep] [An-W + Ten-A] [GO + LP + EH]	5	18.92	11.21	69.88	0.13819	0.30123	0.18918
[GC] [TF] [GO] [LP] [EH]	5	3.62	24.27	72.10	0.25186	0.27897	0.03624
[GC] [An-E + An- C] [An-W] [Ten-Ep] [Ten-A] [GO] [LP] [EH]	8	16.63	11.30	72.07	0.13558	0.27934	0.16631

Geographical locations of populations are shown in Fig. 2.

usually not as marked as here (<10%: Gómez *et al.* 2003; García-Verdugo *et al.* 2010; see review in García-Verdugo *et al.* 2014). Interestingly, considering the palaeo-islands as separate areas reversed the pattern (Table 3). Together with our BARRIER analyses (Fig. 5a in $K = 5$), these results support the hypothesis that the geographical distribution of genetic diversity in *C. canariensis* is structured around the palaeo-islands and that these ancient massifs could be considered as separate units in phylogeographical analyses, a sort of 'islands within-islands'. It indicates that geological barriers within-islands—for example, those associated to catastrophic volcanic events—have probably been more important in structuring patterns of genetic differentiation within species than the oceanic barriers separating the islands (García-Verdugo *et al.* 2014).

Unfortunately, we could not obtain AFLP data for the population in Roque del Conde. Given the mixed composition of the chloroplast compartment in this population (Fig. 2), it is possible that inclusion of this population in our analysis would have led to higher levels of genetic admixture and lower K values in the STRUCTURE analyses—though given the marked east/west split among the remaining populations (Fig. 5), this is unlikely. On the other hand, the two most frequent haplotypes, H1 and H9, in the Conde population are also the most frequent or dominant within the eastern and western clades, respectively, while the divergence of these two haplotypes in the BEAST tree (Fig. 4) can be traced back to the basalmost split in *C. canariensis* (0.8 Ma). This, together with the presence of a unique haplotype (H7) and the fact that the Conde population is identified as the ancestral source of westward and eastward migration events in the BPEC analyses, suggests that the admixture detected in the chloroplast compartment for Roque del Conde is of ancient rather than recent origin. It is thus possible that including this population in the AFLP analysis would have increased rather than decreased levels of genetic diversity within Tenerife for *C. canariensis*.

Which might have caused this level of within-island genetic structure? Many Canarian plant and animal taxa include sister lineages endemic to the Tenerife palaeo-islands (Juan *et al.* 2000; Báez *et al.* 2001), especially among laurel forest species from Teno and Anaga (Table 4). While in some cases, the divergence between these taxa is found at the species level (e.g. *Trechus*, Contreras-Díaz *et al.* 2007; *Pericallis*, Jones *et al.* 2014) and/or predates the merging of the palaeo-islands (*Micromeria*, Puppo *et al.* 2014; *Pholcus*, Dimitrov *et al.* 2008); in others, it is observed within species (e.g. *Tarentola delalandii*, Gübitz *et al.* 2000) and/or postdates the merging of the palaeo-islands (*Eutrichopus*, Moya *et al.* 2004; *Calathus abaxoides*, Emerson *et al.* 1999). The fact that the pattern of divergence is not contemporary across taxa suggests that the role of Tenerife palaeo-islands in structuring genetic variation has been long lasting. In *C. canariensis*, the basal divergence between the basal eastern and western lineages was dated as 0.878 Ma (Fig. 4) substantially, postdating the age of the merging of the palaeo-islands (*c.* 3.5 Ma, Ancochea *et al.* 1990). This 'young' east–west divergence might be explained by historical dispersal events between forest patches followed by in situ diversification. Alternatively, it could be explained by allopatric speciation (vicariance) driven by catastrophic volcanic events within a previously more widespread distribution with low or no reconnection. This might explain the 'young' (<3.5 Ma) east–west Tenerife divergence found in *Canarina*, and seen in other taxa (Table 4). The reciprocal monophyly, similar haplotype diversity levels and deep temporal divergence found between the eastern and western clusters of *C. canariensis* in Tenerife (Fig. 4) favour the vicariant, allopatric scenario. In the last 1 million years, several major landslides and volcanic events have affected the north of Tenerife, extending from the summit to the coast (Boulesteix *et al.* 2012, 2013; Carracedo 2014), for example the Güimar and La Orotava mega-landslides (Fig. 1b). These events could have fragmented the ancient laurel forest corridor that extended across the northern flank of the island (Moya *et al.* 2004), interrupting gene flow within species associated with this laurel forest (e.g. *C. canariensis*) and promoting differentiation among populations. Several studies have supported debris avalanches as important factors driving diversification within terrestrial Canarian organisms (Juan *et al.* 2000; Brown *et al.* 2006; Macías-Hernández *et al.* 2013); others have reported a temporal divergence within northern Tenerife (Thorpe *et al.* 1996; Moya *et al.* 2004) that is similar to the one found in *C. canariensis* and corresponds well with the time of the La Orotava mega-landslide (Ancochea *et al.* 1990; Boulesteix *et al.* 2013).

Whereas central Tenerife was almost completely covered by catastrophic events until as recently as 0.13 Ma (Ancochea *et al.* 1999), the three palaeo-islands of Teno, Anaga and Roque del Conde remained geologically stable since the mid-Pliocene (see Fig. 1). This suggests that these areas could have acted as refugia, allowing the survival of populations that otherwise disappeared in other parts of the island that were affected by volcanic activity. Support for this suggestion comes from the population genetic analysis. According to the central–marginal hypothesis (Eckert *et al.* 2008), spatial structure and genetic diversity should be higher in areas that have acted as refugia for the preservation of genotypes that went extinct in other areas and generally for the long-term persistence of populations (Hewitt 2000;

Table 4 Phylogeographical breaks with divergence times reported in the literature between the palaeo-islands of Tenerife

Organism	Phylogeographical disjunction and groups	Divergence (Ma)	Markers	References
Species complex in darking beetle *Pimelia*	East Tenerife/West Tenerife	5.5 Ma	Mitochondrial (COI) and nuclear (ITS-1)	Juan *et al.* (1996)
Lizard *Gallotia galloti*	Western/North-eastern lineages	0.7 Ma	Mitochondrial cytochrome b	Thorpe *et al.* (1996)
Beetle *Calathus abaxoides*	Teno/Anaga	350 000 years	Two mitochondrial (COI and COII)	Emerson *et al.* (1999)
Skink *Chalcides viridanus*	Teno/Anaga	1.1 Ma	Two mitochondrial (12S and 16S)	Brown *et al.* (2000)
Gecko *Tarentola delalandii*	Clade 1. Teno + Roque del Conde/Anaga. Clade 2. Teno/Roque del Conde	Clade 1. 9–10 Ma Clade 2. 7.6 Ma	Mitochondrial cytochrome b	Gübitz *et al.* (2000)
Mite *Steganacarus carlosi*	Clade 1. Teno/Anaga. Clade 2. Roque del Conde/Anaga	Clade 1. 3.2 Ma Clade 2. 25–3.6 Ma	Mitochondrial cytochrome oxidase 1	Salomone *et al.* (2002)
Beetles *Eutrichopus*	Teno (*E. gonzalezi*)/ Anaga (*E. canariensis*)	0.7 Ma	Mitochondrial (COII marker)	Moya *et al.* (2004)
Species complex in beetle *Tarphius*	Teno/Anaga	1.2 Ma (1–1.4)	Two mitochondrial (COI and COII)	Emerson & Oromí (2005)
Beetle *Trechus*	Clade 1. Anaga + Teno (*T. antonii*, *T. tenoensis*, *T. felix*)/Anaga (*T. fortunatus*) Clade 2. *T. flavocintus*; Teno/East Tenerife	Clade 1. 1.73 Ma (HPD: 1.48–2.01) Clade 2. approx 0.75 Ma	Part of mitochondrial genes cytochrome oxidase I and II (Cox1 and Cox2),and nuclear (ITS 2)	Contreras-Díaz *et al.* (2007)
Spider *Pholcus*	Anaga (*P. malpaisensis*, *P. knoeseli*)/Teno (*P. intricatus*, *P. mascaensis*/*P. tenerifensis* and *P. roquensis*)	3.93 Ma (HPD: 2.2–5.88)	Four mitochondrial (CO1, 16S, NADH and tRNAleu). Morphological data	Dimitrov *et al.* (2008)
Grasshopper *Arminda brunneri*	Anaga + Güimar/Teno + Roque del Conde	1–0.17 Ma	Two mitochondrial (12s rRNA, ND5) and two nuclear gene fragments (28s rRNA, ITS2)	Hochkirch & Goerzig (2009)
Spider *Dysdera verneaui*	Teno/Anaga	3.94 Ma (HPD: 5.1–2.7)	Mitochondrial (*cox*1)	Macías-Hernández *et al.* (2013)
Plant *Pericallis*	Teno (*P. echinata*)/ Anaga (*P. tussilaginis*)	2.87 Ma (HPD: 1.55–4.76)	Nuclear ITS	Jones *et al.* (2014)
Plant *Micromeria*	Teno (*M. densiflora*)/Anaga (*M. teneriffae*, *M. glomerata* and *M. rivas-martinezii*) + Central group	5.2 Ma	8 nuclear loci	Puppo *et al.* (2014)

Tzedakis *et al.* 2013; Feliner 2014; Gavin *et al.* 2014). The higher number of ancestral and endemic alleles, private fragments and larger heterozygosity levels exhibited by the populations of the palaeo-islands of Tenerife (Table 1) are congruent with the idea that these massifs acted as reservoirs of ancient genetic diversity and as refugia against volcanically induced extinction. Interestingly, La Gomera, an island that has been geologically quiescent since the Pliocene (Carracedo & Day 2002), shows the highest Hj and percentage of polymorphic nuclear DNA fragments (Table 1), suggesting that this island might have acted in a similar way to the palaeo-islands of Tenerife.

Further support for the extinction hypothesis comes from the BPEC analysis. Theoretical predictions of coalescent theory states that high-frequency haplotypes have been present for a long time, and more recent ones are rare and derived from the commonest haplotypes (Posada & Crandall 2001). Additionally, a root haplotype is expected to have a higher number of haplotype connections in the network, rather than being close to the tips. However, past extinction of haplotypes can obscure the inference, with younger haplotypes becoming the most prevalent, so accurately identifying the root haplotype is a challenging task. The fact that BPEC does not provide a single estimate of the haplotype network like Statistical Parsimony, but a finite (probability) distribution of haplotype trees—as well as the existence of the underlying migration model—allows this method to incorporate the uncertainty in the haplotype rooting. BPEC estimates the missing, extinct haplotype H11 as the most probable root of the haplotype network (Fig. 3b), although haplotypes H1, H2 and H8 are also associated with high posterior probabilities.

Have palaeo-islands acted as sources of genetic diversity within and towards other islands?

The theory of Pleistocene climate refugia (Hewitt 2000) states that historically environmentally stable areas can act as sources of genetic diversity exporting migrant alleles to other, disturbed regions (Gavin *et al.* 2014). Palaeo-islands could have played the same role in volcanic archipelagos, although in this case catastrophic geological events rather than climatic changes might be responsible for the observed patterns. BPEC provides support to this hypothesis, identifying the populations in the palaeo-islands of Tenerife as the source areas of ancestral migration events to other adjacent islands, such as from Teno to La Gomera or from Anaga to Gran Canaria (Fig. 3a, b; Table 2). A third dispersal event from west Tenerife/La Gomera to La Palma is supported by the BEAST tree (Fig. 4), and a fourth dispersal event from La Palma to El Hierro is inferred by the BPEC analysis (Fig. 3). Similar patterns with the central Canaries as centres of dispersal events within the archipelago have been described in other animal and plant studies (Francisco-Ortega *et al.* 2002; Gómez *et al.* 2003; Sanmartín *et al.* 2008; Mairal *et al.* 2015; Puppo *et al.* 2015). Moreover, a spatio-temporal pattern of colonization comparable to *Canarina*, showing Tenerife as the centre of dispersal events to adjacent islands in the last 1 Ma, can be found in the Canarian lineage of *Cistus* [0.33 (0.88–0.07) Ma; Guzmán & Vargas 2010], *Cheirolophus* (1 Ma; Vitales et al. 2014a), *Cistus monspeliensis* (0.93–0.20 Ma; Fernández-Mazuecos & Vargas 2011) and *Gallotia galloti* (0.8–0.9 Ma; Cox *et al.* 2010). This might be a consequence of the central geographical position of Tenerife within the archipelago—acting as a crossroad for dispersal events—but also of the concentration of plant genetic diversity in the palaeo-islands.

The case of Gran Canaria is especially interesting. It shares the same haplotype with Tenerife (H1, Figs 2 and 3). Low haplotypic diversity (Hn and Hd in Table 1) could be explained by a recent colonization after a catastrophic event. The island was subject to intense volcanic activity during the Holocene (24 eruption events; Rodríguez-Gonzalez *et al.* 2009), so extinction might explain its present low genetic diversity. Another possibility is related to the topography of Gran Canaria, where a network of ravines (locally known as 'barrancos') connecting at their summits could have facilitated gene flow among populations. Dispersal of *C. canariensis* seeds by *Gallotia* lizards (Rodríguez *et al.* 2008), probably using forest gaps and edges of roads as dispersal corridors (Delgado *et al.* 2007), might have helped to connect populations in the highly altered laurel forest of Gran Canaria.

In addition to dispersal events between-islands, the palaeo-islands of Tenerife might have acted as sources of genetic diversity within Tenerife, exporting migrant alleles to other geologically unstable, disturbed areas. Our demographic analyses indicated a recent population expansion in two populations of east Tenerife close to Anaga (TF-Br and TF-Bj). Although this result should be taken with caution (the EBSP analysis supported a constant effective population size; Fig. S2, Supporting information), these two populations exhibited also a single cpDNA haplotype, which agrees with the idea of a recent colonization. The areas where these populations are located (La Orotava Valley and Güimar Valley, respectively) have been subjected to catastrophic volcanic events (Fig. 1b). Some authors (Thorpe *et al.* 1996; Gübitz *et al.* 2000) have proposed the existence of a corridor of suitable habitat along the northern coastal fringe of Tenerife to explain migrations of the reptiles *G. galloti* and *T. delalandii* from Anaga to the west. These dispersal events could also explain the patterns found in our BPEC analyses, which suggest Anaga populations as sources of migration events to other populations in eastern Tenerife and Gran Canaria (Fig. 3; Table 2).

Ongoing genetic connectivity between the populations in the palaeo-islands of Tenerife and those from other areas is supported by the nuclear genome, which shows genetic admixture between west-Anaga, Teno and La Gomera (Fig. 5a). Admixture could be explained by the carrying of pollen by nectar-feeding birds between forest patches (e.g. ringing techniques have confirmed migration of the main pollinator (*Phylloscopus canariensis*) between Teno and Anaga; A. González, personal communication). The fact that this connectivity is to some extent lost in the cpDNA might be explained by the cpDNA not being transported via pollen or, alternatively, by the small size and haploid nature of the chloroplast genome, which imply shorter coalescent times and less time to fix novel mutations for chloroplast markers (Avise 2000; Jakob & Blattner

2006). On the other hand, the widespread distribution of some cpDNA haplotypes across the archipelago (H1, H3 and H9, Fig. 2a) supports some gene flow among populations driven by seed dispersal. In the Canary Islands, birds have been cited as important vectors for the dispersal of fleshy fruits (Arevalo *et al.* 2007; Padilla *et al.* 2012), and the latter has been associated with frequent gene flow preventing speciation within widespread island endemics (García-Verdugo *et al.* 2014), such as in *Canarina*. This fits with what is known on the reproductive biology of *C. canariensis*, which is pollinated by generalist birds (Rodríguez-Rodríguez & Valido 2011), while its fleshy fruits are dispersed by vertebrates such as *Gallotia* lizards (Valido *et al.* 2003; Rodríguez *et al.* 2008). An additional factor to explain frequent dispersal between-islands are the eustatic sea level shifts during the Pleistocene, which might have decreased geographical distance between-islands (Rijsdijk *et al.* 2014).

Finally, in addition to exporting migrant alleles to other islands and disturbed areas, the palaeo-islands of Tenerife might have acted as cradles or sources or new genetic diversity. The higher haplotype and nucleotide diversity and higher Nei's gene diversity Hj exhibited by populations in these areas (Table 1) agree with their role as ancient refugia but also as sources of novel genetic diversity. Maximum topographic complexity is one of the main factors explaining species richness and high speciation rates within-islands (Whittaker *et al.* 2007). The rugged nature of the Tenerife palaeo-islands has likely promoted genetic differentiation within these massifs. For example, divergences found between populations in east and west forest ranges within Anaga have been explained by the existence of deep ravines and the volcanic arc of Taganana (Fig. 1b), acting as geographical barrier to gene flow (Macías-Hernández *et al.* 2013). These divergences were also detected in *C. canariensis*.

Conclusions

Traditionally, the distribution of genetic diversity within archipelagos is assumed to be structured around oceanic barriers, with between-island divergences expected to be larger than within-island differentiation. Here, we showed that within-island genetic patterns might be as strong as or stronger than those observed between-islands when they are associated with historical volcanic events. In *C. canariensis*, geographical patterns of genetic variation are structured around the palaeo-islands of Tenerife, with a minor secondary effect due to oceanic barriers. Carine & Schaefer (2010) argued that although relatively short oceanic distances separate the Canary Islands, they might be responsible for the high diversity levels found in the archipelago, acting as effective barriers to dispersal and promoting allopatric speciation. However, our results suggest that this hypothesis might not be valid for endemic species with widespread distributions across several islands (e.g. *Canarina*), for whom stretches of ocean are apparently less of a barrier than topographic relief within volcanic islands. Phylogeographical studies on other MIEs (multiple island endemics) are needed to confirm this hypothesis. The palaeo-islands of Tenerife have probably acted as both genetic refugia and sources of new diversity within- and between-islands. The preservation of genotypes that became extinct everywhere else and the topographic complexity of the palaeo-islands makes them potential 'phylogeographical hotspots' (Médail & Diadema 2009) and reservoirs of unique genetic diversity, whose conservation should be prioritized.

Acknowledgements

We are grateful to Mitchell Cruzan and four anonymous reviewers, whose comments helped to significantly improve our manuscript. We would like to thank Alberto Herrero, Emilio Cano and Fatima Durán for their laboratory assistance. Field work could not have been conducted without the cooperation of the staff at the Jardín Botánico Canario Viera y Clavijo (Gran Canaria); Jacinto Leralta and Ángel Fernández (La Gomera); Félix Manuel Medina from the Cabildo of La Palma, and the Cabildo of Tenerife and El Hierro are thanked for helping with sampling permits and accommodation during field expeditions. We are grateful to Gonzalo Nieto-Feliner for a critical review of the manuscript. We also thank Águedo Marrero, Alfredo Valido, Aurelio Martín, Carlos García-Verdugo, Isabel Gómez, Juli Caujapé-Castells, Lisa Pokorny, María Candelaria Rodríguez-Rodríguez, Marta Martínez, Moisés Soto, Nadir Álvarez, Nils Arrigo, Rubén Barone, Xavier Picó and Yulán Úbeda for all the fruitful discussions. This work was funded by Spanish Ministry of Science through projects CGL2009-1332-C03-01 and CGL2012-40129-C02-01). M. Alarcón was funded by the JAE-Doc program (CSIC/FSE). MM and VC were supported by MINECO FPI predoctoral fellowships (BES-2010-037261 and BES-2013-065389, respectively).

References

Ancochea E, Fuster J, Ibarrola E *et al.* (1990) Volcanic evolution of the island of Tenerife (Canary Islands) in the light of new K-Ar data. *Journal of Volcanology and Geothermal Research*, **44**, 231–249.

Ancochea E, Huertas MJ, Cantagrel JM *et al.* (1999) Evolution of the Cañadas edifice and its implications for the origin of the Cañadas Caldera (Tenerife, Canary Islands). *Journal of Volcanology and Geothermal Research*, **88**, 177–199.

Arevalo JR, Delgado JD, Fernández-Palacios JM (2007) Variation in fleshy fruit fall composition in an island laurel forest of the Canary Islands. *Acta Oecologica*, **32**, 152–160.

Avise JC (2000) *Phylogeography: The History and Formation of Species*. Harvard University Press, Cambridge, Massachusetts.

Baele G, Lemey P, Bedford T, Rambaut A, Suchard MA, Alekseyenko AV (2012) Improving the accuracy of demographic and molecular clock model comparison while accommodating phylogenetic uncertainty. *Molecular Biology and Evolution*, **29**, 2157–2167.

Báez M, Martín JL, Oromí P (2001) Diversidad taxonómica terrestre. En (Ed.) Fernández Palacios JM & Martín Esquivel JL, Naturaleza de las Islas Canarias. Ecología y Conservación.Tenerife: Ed. Turquesa, 65–76.

Bloomquist EW, Lemey P, Suchard MA (2010) Three roads diverged? Routes to phylogeographic inference. *Trends in Ecology & Evolution*, **25**, 626–632.

Bonin A, Bellemain E, Bronken Eidesen P, Pompanon F, Brochmann C, Taberlet P (2004) How to track and assess genotyping errors in population genetics studies. *Molecular Ecology*, **13**, 3261–3273.

Bottin L, Verhaegen D, Tassin J, Olivieri I, Vaillant A, Bouvet JM (2005) Genetic diversity and population structure of an insular tree, Santalum austrocaledonicum in New Caledonian archipelago. *Molecular Ecology*, **14**, 1979–1989.

Boulesteix T, Hildenbrand A, Gillot PY, Soler V (2012) Eruptive response of oceanic islands to giant landslides: new insights from the geomorphologic evolution of the Teide-Pico Viejo volcanic complex (Tenerife, Canary). *Geomorphology*, **138**, 61–73.

Boulesteix T, Hildenbrand A, Soler V, Quidelleur X, Gillot PY (2013) Coeval giant landslides in the Canary Islands: implications for global, regional and local triggers of giant flank collapses on oceanic volcanoes. *Journal of Volcanology and Geothermal Research*, **257**, 90–98.

Brown RP, Campos-Delgado R, Pestano J (2000) Mitochondrial DNA evolution and population history of the Tenerife skink *Chalcides viridanus*. *Molecular Ecology*, **9**, 1061–1067.

Brown RP, Hoskisson PA, Welton JH, Baez M (2006) Geological history and within-island diversity: a debris avalanche and the Tenerife lizard *Gallotia galloti*. *Molecular Ecology*, **15**, 3631–3640.

Cantagrel JM, Arnaud NO, Ancochea E, Fúster JM, Huertas MJ (1999) Repeated debris avalanches on Tenerife and genesis of Las Cañadas caldera wall (Canary Islands). *Geology*, **27**, 739–742.

Carine MA, Schaefer H (2010) The Azores diversity enigma: why are there so few Azorean endemic flowering plants and why are they so widespread? *Journal of Biogeography*, **37**, 77–89.

Carracedo J (1994) The Canary Islands: an example of structural control on the growth of large oceanic-island volcanoes. *Journal of Volcanology and Geothermal Research*, **60**, 225–241.

Carracedo JC (2014) The Teide volcano, Tenerife, Canary Islands. In: *Landscapes and Landforms of Spain*, pp, 257–272. Springer, Netherlands

Carracedo JC, Day S (2002) *Canary Islands*. Dunedin Academic Press Ltd, Lewiston, New York.

Carracedo JC, Day S, Guillou H, Rodríguez Badiola E, Canas JA, Pérez Torrado FJ (1998) Hotspot volcanism close to a passive continental margin: the Canary Islands. *Geological Magazine*, **135**, 591–604.

Carracedo JC, Badiola ER, Guillou H *et al.* (2007) Eruptive and structural history of Teide Volcano and rift zones of Tenerife, Canary Islands. *Geological Society of America Bulletin*, **119**, 1027–1051.

Carson HL, Templeton AR (1984) Genetic revolutions in relation to speciation phenomena: the founding of new populations. *Annual Review of Ecology and Systematics*, **15**, 97–131.

Carson HL, Lockwood JP, Craddock EM (1990) Extinction and recolonization of local populations on a growing shield volcano. *Proceedings of the National Academy of Sciences*, **87**, 7055–7057.

Caujapé-Castells J (2010) General G_{ST} and θ inflation due to biased intra-population sampling, and its consequences for the conservation of the Canarian Flora. *Conservation Genetics*, **11**, 709–720.

Clement M, Posada DCKA, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, **9**, 1657–1659.

Contreras-Díaz HG, Moya O, Oromí P, Juan C (2007) Evolution and diversification of the forest and hypogean ground-beetle genus *Trechus* in the Canary Islands. *Molecular Phylogenetics and Evolution*, **42**, 687–699.

Cox SC, Carranza S, Brown RP (2010) Divergence times and colonization of the Canary Islands by *Gallotia* lizards. *Molecular Phylogenetics and Evolution*, **56**, 747–757.

Delgado JD, Arévalo JR, Fernández-Palacios JM (2007) Road edge effect on the abundance of the lizard *Gallotia galloti* (Sauria: Lacertidae) in two Canary Islands forests. *Biodiversity and Conservation*, **16**, 2949–2963.

Dimitrov D, Arnedo MA, Ribera C (2008) Colonization and diversification of the spider genus *Pholcus* Walckenaer, 1805 (Araneae, Pholcidae) in the Macaronesian archipelagos: evidence for long-term occupancy yet rapid recent speciation. *Molecular Phylogenetics and Evolution*, **48**, 596–614.

Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, **7**, 214.

Earl DA (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, **4**, 359–361.

Eckert CG, Samis KE, Lougheed SC (2008) Genetic variation across species' geographical ranges: the central–marginal hypothesis and beyond. *Molecular Ecology*, **17**, 1170–1188.

Emerson BC, Oromí P (2005) Diversification of the forest beetle genus *Tharpius* on the Canary Island, and the evolutionary origins of island endemics. *Evolution*, **59**, 586–598.

Emerson BC, Oromı́ P, Hewitt GM (1999) MtDNA phylogeography and recent intra-island diversification among Canary Island *Calathus* beetles. *Molecular Phylogenetics and Evolution*, **13**, 149–158.

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

Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, **1**, 47.

Falush D, Stephens M, Pritchard JK (2007) Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Molecular Ecology Notes*, **7**, 574–578.

Feliner GN (2014) Patterns and processes in plant phylogeography in the Mediterranean basin. A review. *Perspectives in Plant Ecology, Evolution and Systematics*, **16**, 265–278.

Fernández-Mazuecos M, Vargas P (2011) Genetically depauperate in the continent but rich in oceanic islands: *Cistus monspeliensis* (Cistaceae) in the Canary Islands. *PLoS One*, **6**, e17172.

Fernández-Palacios JM, de Nascimento L, Otto R *et al.* (2011) A reconstruction of Palaeo-Macaronesia, with particular reference to the long-term biogeography of the Atlantic island laurel forests. *Journal of Biogeography*, **38**, 226–246.

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

Francisco-Ortega J, Fuertes-Aguilar J, Kim SC, Santos-Guerra A, Crawford DJ, Jansen RK (2002) Phylogeny of the Macaronesian endemic *Crambe* section Dendrocrambe (Brassicaceae) based on internal transcribed spacer sequences of nuclear ribosomal DNA. *American Journal of Botany*, **89**, 1984–1990.

Fu YX (1996) New statistical tests of neutrality for DNA samples from a population. *Genetics*, **143**, 557–570.

Fu YX, Li WH (1993) Statistical tests of neutrality of mutations. *Genetics*, **133**, 693–709.

García-Verdugo C, Forrest AD, Fay MF, Vargas P (2010) The relevance of gene flow in metapopulation dynamics of an oceanic island endemic, *Olea europaea* subsp. guanchica. *Evolution*, **64**, 3525–3536.

García-Verdugo C, Baldwin BG, Fay MF, Caujapé-Castells J (2014) Life history traits and patterns of diversification in oceanic archipelagos: a meta-analysis. *Botanical Journal of the Linnean Society*, **174**, 334–348.

García-Verdugo C, Sajeva M, La Mantia T, Harrouni C, Msanda F, Caujapé-Castells J (2015) Do island plant populations really have lower genetic variation than mainland populations? Effects of selection and distribution range on genetic diversity estimates. *Molecular Ecology*, **24**, 726–741.

Gavin DG, Fitzpatrick MC, Gugger PF *et al.* (2014) Climate refugia: joint inference from fossil records, species distribution models and phylogeography. *New Phytologist*, **204**, 37–54.

Gillespie R (2004) Community assembly through adaptive radiation in Hawaiian spiders. *Science*, **303**, 356–359.

Gillespie RG, Clague DA (2009) *Encyclopedia of Islands*. University of California Press, Berkeley, California.

Gillespie RG, Roderick GK (2002) Arthropods on islands: colonization, speciation, and conservation. *Annual Review of Entomology*, **47**, 595–632.

Gillespie RG, Roderick GK (2014) Evolution: geology and climate drive diversification. *Nature*, **509**, 297–298.

Gómez A, González-Martínez SC, Collada C, Climent J, Gil L (2003) Complex population genetic structure in the endemic Canary Island pine revealed using chloroplast microsatellite markers. *Theoretical and Applied Genetics*, **107**, 1123–1131.

Gübitz T, Thorpe RS, Malhotra A (2000) Phylogeography and natural selection in the Tenerife gecko *Tarentola delalandii*: testing historical and adaptive hypotheses. *Molecular Ecology*, **9**, 1213–1221.

Gübitz T, Thorpe RS, Malhotra A (2005) The dynamics of genetic and morphological variation on volcanic islands. *Proceedings of the Royal Society B: Biological Sciences*, **272**, 751–757.

Guillot G, Estoup A, Mortier F, Cosson JF (2005) A spatial statistical model for landscape genetics. *Genetics*, **170**, 1261–1280.

Guillou H, Carracedo JC, Paris R, Pérez Torrado FJ (2004) Implications for the early shield-stage evolution of Tenerife from K/Ar ages and magnetic stratigraphy. *Earth and Planetary Science Letters*, **222**, 599–614.

Guzmán B, Vargas P (2010) Unexpected synchronous differentiation in Mediterranean and Canarian *Cistus* (Cistaceae). *Perspectives in Plant Ecology, Evolution and Systematics*, **12**, 163–174.

Harpending H (1994) Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. *Human Biology*, **66**, 591–600.

Harter DE, Irl SD, Seo B *et al.* (2015) Impacts of global climate change on the floras of oceanic islands–Projections, implications and current knowledge. *Perspectives in Plant Ecology, Evolution and Systematics*, **17**, 160–183.

Hewitt G (2000) The genetic legacy of the Quaternary ice ages. *Nature*, **405**, 907–913.

Ho SY, Phillips MJ, Cooper A, Drummond AJ (2005) Time dependency of molecular rate estimates and systematic overestimation of recent divergence times. *Molecular Biology and Evolution*, **22**, 1561–1568.

Hochkirch A, Goerzig Y (2009) Colonization and speciation on volcanic islands: phylogeography of the flightless grasshopper genus *Arminda* (Orthoptera, Acrididae) on the Canary Islands. *Systematic Entomology*, **34**, 188–197.

von Humboldt A (1814) *Voyage aux regions equinoxiales du nouveaux continent, fait en 1799–1804 par Al. de Humboldt et A. Bonpland*. Part 1. París.

Huson DH, Bryant D (2006) Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution*, **23**, 254–267.

Jakob SS, Blattner FR (2006) A chloroplast genealogy of *Hordeum* (Poaceae): long-term persisting haplotypes, incomplete lineage sorting, regional extinction, and the consequences for phylogenetic inference. *Molecular Biology and Evolution*, **23**, 1602–1612.

Jones KE, Reyes-Betancort JA, Hiscock SJ, Carine MA (2014) Allopatric diversification, multiple habitat shifts, and hybridization in the evolution of *Pericallis* (Asteraceae), a Macaronesian endemic genus. *American Journal of Botany*, **101**, 637–651.

Juan C, Ibrahim KM, Oromi P, Hewitt GM (1996) Mitochondrial DNA sequence variation and phylogeography of *Pimelia* darkling beetles on the island of Tenerife (Canary Islands). *Heredity*, **77**, 589–598.

Juan C, Emerson BC, Oromí P, Hewitt GM (2000) Colonization and diversification: towards a phylogeographic synthesis for the Canary Islands. *Trends in Ecology & Evolution*, **15**, 104–109.

Katoh K, Misawa K, Kuma KI, Miyata T (2002) MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research*, **30**, 3059–3066.

Kelchner SA (2000) The evolution of non-coding chloroplast DNA and its application in plant systematics. *Annals of the Missouri Botanical Garden*, **87**, 482–498.

Kim SC, McGowen MR, Lubinsky P, Barber JC, Mort ME, Santos-Guerra A (2008) Timing and tempo of early and successive adaptive radiations in Macaronesia. *PLoS One*, **3**, e2139.

Knowles LL (2008) Why does a method that fails continue to be used? *Evolution*, **62**, 2713–2717.

Kunkel G (1991) Flora y vegetación del Archipiélago Canario. Tratado florístico. 2ª parte. Dicotiledóneas. 312 pp. Edirca. Madrid.

Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, **25**, 1451–1452.

Lyell C (1855) *A Manual of Elementary Geology*. John Murray, London.

Lynch M, Milligan BG (1994) Analysis of population genetic structure with RAPD markers. *Molecular Ecology*, **3**, 91–99.

Macías-Hernández N, Bidegaray-Batista L, Emerson BC, Oromí P, Arnedo M (2013) The imprint of geologic history on within-island diversification of woodlouse-hunter spiders (Araneae, Dysderidae) in the Canary Islands. *Journal of Heredity*, **104**, 341–356.

Mairal M, Pokorny L, Aldasoro JJ, Alarcón M, Sanmartín I (2015) Ancient vicariance and climate-driven extinction continental-wide disjunctions in Africa: the case of the Rand Flora genus Canarina (Campanulaceae). *Molecular Ecology*, **24**, 1335–1354.

Manni F, Guerard E, Heyer E (2004) Geographic patterns of (genetic, morphologic, linguistic) variation: how barriers can be detected by using Monmonier's algorithm. *Human Biology*, **76**, 173–190.

Manolopoulou I, Emerson BC (2012) Phylogeographic ancestral inference using the coalescent model on haplotype trees. *Journal of Computational Biology*, **19**, 745–755.

Manolopoulou I, Legarreta L, Emerson BC, Brooks S, Tavaré S (2011) A Bayesian approach to phylogeographic clustering. *Interface Focus*, **1**, 909–921.

Mansion G, Parolly G, Crowl AA *et al.* (2012) How to handle speciose clades? Mass taxon-sampling as a strategy towards illuminating the natural history of *Campanula* (Campanuloideae). *PLoS One*, **7**, e50076.

Médail F, Diadema K (2009) Glacial refugia influence plant diversity patterns in the Mediterranean basin. *Journal of Biogeography*, **36**, 1333–1345.

Médail F, Quézel P (1997) Hot-spots analysis for conservation of plant biodiversity in the Mediterranean basin. *Annals of the Missouri Botanical Garden*, **84**, 112–127.

Moya Ó, Contreras-Díaz HG, Oromí P, Juan C (2004) Genetic structure, phylogeography and demography of two ground-beetle species endemic to the Tenerife laurel forest (Canary Islands). *Molecular Ecology*, **13**, 3153–3167.

Nei M, Li WH (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences*, **76**, 5269–5273.

Nylander J (2004) *MrModeltest v2. Program Distributed by the Author*. Evolutionary Biology Centre, Uppsala University 2, Uppsala.

Olesen JM, Alarcon M, Ehlers BK, Aldasoro JJ, Roquet C (2012) Pollination, biogeography and phylogeny of oceanic island bellflowers (Campanulaceae). *Perspectives in Plant Ecology, Evolution and Systematics*, **14**, 169–182.

Padilla DP, González-Castro A, Nogales M (2012) Significance and extent of secondary seed dispersal by predatory birds on oceanic islands: the case of the Canary archipelago. *Journal of Ecology*, **100**, 416–427.

Patiño J, Carine M, Mardulyn P *et al.* (2015) Approximate Bayesian computation reveals the crucial role of oceanic islands for the assembly of continental biodiversity. *Systematic Biology*, **64**, 579–589.

de Paz JP, Caujapé-Castells J (2013) A review of the allozyme data set for the Canarian endemic flora: causes of the high genetic diversity levels and implications for conservation. *Annals of Botany*, **111**, 1059–1073.

Pokorny L, Oliván G, Shaw AJ (2011) Phylogeographic patterns in two southern hemisphere species of Calyptrochaeta (Daltoniaceae, Bryophyta). *Systematic Botany*, **36**, 542–553.

Posada D, Crandall KA (2001) Intraspecific gene genealogies: trees grafting into networks. *Trends in Ecology & Evolution*, **16**, 37–45.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.

Puppo P, Curto M, Velo-Antón G, Pérez de Paz PL, Meimberg H (2014) The influence of geological history on diversification in insular species: genetic and morphological patterns of Micromeria Benth. (Lamiaceae) in Tenerife (Canary archipelago). *Journal of Biogeography*, **41**, 1871–1882.

Puppo P, Curto M, Gusmão-Guedes J *et al.* (2015) Molecular phylogenetics of *Micromeria* (Lamiaceae) in the Canary Islands, diversification and inter-island colonization patterns inferred from nuclear genes. *Molecular Phylogenetics and Evolution*, **89**, 160–170.

Rambaut A, Drummond A (2009) FigTree v1. 3.1. Computer program and documentation distributed by the author at http://tree.bio.ed.ac.uk/software.

Rambaut A, Drummond AJ, Suchard M (2007) *Tracer v1. 6.* http://tree.bio.ed.ac.uk/software/tracer/.

Reyes-Betancort JA, Santos-Guerra A, Guma IR, Humphries CJ, Carine MA (2008) Diversity, rarity and the evolution and conservation of the Canary Islands endemic flora. *Anales del Jardín Botánico de Madrid*, **65**, 25.

Rijsdijk KF, Hengl T, Norder SJ *et al.* (2014) Quantifying surface-area changes of volcanic islands driven by Pleistocene sea-level cycles: biogeographical implications for the Macaronesian archipelagos. *Journal of Biogeography*, **41**, 1242–1254.

Rodríguez A, Nogales M, Rumeu B, Rodríguez B (2008) Temporal and spatial variation in the diet of the endemic lizard *Gallotia galloti* in an insular Mediterranean scrubland. *Journal of Herpetology*, **42**, 213–222.

Rodríguez-Gonzalez A, Fernandez-Turiel JL, Perez-Torrado FJ *et al.* (2009) The Holocene volcanic history of Gran Canaria island: implications for volcanic hazards. *Journal of Quaternary Science*, **24**, 697–709.

Rodríguez-Rodríguez MC, Valido A (2011) Consequences of plant-pollinator and floral-herbivore interactions on the reproductive success of the Canary Islands endemic *Canarina canariensis* (Campanulaceae). *American Journal of Botany*, **98**, 1465–1474.

Rohlf F (1998) *NTSYS-pc Version 2.0. Numerical Taxonomy and Multivariate Analysis System*. Exeter Software, Setauket, New York.

Salomone N, Emerson BC, Hewitt GM, Bernini F (2002) Phylogenetic relationships among the Canary Island Steganacaridae (Acari, Oribatida) inferred from mitochondrial DNA sequence data. *Molecular Ecology*, **11**, 79–89.

Sanmartín I, Van Der Mark P, Ronquist F (2008) Inferring dispersal: a Bayesian approach to phylogeny-based island biogeography, with special reference to the Canary Islands. *Journal of Biogeography*, **35**, 428–449.

Santos-Guerra A (2001) Flora vascular nativa. In: *Naturaleza de las Islas Canarias. Ecología y Conservación* (eds Fernández-Palacios JM, Martín-Esquivel JM), pp. 185–198. Publicaciones Turquesa, Santa Cruz de Tenerife.

Stuessy TF, Takayama K, López-Sepúlveda P, Crawford DJ (2014) Interpretation of patterns of genetic variation in endemic plant species of oceanic islands. *Botanical Journal of the Linnean Society*, **174**, 276–288.

Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, **123**, 585–595.

Templeton AR, Crandall KA, Sing CF (1992) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics*, **132**, 619–633.

Thorpe RS, Black H, Malhotra A (1996) Matrix correspondence tests on the DNA phylogeny of the Tenerife lacertid elucidate both historical causes and morphological adaptation. *Systematic Biology*, **45**, 335–343.

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

Tzedakis PC, Emerson BC, Hewitt GM (2013) Cryptic or mystic? Glacial tree refugia in northern Europe. *Trends in Ecology & Evolution*, **28**, 696–704.

Valido A, Nogales M, Medina FM (2003) Fleshy fruits in the diet of Canarian lizards *Gallotia galloti* (Lacertidae) in a xeric habitat of the island of Tenerife. *Journal of Herpetology*, **37**, 741–747.

Vekemans X, Beauwens T, Lemaire M, Roldán-Ruiz I (2002) Data from amplified fragment length polymorphism (AFLP) markers show indication of size homoplasy and of a relationship between degree of homoplasy and fragment size. *Molecular Ecology*, **11**, 139–151.

Vitales D, García-Fernández A, Pellicer J *et al.* (2014a) Key processes for *Cheirolophus* (Asteraceae) diversification on oceanic islands inferred from AFLP data. *PLoS One*, **9**, e113207.

Vitales D, Garnatje T, Pellicer J, Vallès J, Santos-Guerra A, Sanmartín I (2014b) The explosive radiation of *Cheirolophus* (Asteraceae, Cardueae) in Macaronesia. *BMC Evolutionary Biology*, **141**, 118.

Vos P, Hogers R, Bleeker M *et al.* (1995) AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research*, **23**, 4407–4414.

Wang Q, Zhou SL, Hong DY (2013) Molecular phylogeny of the platycodonoid group (Campanulaceae s. str.) with special reference to the circumscription of *Codonopsis*. *Taxon*, **62**, 498–504.

Whitlock R, Hipperson H, Mannarelli M, Butlin RK, Burke T (2008) An objective, rapid and reproducible method for scoring AFLP peak-height data that minimizes genotyping error. *Molecular Ecology Resources*, **8**, 725–735.

Whittaker RJ, Ladle RJ, Araújo MB, Fernández-Palacios JM, Delgado J, Arévalo J (2007) The island immaturity–speciation pulse model of island evolution: an alternative to the "diversity begets diversity" model. *Ecography*, **30**, 321–327.

Zaczek K, Troll V, Cachao M *et al.* (2015) Nannofossils in 2011 El Hierro eruptive products reinstate plume model for Canary Islands. *Scientific Reports*, **5**, 1–5.

Zhivotovsky LA (1999) Estimating population structure in diploids with multilocus dominant DNA markers. *Molecular Ecology*, **8**, 907–913.

Zink RM, Barrowclough GF (2008) Mitochondrial DNA under siege in avian phylogeography. *Molecular Ecology*, **17**, 2107–2121.

M.M., M.A., J.J.A. and I.S. designed the study; M.M. carried out the field work, M.A. and J.J.A. contributed some samples; M.M. performed the research and analysed the data under the supervision of M.A. and I.S.; I.M., M.M., I.S. and V.C. performed the BPEC analyses; M.M. wrote the paper with help from I.S., M.A. and J.J.A., M.M. and I.S. revised the manuscript. All authors contributed with comments and approved the final version.

Data accessibility

DNA sequences: Genbank Accession nos KP797991–KP798432.

GenBank accessions, sampling locations and/or online-only appendices uploaded as online supplemental material.

NEXUS files for the concatenated cpDNA data set and original AFLPs Matrix of *Canarina canariensis*, Dryad doi:10.5061/dryad.j4103.

Supporting information

Additional supporting information may be found in the online version of this article.

Table S1 Plant species which are restricted and shared between two or more of the ancient static areas considered: the three paleo-islands of Tenerife and La Gomera.

Table S2 Sampling information and GenBank accession numbers for all of the taxa included in this study.

Table S3 Results from the sensitivity analyses to assess confidence in our age estimates.

Table S4 Haplotypes of *Canarina canariensis*, as determined by using the variable sites and indels found in the sixteen *C. canariensis* populations with markers *Pet*B-*Pet*D (937 bp), *Rpl*32-*trn*L (654 bp) and *trn*S-*trn*G (658 bp).

Table S5 G_{ST} and number of migrants (in parentheses) obtained from cloroplastic data in *Canarina canariensis*.

Table S6 Mean ages and 95% HPD confidence intervals from the BEAST analysis.

Table S7 Results from the DNA polymorphism, neutrality test and mismatch raggedness for *Canarina canariensis* haplotypes.

Table S8 Pairwise F_{ST} between populations of *Canarina canariensis* with nuclear data.

Table S9 Hierarchical analysis of molecular variance (AMOVA) for *C. canariensis* based on allelic variation at different levels: (a) amongst groups, (b) amongst populations within groups and (c) within populations.

Fig. S1 Clusters and networks obtained with bayesian phylogeographic and environmental clustering (BPEC) with plastid DNA sequences.

Fig. S2 Mismatch distribution (left handside) and Extended Bayesian Skyline Plot analysis (right handside) for the different groups.

Fig. S3 The MCC tree, from the BEAST analysis, showing nodes with mean ages and the 95% HPD confidence intervals (values specified in Table S5).

Fig. S4 Principal coordinates analysis of AFLP data for individuals of *C. canariensis*.

Fig. S5 The estimated probability of the likelihood function according to the Evanno method for: (a) STRUCTURE analysis for *C. canariensis*; (b) STRUCTURE analysis using only Teneriffean populations.

CHAPTER 4

A tale of two forests: Population decline threating 1 biodiversity in evergreen forest archipelagos in two relict Afro-Macaronesian species (*Canarina*, Campanulaceae)

Mario Mairal, Juli Caujapé-Castells, Myriam Heuertz, Loïc Pellissier,
Ruth Jaén-Molina, Nadir Álvarez, Mila Pakjovic, Isabel Sanmartín

This chapter is going to be submitted to the peer-review journal
Conservation Biology

Annals of Botany **XX**: 1–13, 00
doi: 10.1093/aob/mcy107, available online at www.academic.oup.com/aob

ANNALS OF BOTANY
Founded 1887

A tale of two forests: ongoing aridification drives population decline and genetic diversity loss at continental scale in Afro-Macaronesian evergreen-forest archipelago endemics

Mario Mairal[1,2,*], Juli Caujapé-Castells[2], Loïc Pellissier[3,4], Ruth Jaén-Molina[2], Nadir Álvarez[5], Myriam Heuertz[6] and Isabel Sanmartín[1]

[1]Real Jardín Botánico (RJB), CSIC, Plaza de Murillo 2, 28014 Madrid, Spain, [2] Departamento de Biodiversidad Molecular y Banco de ADN, Jardín Botánico 'Viera y Clavijo' – Unidad Asociada CSIC (Cabildo de Gran Canaria), Camino del Palmeral 15, 35017, Las Palmas de Gran Canaria, Spain, [3]Landscape Ecology, Institute of Terrestrial Ecosystems, ETH Zürich, CH-8092 Zürich, Switzerland, [4]Swiss Federal Research Institute WSL, CH-8903 Birmensdorf, Switzerland, [5]Department of Ecology and Evolution, Institute of Biology, University of Lausanne, Biophore Dorigny, CH-1015 Lausanne, Switzerland and [6]Biogeco, INRA, University of Bordeaux, F-33610 Cestas, France
**For correspondence. Present address: Department of Botany and Zoology, Stellenbosch University, Private Bag X1, Matieland, 7602, South Africa. E-mail mariomairal@gmail.com.*

Received: 30 December 2017 Returned for revision: 2 March 2018 Editorial decision: 22 May 2018 Accepted: 25 May 2018

- **Background and Aims** Various studies and conservationist reports have warned about the contraction of the last subtropical Afro-Macaronesian forests. These relict vegetation zones have been restricted to a few oceanic and continental islands around the edges of Africa, due to aridification. Previous studies on relict species have generally focused on glacial effects on narrow endemics; however, little is known about the effects of aridification on the fates of previously widespread subtropical lineages.
- **Methods** Nuclear microsatellites and ecological niche modelling were used to understand observed patterns of genetic diversity in two emblematic species, widely distributed in these ecosystems: *Canarina eminii* (a palaeoendemic of the eastern Afromontane forests) and *Canarina canariensis* (a palaeoendemic of the Canarian laurel forests). The software DIYABC was used to test alternative demographic scenarios and an ensemble method was employed to model potential distributions of the selected plants from the end of the deglaciation to the present.
- **Key Results** All the populations assessed experienced a strong and recent population decline, revealing that locally widespread endemisms may also be alarmingly threatened.
- **Conclusions** The detected extinction debt, as well as the extinction spiral to which these populations are subjected, demands urgent conservation measures for the unique, biodiversity-rich ecosystems that they inhabit.

Key words: Islands, nuclear microsatellites, subtropical flora, decline, aridification, extinction, genetic conservation.

INTRODUCTION

Rare species are defined as those characterized by small population sizes, narrow geographic distributions and high habitat specificity, or any combination of these criteria (Rabinowitz, 1981). These species often display low genetic diversity, which makes them vulnerable to stochastic demographic phenomena (Kruckeberg and Rabinowitz, 1985; Hewitt, 2004). However, this relationship between rarity and genetic diversity is not an axiom: rare species may present high levels of genetic diversity (Stebbins, 1980; Gitzendanner and Soltis, 2000; Pérez de Paz and Caujapé-Castells, 2013; García-Verdugo *et al.*, 2015), whereas some widespread and common species, such as exotic invasive plants, may exhibit low levels of genetic diversity (Tsutsui *et al.*, 2000).

Relict species are rare species whose distribution ranges have been reduced to a few populations due to different processes (e.g. aridification), but with a degree of threat that gives them a great value in conservation biology (Habel and Assmann, 2009). Especially interesting are palaeoendemics, i.e. relict species that have become rare due to widespread extinction in their original distribution range (Stebbins and Major, 1965), and which frequently act as reservoirs of phylogenetic exclusivity, as the only surviving representatives of larger clades (Cronk, 1992; Faith, 1992). Biologists often characterize these species as 'living fossils', an irreplaceable heritage in the Tree of Life (Grandcolas *et al.*, 2014). The need to conserve these rare and relict species and their habitat-restricted communities has increased in the current context of rapid human-induced climate change.

Studies on relict species have generally focused on glacial relicts, which are now constrained to some reduced ranges in mountain peaks or high latitudes (Petit *et al.*, 2003; Hewitt, 2004). By contrast, relict species that have become rare as a result of aridification have received far less attention. Genetic studies on these species, however, are of particular relevance because they may provide insights into the capacity for resilience and the foreseeable evolutionary fates of species subject to global warming.

Here we focus on one such taxon: the angiosperm plant genus *Canarina*, belonging to the bellflower family, Campanulaceae (tribe Platycodoneae). This genus comprises only three species separated by >7000 km across the Sahara desert: *C. canariensis*, associated with the endemic laurel forests of the Canary Islands, and *C. eminii* and *C. abyssinica*, which are endemic to the relict Afromontane forests of East Africa. *Canarina* is the only African representative of Platycodoneae, an early-diverging tribe within family Campanulaceae (Roquet *et al.*, 2008), while the remaining members of the tribe are distributed throughout Central and East Asia (Wang *et al.*, 2013). Thus, *Canarina* complies with the criterion of phylogenetic exclusivity in a relict rare species (Faith, 1992).

Mairal *et al.* (2015a) reconstructed the biogeographic history of *Canarina* and showed that ancestors of this genus migrated from Asia to East Africa in the Middle Miocene. They linked the disjunct sister-group relationship between *C. canariensis* and *C. eminii* (dated ~7–6 Ma) to large-scale extinction and a gradual aridification process that would have started in North Africa in the Late Miocene (Sepulchre *et al.*, 2006; Senut *et al.*, 2009). This was followed by cycles of contraction and expansion affecting both subtropical and tropical taxa (Maley, 2000; Bonnefille, 2011), in which mountainous areas in the Sahara served as refuges during arid periods (Osborne *et al.*, 2008). This Afro-Macaronesian biogeographic pattern has been observed in many other angiosperm genera, and is known as the Rand Flora pattern (Sanmartín *et al.*, 2010; Pokorny *et al.*, 2015; Mairal *et al.*, 2017*a*). Thus, *Canarina* also fulfils the criterion of a palaeoendemic relict – the palaeoendemic character can also be seen in the *Canarina* populations of both species, in which reservoirs of ancient genetic diversity appear in the oldest massifs and palaeo-islands (Mairal *et al.*, 2015b, 2017b).

Rare or relictual species are also generally characterized by high habitat specificity (Kruckeberg and Rabinowitz, 1985), either because the species have low environmental tolerance or a narrow niche, or because their habitats are currently under contraction. The habitats now occupied by *Canarina* (the Canarian laurel forest and the isolated patches of subtropical Afromontane forest in East African 'sky islands') have often been interpreted as the last remnants of a more humid vegetation that extended over the Mediterranean Basin and North Africa during the Early–Middle Miocene (Heald, 1951; Axelrod and Raven, 1978; White, 1983; Bramwell, 1985; McCormack *et al.*, 2009; Fernández-Palacios *et al.*, 2011); but see Kondraskov *et al.* (2015) for a different view on the relict character of the Canarian laurel forests. Plant lineages that took shelter in these enclaves have acquired a high degree of relictualism and contribute substantially to the exceptional endemicity of these areas (Fjeldså and Lovett, 1997; Juan *et al.*, 2000), which are now classified as biodiversity hotspots; i.e. biogeographic regions whose great biodiversity is threatened by habitat loss (Myers *et al.*, 2000; Mittermeier *et al.*, 2004).

Both *C. eminii* and *C. canariensis* have widespread but disjunct distribution ranges and fulfil Rabinowitz's (1981) 'large and narrow' criterion of rarity: species that are constantly sparse in a specific habitat but over a large range. *Canarina canariensis* is a terrestrial herb that propagates by seeds and shoots from its tuber (Bramwell and Bramwell, 2001); it occurs in the endemic Canarian laurel forests and (more rarely) in the adjacent thermophilous forests in the central and western islands of the Canarian Archipelago. *Canarina eminii* propagates from a long and thick root and grows as a terrestrial or (mainly) as an epiphyte of Afromontane forest trees (*Podocarpus*, *Hagenia*, *Conopharyngia*; Hedberg, 1961) in East Africa, from Ethiopia in the north to Malawi in the south (Fig. 1).

The steep topography and patchy distribution of the Eastern African forests (divided by deserts, savannahs and cultivated lands) and the deep ocean floor that lies between islands in the Canarian Archipelago have added to a high degree of geographic and ecological isolation (Fig. 1B), manifested by high genetic differentiation among populations and a marked geographic structure (Mairal *et al.*, 2015a,b, 2017a). This great interpopulational diversity is threatened in *C. eminii* by human activities, such as logging and forest clearance (fires) linked to a rapid growth of the human population and agricultural development (Hedberg, 1961; Alemayehu, 2006), which have contributed to further fragmentation and isolation among forest patches (Fahrig, 2003). In the case of *C. canariensis*, although big patches of the laurisilva are currently under legal protection, this forest type has experienced a strong contraction during the last two centuries with respect to its estimated original range, due mostly to agricultural practices (Fernández-Palacios *et al.*, 2011; but see de Nascimento *et al.*, 2015 for an alternative view of Gran Canaria's forests).

In this study, we used molecular and ecological niche modelling approaches to explore the effects of historical climate change and recent anthropogenic activity on the demographic evolution of *C. canariensis* and *C. eminii*. We developed new molecular markers from a library of nuclear microsatellites (SSRs), and used these to document levels of genetic diversity within and among populations. We also tested the alternative scenarios of population dynamics of (1) constant size, (2) bottleneck, (3) demographic expansion and (4) decline, using approximate Bayesian computation techniques (Beaumont, 2010) based on SSR data. Finally, we employed ecological niche modelling to reconstruct the climatic requirements (niche breadth) of African and Canarian populations and project them back into the past (Last Glacial Maximum, Holocene) to understand the impact of recent climate change on the geographic distribution of populations. If climate change since the end of the deglaciation or human activities have driven the retreat of these forests, we should find a genetic signal linked to a recent demographic decrease; conversely, this signal would not appear if the demographic processes were older. A demographic decline would confirm the suspicions about the serious state of the threat to these forest patches. Accordingly, our main goals are: (1) to compare levels of population genetic diversity between these two species, and with those estimated by other markers [chloroplast sequences and amplified fragment length polymorphisms (AFLPs); Mairal *et al.*, 2015b, 2017b]; (2) to search for the presence of population decline or bottlenecks that may indicate ongoing loss of genetic diversity; (3) to relate population dynamics to changes in the potential geographic distribution of *Canarina* populations over time; and (4) to provide a genetic basis to enable the design of suitable strategies for the genetic conservation of these taxa.

MATERIALS AND METHODS

Plant material, population sampling and DNA sequencing

Several sampling field campaigns were performed in East Africa and the Canary Islands (2009–14), which led to the

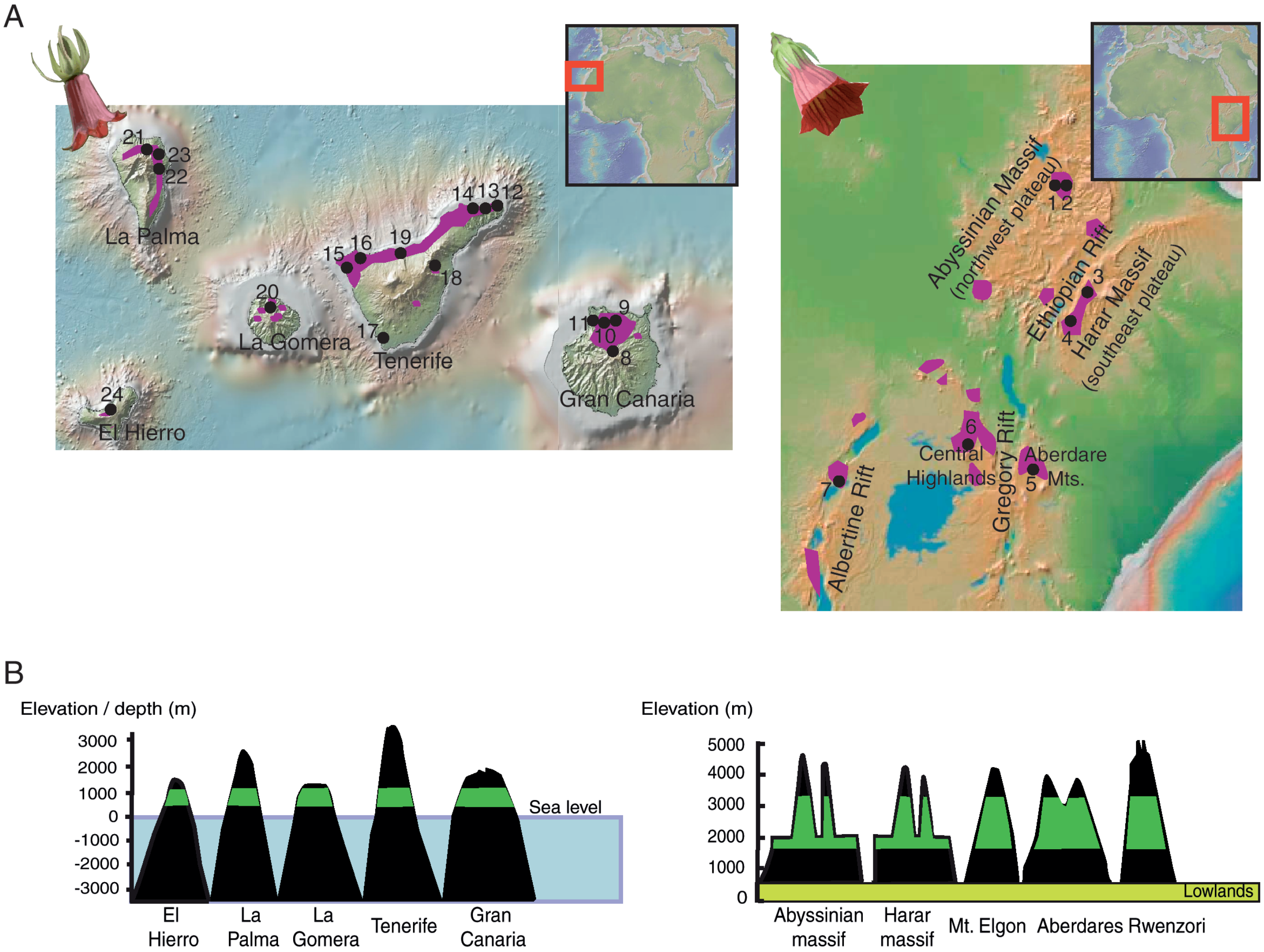


FIG. 1. Distribution and scheme of the insular systems inhabited by the studied species. (A) Geographic distribution of *C. canariensis* in the Canary Islands (left) and *C. eminii* in the Afromontane forests (right). The distribution areas are shaded in purple. Sampled populations are represented with black dots accompanied by a population code number, as specified in Table 1. (B) Schematic representations of the insular systems and the subtropical vegetation belts inhabited by the studied species: oceanic islands and laurisilva forests for *C. canariensis* (left) and sky islands and Afromontane forests for *C. eminii* (right).

collection of fresh material from 223 individuals in 24 different populations of *Canarina*: 17 of *C. canariensis* (N = 159) and seven of *C. eminii* (N = 64) (Table 1). For the largest populations, a minimum of ten samples was collected through the entire area of occupation to reduce the possible inflation in the genetic structure descriptors (Caujapé-Castells, 2010). DNA was extracted from silica-gel-dried leaves using the DNeasy Plant Mini Kit (Qiagen, CA, USA). The quality of DNA extractions was checked on 1 % agarose gels, and DNA concentration was measured using a NanoDrop spectrophotometer. All samples were diluted to 10 ng μL^{-1} prior to PCR amplification.

Development of nuclear microsatellite markers and scoring

A microsatellite library was developed for *C. canariensis* at the Savannah River Ecology Laboratory (University of Georgia). The extracted plant DNA was enriched using magnetic streptavidin beads and biotinylated oligonucleotides representing CT and GT repetitions. The library was sequenced using a Titanium Sequencing Kit (Roche Life Sciences) on the Roche 454 GS-FLX Titanium platform. We selected tri- and tetra-nucleotide repetition motifs to reduce the risk of misinterpreting as homozygotes the heterozygotes with alleles differing by a single repeat. Primer design and PCR amplifications were performed in the laboratories of Jardín Botánico Canario 'Viera y Clavijo', Unidad Asociada CSIC del Cabildo de Gran Canaria (JBCVC-CSIC).

We selected and tested 26 primer pairs for *C. canariensis*, which were also tested for cross-amplification in *C. eminii.* Fifteen loci were discarded because of amplification failure in most samples, poor electrophoretic profiles or low levels of variation. Finally, 11 SSR loci were retained for *C. canariensis* and seven for *C. eminii*; these showed good quality profiles and were used to genotype all sampled individuals from both species. DNA amplification (PCR) was performed in a final volume of 20 μL, containing 1 μL of DNA, 12.55 μL of sterilized water, 1.5 μL of 10× buffer, 0.8 μL of $MgCl^{2+}$, 0.3 μL of dNTP, 1 μL of BSA, 0.25 μL Taq DNA polymerase

TABLE 1. *Genetic diversity descriptors and tests used for each of the populations of* C. canariensis *and* C. eminii. *For Tenerife, names of paleo-islands are indicated in parentheses*

Population		N	A	H_e (s.e.)	N_{pa}	M
Canarina eminii, continental islands						
1. Gifta	Abyssinian massif	8	1.58	0.167 (0.220)	0.00	0.442
2. Dembecha	Abyssinian massif	8	1.80	0.244 (0.270)	0.07	0.349
3. Yirga	Harar massif	8	2.00	0.301 (0.349)	0.17	0.081
4. Harenna	Harar massif	9	2.28	0.287 (0.218)	0.33	0.299
5. Aberdare	Aberdare Mountains	8	1.83	0.382 (0.222)	0.01	0.128
6. Elgon	Mount Elgon	13	1.99	0.339 (0.295)	0.01	0.123
7. Rwenzori	Rwenzori Mountains	10	2.24	0.410 (0.272)	0.01	0.161
Canarina canariensis, oceanic islands						
8. Cueva Corcho	Gran Canaria	8	1.96	0.305 (0.34)	0.14	0.111
9. Tilos de Moya	Gran Canaria	12	2.04	0.325 (0.27)	0.06	0.158
10. El Sao	Gran Canaria	10	1.84	0.282 (0.26)	0.00	0.229
11. Andenes de Guayedra	Gran Canaria	12	1.72	0.292 (0.25)	0.00	0.224
12. Camino Chamuscadas	Tenerife (Anaga)	10	4.07	0.417 (0.29)	0.37	0.235
13. El Bailadero	Tenerife (Anaga)	11	3.89	0.378 (0.31)	0.03	0.164
14. Tope del Carnero	Tenerife (Anaga)	7	3.94	0.412(0.29)	0.11	0.229
15. Teno Alto	Tenerife (Teno)	8	2.37	0.400 (0.26)	0.07	0.171
16. El Palmar	Tenerife (Teno)	8	2.54	0.485 (0.27)	0.25	0.198
17. Barranco del Infierno	Tenerife (Adeje)	11	2.50	0.489 (0.23)	0.00	0.184
18. Barranco de Badajoz	Tenerife	5	3.50	0.316 (0.35)	0.02	0.178
19. Barranco Ruiz	Tenerife	6	2.20	0.411 (0.22)	0.02	0.158
20. Tamargada	Gomera	13	1.67	0.235 (0.26)	0.04	0.250
21. Los Tilos	La Palma	10	2.26	0.394 (0.25)	0.14	0.202
22. Barranco del Agua	La Palma	10	2.02	0.332 (0.24)	0.01	0.188
23. Barranco La Barata	La Palma	4	1.88	0.365 (0.26)	0.00	0.258
24. El Hierro	El Hierro	12	2.24	0.415 (0.20)	0.17	0.183

A, allelic richness; N_{pa}, number of private alleles standardized to sample size; M, Garza and Williamson's M-ratio test.

(Bioline), 1 μL of forward primer (0.2 μM), 0.8 μL of reverse primer (5 μM) and 0.8 μL of reverse-tailed primer. PCR products were labelled using the fluorescent dyes FAM, VIC, NED or PET (labels according to the original kit of Applied Biosystems), and an additional 19-bp fluorescently labelled M13 primer (5′-CACGACGTTGTAAAACGAC-3′) according to the methods of Boutin-Ganache *et al.*, (2001). A tail sequence (5′-GTGTCTT-3′) was added to the 5′ end of the reverse primer to improve adenylation (addition of AMP) and to facilitate genotyping. Samples were incubated in a Verity 96 thermocycler using the following conditions: 2 min of denaturation at 94.4 °C, 30 cycles of 30 s of denaturation at 94.4 °C, 40 s at different annealing temperatures for each locus (Supplementary Data Table S1), 30 s of elongation at 72 °C and final extension for 8 min at 72 °C. Reactions were separated on 5 % polyacrylamide gels. The amplified fragments were scored using GeneMapper 4.0 software (Applied Biosystems) with the LIZ 500 size standard. We combined automatic detection of each allele with visual inspection of each sample, following Dewoody (2006) to reduce scoring errors. Finally, we identified peak profiles for each locus and allele, and we assigned a genotype to each individual.

To detect possible genotyping errors (stuttering, null alleles and allelic dropout), we analysed the genotyping matrix with the programs MICRO-CHECKER 2.2.1. (Van Oosterhout *et al.*, 2004) and INEst 1.0 (Chybicki and Burczyk, 2009). A single locus in *C. eminii* and one locus in *C. canariensis* displayed null alleles with frequencies >0.05 and high F_{IS} values, and were therefore removed from further analyses. Other loci were excluded due to low quality of electrophoretograms, especially in *C. eminii*, where several loci were removed due to the effect of cross-amplification. The final data matrices contained nine loci for *C. canariensis* and six loci for *C. eminii*.

Genetic diversity and population structure

Deviations from Hardy–Weinberg equilibrium and linkage disequilibrium were tested using POPGENE version 4.2 (Yeh *et al.*, 1997). Standard genetic diversity statistics, such as expected heterozygosity (H_e), were calculated using the software Arlequin (Excoffier *et al.*, 2005). To account for sample size variation among populations, we estimated the allelic richness and the number of private alleles using rarefaction analysis implemented in the software HP-Rare v.1.0 (Kalinowski, 2005). To identify the signal of old or severe bottlenecks, we used Garza and Williamson's M-ratio test (GW test) (2001), which is better suited to the detection of these events than methods based on the deficit of rare alleles (Williamson-Natesan, 2005). The GW test detects the presence of bottlenecks by determining the ratio between the number of alleles and the allele size range, according to the statistic $M = k/r$, where k is the number of alleles and $r = S_{max} - S_{min} + 1$, where S is allele size.

To detect the genetic composition of populations, we used the Bayesian clustering method implemented in STRUCTURE v.2.3 (Pritchard *et al.*, 2000). This approach assumes that loci are in Hardy–Weinberg equilibrium and that there is linkage equilibrium within populations. Analyses were performed separately for each species using an admixture model with correlated allele frequencies among groups. We ran 500 000 Markov chain Monte Carlo iterations after a burn-in of 100 000 iterations for K values of 1–10, with ten repetitions for each K. The most probable value of K was determined with the method of Evanno *et al.* (2005) implemented in STRUCTURE Harvester (Earl, 2012). The matrices were adapted to the specific format of each of the programs using Transformer-4 (Caujapé-Castells *et al.*, 2013). Microsatellite matrices are available in the DEMIURGE information system (http://www.demiurge-project.org/) with codes D-NMICR-122 (matrix of *C. canariensis*) and D-NMICR-123 (matrix of *C. eminii*). To quantify the amount of genetic variance attributable to geographic and population subdivision, a hierarchical analysis of molecular variance (AMOVA) was performed using the software ARLEQUIN v.3.0 (Excoffier *et al.*, 2005). Exploratory analyses were performed considering, alternatively, the individual populations in palaeoislands and sky islands separately, or grouped into geographic units as identified by STRUCTURE.

TABLE 2. *Prior distribution of demographic and historical parameters used as summary statistics for simulations of population size change in DIYABC (Fig. S1). Times of changes in effective population size are considered from the present (0) back in time:* t_1, t_2, t_3, t_4. *Mean* μ *is the mean mutation rate of microsatellites; Mean* P *is the mean of the geometric distribution parameter. All time events are expressed in numbers of generations. Conditions:*

$N_1 < N_e$, $N_3 < N_e$, $N_4 > N_e$

Parameter	Prior distribution
N_e	Uniform (10–10 000)
t_1	Uniform (1–10 000)
N_1	Uniform (1–100)
t_2	Uniform (1–10 000)
t_3	Uniform (1–10 000)
N_3	Uniform (1–100)
t_4	Uniform (1–10 000)
N_4	Uniform (4000–100 000)
Mean μ	Log-uniform (1.00E–004 to 1.00E–003)
Mean P	Log-uniform (1.00E–001 to 3.00E–001)

N_e, effective population size; N_1 past effective population size in a bottleneck scenario; N_3, past effective population size in an expansion scenario; N_4 past effective population size in a decline scenario.

Demographic analysis

We used approximate Bayesian computation techniques as implemented in the software DIYABC v1.0.4.46beta (Cornuet *et al.*, 2008, 2010) to statistically evaluate alternative scenarios for the demographic history of *C. eminii* and *C. canariensis*. DIYABC uses a coalescent framework to simulate complex evolutionary scenarios without the need to estimate the underlying likelihood function (Cornuet *et al.*, 2008, 2010). Instead, simulated scenarios are summarized using variable estimates (summary statistics), which are then compared with observed values from the genetic data to estimate the posterior probability of the model parameters using Bayesian inference, and to compute measures of bias and precision for each scenario [95 % high posterior density (HPD) credibility intervals]. Due to the strong population structure observed in the two study species (Mairal *et al.*, 2015b, 2017b) and the problems that can arise by assuming panmixia, we conducted separate analyses for each population, and we grouped only those populations that were geographically (<10 km) and genetically very close according to the STRUCTURE results, i.e. Gifta and Dembecha probably form an apomictic population, and were part of the same Afromontane forest patch, nowadays fragmented because of agriculture.

We considered four competing population history scenarios: (1) a null model of constant population size; (2) a bottleneck in the past with subsequent recovery; (3) a demographic expansion; and (4) a recent population decline without recovery (Fig. 1A). We assumed different effective population sizes for each historical scenario (N_e, N_1, N_3, N_4). Table 2 shows the prior distributions of the demographic parameters for the four simulations, each simulating a change in the population size. We assumed that each locus followed a generalized stepwise mutation (GSM) model for microsatellite markers (Estoup *et al.*, 2002), using a log-uniform distribution for the geometric distribution parameter ($P = 10^{-1}$–30^{-1}) and the mutation rate (μ = 10^{-4}–10^{-3}). In each of the tested scenarios, we simulated one million datasets and assumed absence of migration between populations. First, we estimated the posterior probability of each scenario by logistic regression, using the 1 % simulated datasets with summary statistics that were closest to the observed values. Bayesian credibility ranges (95 % HPD) of each simulated scenario were compared (Cornuet *et al.*, 2008). We also estimated the posterior probability of the parameters from 1 % of the best-simulated datasets, using local linear regression and logit transformation of parameters. To make sure that the best scenario was not far off the observed data, we checked the goodness of fit by simulating 1000 data sets from the posterior predictive distribution of the parameter and compared SumStats with the observed data (Cornuet *et al.*, 2010; Budde *et al.*, 2013).

Ecological niche modelling

We modelled the current climatic niche occupied by *C. eminii* and *C. canariensis* using extant occurrence data together with climatic maps. Data points were obtained from published monographs and inventories (Hedberg, 1961), three online databases (www.jardincanario.org/flora-de-gran-canaria, www.gbif.org and www.anthos.es) and data compiled through field trips. In all, we used 108 records for *C. canariensis* and 80 records for *C. eminii* (Supplementary Data Tables S2, S3), covering the entire distributional ranges of the species. Climatic data for current conditions were obtained from WorldClim (www.worlclim.org; Hijmans *et al.*, 2005). For past climate scenarios we downscaled WorldClim data from a 30 arc-second resolution to 1 km resolution and projected current and past climate conditions for several periods after the Last Glacial Maximum, since the end of the deglaciation to the present (18, 10, 5 and 0 kiloyears; for more information see Espíndola *et al.*, 2012). Pleistocene sea-level changes [~18 kiloyears ago the sea levels were 110 m below the actual level (Rijsdijk *et al.*, 2014)] strongly altered the surface area of the Canary Islands, and thus the extent of potentially suitable areas for *C. canariensis*. To account for these changes, we used a bathymetry layer and constructed new rasters using the sea levels modelled by Rijsdijk *et al.* (2014) and projected climate data, including sea levels (−110 m for 18 kiloyears, −46 m for 10 kiloyears, −4.7 m for 5 kiloyears). To model the distributions of the two species, we combined the available occurrences with a set of six bioclimatic variables that could be estimated for all past scenarios: total annual precipitation, maximum and minimum monthly precipitation, annual mean temperature, and maximum and minimum monthly temperature. Pseudoabsences were generated by selecting 5000 random points. We used ensemble modelling (Araújo and New, 2007) to generate our predictions. Three modelling techniques – generalized additive models (GAMs), the general boosting method (GBM) and random forests (RF) – were run and summarized using R packages (R Core Team, 2014): biomod2 (Thuiller *et al*., 2013), foreign, raster (Hijmans and van Etten, 2016), SDMTools (VanDerWal *et al*., 2011), rms (Harrell, 2016), gbm (Ridgeway, 2015), gam (Hastie, 2016), rJava (Urbanek, 2010), dismo (Hijmans *et al*., 2016) and randomForest (Liaw and Wiener, 2002) (references for R packages are given in the Supplementary Data; Mairal *et al*., 2017a). We used repeated split sampling to evaluate the performance of the models, successively splitting the dataset into 70 % for calibration and 30 % for evaluation by measuring the area under the curve (AUC).

RESULTS

Genetic diversity and population structure

Based on the nine SSR loci selected for *C. canariensis* and the six loci for *C. eminii*, we detected 54 and 23 alleles, respectively, in the two species. The greatest number of private alleles in *C. canariensis* was detected in populations from the three palaeo-islands of Tenerife (e.g. 0.37 in Camino Chamuscadas; 0.25 in El Palmar; Table 1), which predate the age of the island in its current configuration. In *C. eminii*, the greatest number of private alleles was detected in populations from the Harar massif (0.33; Table 1). These Canarian and East African populations were inferred as sources of allele migration events in previous studies based on AFLP and chloroplast sequence data (Mairal *et al*., 2015b, 2017b). The average number of alleles per locus ranged from 1.67 to 4.07 in *C. canariensis* and from 1.58 to 2.28 in *C. eminii*. Genetic diversity values, estimated as Nei's heterozygosity (H_e), were relatively low compared with values generally obtained with microsatellite markers in endemic species ($H_e = 0.42$; Nybom, 2004): 0.167–0.41 in *C. eminii* and 0.235–0.489 in *C. canariensis*. Linkage disequilibrium was not significant ($P > 0.05$).

The method of Evanno *et al.* (2005) indicated that the most likely number of genetic groups was $K = 2$ for *C. canariensis* ($\Delta K = 1301$) and $K = 2$ for *C. eminii* ($\Delta K = 108$) (Supplementary Data Tables S4, S5, Fig. S2). A more complex genetic structure was revealed for *C. canariensis* by $K = 4$, delimiting two additional groups [$\Delta K = 60$ (Supplementary Data Table S4)]. In general, populations of *C. canariensis* displayed a clear east/west structure, with a stronger admixture within the islands of Tenerife and Gran Canaria (Fig. 2A). In *C. eminii*, the structure of $K = 2$ supported a north/south division among populations (Fig. 2B).

Hierarchical AMOVA analyses showed the largest proportion of genetic variation among groups, when each group was considered separately (23.22 % in *C. eminii*; 22.21 % in *C. canariensis*; Supplementary Data Table S6). Additionally, the STRUCTURE sublevels among groups ($K = 2$ for *C. eminii* and $K = 2$, $K = 4$ for *C. canariensis*) were consistent with the AMOVA analyses (Supplementary Data Table S6).

Population dynamics

All values for the GW test were M <0.68, the critical value below which populations are assumed to have suffered a recent population decline (with datasets of seven loci or more; Garza and Williamson, 2001). Though we could only reliably score six loci in *C. eminii* (see the Materials and methods section), the results of the demographic analysis (below) and the low values obtained with the GW test (Table 1) support a demographic decline in all populations of *C. canariensis* and *C. eminii*.

In both species, the 'recent population decline' scenario (Supplementary Data Fig. S1) was associated with the highest posterior probabilities and the narrowest 95 % HPD intervals in the DIYABC analysis (Table 3). For some populations, we could not discriminate among alternative hypotheses, as we did not obtain robust statistics for any given scenario [e.g. Gifta, Harenna, Yirga and La Gomera (Tamargada); Table 3, Supplementary Data Table S7]. The current estimated effective population size N_e was estimated between 500 and 1300 across populations, and was about two orders of magnitude smaller than the effective population size before the decline. Population decline was estimated to have started ~2000–5000 generations ago.

Ecological niche modelling

Species distribution models indicated that the inferred potential distribution for the present largely coincided with the extant distribution of both species. The AUC values were generally high (with values of 0.71 and 0.96 for *C. canariensis* and *C. eminii* respectively), suggesting that the models are consistent. Our hindcast climate niche projections showed that the areas with favourable climatic conditions for *C. eminii* and *C. canariensis* experienced a reduction from the start of the deglaciation (around 18 kiloyears ago) to the present (Fig. 3). Present projections show patches of climatic suitability (climatic refugia) that largely coincide with the extant distribution of the species (Fig. 3).

DISCUSSION

Low genetic diversity and strong population structure in Canarina *endemics*

Population genetic theory predicts that species subject to habitat fragmentation, e.g. due to habitat loss, are more susceptible to stochastic demographic fluctuations, such as genetic bottlenecks or declines in effective population size. This is often accompanied by an upward surge of inter-population genetic differentiation, higher inbreeding and loss of genetic variability (Stebbins, 1980; Kruckeberg and Rabinowitz, 1985; Lande, 1993), which can eventually drive a species to extinction. This 'fragmentation

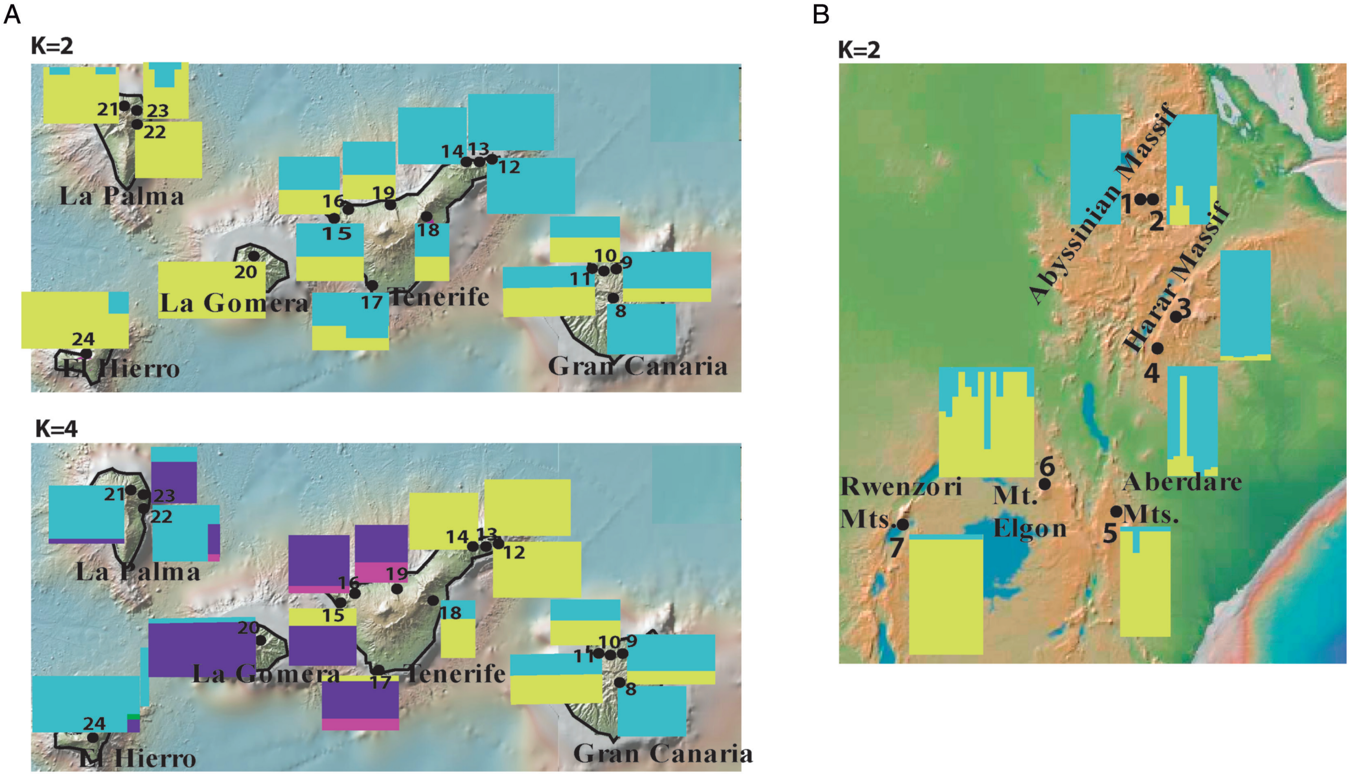


FIG. 2. Results of the genetic structure analyses for (A) *C. canariensis* and (B) *C. eminii*. Histograms show the Bayesian clustering of individuals within populations (STRUCTURE). Colours represent the individual proportion for each of the inferred Bayesian groups. All other conventions as in Fig. 1.

syndrome' (Lande, 1993) seems to be evidenced in *Canarina*. Our ecological niche models indicate that both *C. eminii* and *C. canariensis* have experienced a contraction of their climatic suitability areas since 18 kiloyears ago, which eventually led to the fragmented, patchy distribution observed today (Fig. 3). The low levels of genetic variability detected at the population level (Table 1) are in agreement with those obtained with other types of molecular marker (chloroplast DNA sequences and AFLPs; Mairal *et al.*, 2015b, 2017b). They are also generally below those obtained with microsatellites in endemic species (Nybom, 2004), and specifically below those reported in other taxa distributed in the study areas (Table 4). Note, however, that comparing genetic statistics across distantly related phylogenetic groups is problematic because their values can be influenced by taxon-specific biological differences (Gitzendanner and Soltis, 2000).

The loss of genetic diversity is also accompanied in *Canarina* by a large among-population differentiation (see AMOVA in Supplementary Data Table S6) and great genetic exclusivity in some populations, especially those that may have acted as refugia or genetic reservoirs in the past (e.g. the palaeo-islands of Tenerife or Mount Elgon in Kenya), though this diversity, detected also by AFLPs and plastid sequences, seems to be of older origin than the one reported here (Mairal *et al.*, 2015b, 2017a).

In *C. canariensis*, our SSR analyses detected a strong population structure akin to the one revealed by AFLPs and DNA sequences: an east/west genetic split, with the strongest difference found within Tenerife (Fig. 2A; Mairal *et al.*, 2015b, 2017a). The only difference is that SSRs detected higher levels of genetic admixture in the population of the paleo-island of Adeje than the other markers (Fig. 2A; this had been predicted in Mairal *et al.*, 2015b). In *C. eminii*, SSRs detect a north/south genetic diversity structure, which was not detected by AFLPs or sequences (Mairal *et al.*, 2017b). This might reflect more recent population dynamics in agreement with similar contact patterns found in other Afromontane plants (Assefa *et al.*, 2007; Masao *et al.*, 2013). Alternatively, it could be an artefact of the lack of specificity in our SSR markers, which were designed for *C. canariensis*. Additionally, we found genetic admixture in Mount Elgon (Fig. 2B), which agrees well with its purported role as a crossroad among populations at both sides of the Rift Valley (Mairal *et al.*, 2017b), and with the descent of vegetation belts during the Last Glacial Maximum (Chala *et al.*, 2017).

Historical population declines in widespread rare species

The time lag between range reduction and the collapse of populations is known as 'extinction debt' (Tilman *et al.*, 1994). Initially, if studies are conducted shortly after habitat loss has occurred, the evolutionary cost of the extinction debt is not obvious and thus difficult to detect (Triantis *et al.*, 2010). Several generations are needed to detect the total impact of habitat loss and fragmentation (Tilman *et al.*, 1994; Tilman, 1999). This seems to be the case in *Canarina*: although this endemic genus of evergreen subtropical forests is still widely distributed in regions of East Africa and Macaronesia, it also shows a worrying population decline scenario. Indeed, demographic analyses based on SSRs (Tables 1 and 3) and ecological niche model

TABLE 3. *Results of DIYABC analyses. In each population the best supported scenario is indicated according to its posterior probability [PP; mean and 95 % Bayesian highest posterior density interval (HPD)], obtained by logistic regression from 1 % of the simulations that best approximate the observed values. NA (not available) is shown for populations for which it was not possible to discriminate a statistically robust scenario*

Population	Population	No. of individuals	Best supported scenario	PP (HPD)	N_e	N_4	t_4	N_eμmic
Canarina eminii, continental islands								
1. Gifta and Dembecha	Abyssinian massif	16	NA	–	–	–	–	–
3. Yirga	Harar massif	13	NA	–	–	–	–	–
4. Harenna	Harar massif	9	NA	–	–	–	–	–
5. Aberdare	Aberdare Mountains	8	Decline	0.7511 (0.6710, 0.8312)	5.17E+02	6.33E+E04	2.03E+03	9.73E–02
6. Elgon	Mount Elgon	8	Decline	0.6912 (0.6544, 0.7280)	5.35E+02	5.94E+E04	2.89E+03	1.39E–01
7. Rwenzori	Rwenzori Mountains	10	Decline	0.6986 (0.6631, 0.7342)	1.03E+03	5.84E+E04	3.56E+03	1.92E–01
Canarina canariensis, oceanic islands								
8. Cueva Corcho	Gran Canaria	8	Decline	0.9293 (0.9066, 0.9519)	5.50E+02	6.59E+E04	1.72E+03	9.91E–02
9. Tilos de Moya	Gran Canaria	12	Decline	0.9400 (0.7882, 1.0000)	6.60E+02	6.33E+E04	2.01E+03	1.22E–01
10. El Sao	Gran Canaria	10	Decline	0.3902 (0.3325, 0.4478)	6.83E+02	5.11E+E04	4.51E+03	1.05E–01
11. Andenes de Guayedra	Gran Canaria	12	Decline	0.6930 (0.6497, 0.7363)	5.00E+02	5.77E+E04	2.90E+03	9.07E–02
12. Camino Chamuscadas	Tenerife (Anaga)	10	Decline	0.5972 (0.5646, 0.6298)	1.25E+03	5.06E+E04	4.96E+03	4.01E–01
13. El Bailadero	Tenerife (Anaga)	11	Decline	0.9265 (0.7775, 1.0000)	1.03E+03	6.07E+E04	4.02E+03	2.10E–01
14. Tope del Carnero	Tenerife (Anaga)	7	Decline	0.6000 (0.1706, 1.0000)	1.40E+03	7.48E+E03	3.51E+03	2.78E–01
15. Teno Alto	Tenerife (Teno)	8	Decline	0.9523 (0.9360, 0.9686)	1.06E+03	6.90E+E04	2.87E+03	2.09E–01
16. El Palmar	Tenerife (Teno)	8	Decline	0.9511 (0.9354, 0.9668)	1.34E+03	6.25E+E04	3.46E+03	2.61E–01
17. Barranco del Infierno	Tenerife (Adeje)	11	Decline	0.9940 (0.9880, 1.0000)	9.31E+02	6.89E+E04	1.87E+03	1.73E–01
18. Barranco de Badajoz	Tenerife	5	Decline	0.8499 (0.8096, 0.8902)	7.41E+02	6.65E+E04	3.39E+03	1.56E–01
19. Barranco Ruiz	Tenerife	6	Decline	0, 8775 (0.8423, 0.9127)	8.70E+02	6.18E+E04	2.35E+03	1.45E–01
20. Tamargada	La Gomera	13	NA	–	–	–	–	–
21. Los Tilos	La Palma	10	Decline	0.8024 (0.7579, 0.8469)	9.00E+02	5.07E+E04	2.83E+03	1.64E–01
22. Barranco del Agua	La Palma	10	Decline	0.7253 (0.6768, 0.7738)	7.70E+02	5.71E+E04	2.37E+03	1.43E–01
24. El Hierro	El Hierro	12	Decline	0.7498 (0.6942, 0.8054)	9.18E+02	5.43E+E04	2.44E+03	1.51E–01

N_e, current effective population size; N_4, effective population size at the beginning of the decline; t_4, start time of decline, measured in number of generations; N_eμmic, current effective population size × micros mutation rate.

projections (Fig. 3) suggest that both *C. eminii* and *C. canariensis* underwent a strong and recent population decline.

When did this decline start and what were its causes? Our previous studies showed that the current disjunct distribution of genus *Canarina* across the Sahara Desert can be explained by the climatic aridification during the Late Miocene/Pliocene and the resulting contraction in the distribution ranges of subtropical forests in Africa and Macaronesia (Mairal *et al.*, 2015a, b, 2017a, b; Pokorny *et al.*, 2015). Although this hypothesis may explain the current continental disjunction, the population decline detected here is probably more recent. The DIYABC analysis detected a population decline ~2000–5000 generations ago, which could not be recovered with coalescent-based Bayesian skyline plots (Drummond *et al.*, 2005; Mairal *et al.*, 2015b). This is in agreement with observations that SSR markers detect more recent signals of genetic diversity than chloroplast DNA and AFLP markers (Chakraborty *et al.*, 2014). Since the generation time of *Canarina* is not known precisely due to its tuberous habit (it can sometimes behave as a biennial or even show longer generation times), it is difficult to assign an absolute temporal framework for this decline; also, the time scale here differed slightly depending on the population. However, the temporal range of the decline detected here was consistent within the past 10 000 years (if we consider the plants as biennial) or 50 000 years (if we consider generation times of 10 years), i.e. it probably took place during the last glacial cycle (Pleistocene–Holocene).

This decline coincides with palaeoclimate projections from our ecological niche models, which support a contraction of the distribution ranges of both species from 18 kiloyears ago to the present (Fig. 3). Caution should be exercised in the interpretation of these models due to the low-resolution climate datasets available, which could underestimate potential microrefugia (Gavin *et al.*, 2014). However, though geographic locations may be biased, the models are highly congruent with the extant distribution of the species and furnish evidence of a population decline. Furthermore, a recent study (Chala *et al.*, 2017) supports our models by projecting the descent of vegetation belts during the Last Glacial Maximum.

Interestingly, DIYABC detected a less pronounced population decline in the populations with the highest levels of genetic diversity (Table 1 and N_4 versus N_e in Table 3). These populations are located in areas that served as refugia during climatic oscillations or tectonic and volcanic events, such as Mount Elgon in Africa or the palaeo-islands of Tenerife, thereby supporting their role as reservoirs of genetic diversity against extinction, both during the Pleistocene and in more recent times (Mairal *et al.*, 2015b, 2017b; this study).

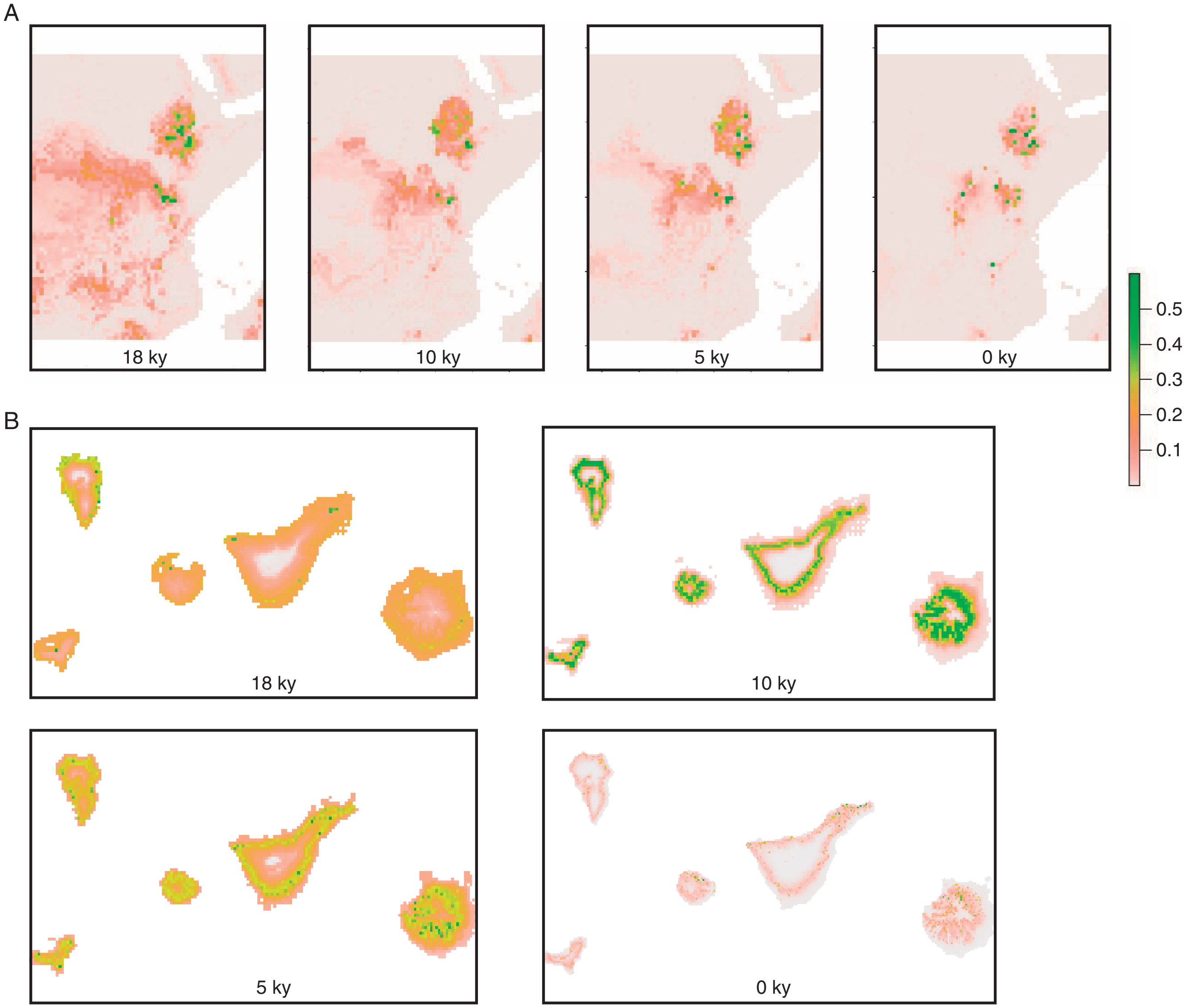


FIG. 3. Geographic projections of the climatic niche model of (A) *Canarina eminii* and (B) *Canarina canariensis* over four time periods: 18, 10, 5 and 0 kiloyears (ky) (modelling sea-level changes). Colour scale indicates climatic suitability values; brown colour indicates low values and dark green indicates high values.

By contrast, a much sharper decline is observed in the most fragmented laurel forest populations in the Canaries, i.e. Barranco de Badajoz and Barranco Ruiz in Tenerife (Table 3). In *C. eminii*, the sky islands located at lower latitudes (Rwenzori Mountains, Mount Elgon, Aberdare Mountains) harbour greater genetic diversity than populations in the massif of Abyssinia, which were probably more strongly affected by glaciations and human activities (e.g. Gifta and Dembecha in Table 1; cf. Mairal *et al.*, 2017b). Loss of ecological interactions in these areas, such as deficiency of pollinators, may also have had a direct impact on genetic diversity in rare species (Lavergne *et al.* 2004).

Conservation biology of Canarina

Our ecological niche models may render an approximation of the effect of climate change on habitat loss and the associated loss of genetic diversity during the last glacial/interglacial cycles. Nowadays, however, human activities constitute an additional threat to the maintenance of genetic diversity in *Canarina*. Afro-Macaronesian forests have been logged in the past centuries for timber, agriculture and human settlements, bringing them to the verge of extinction, especially in topographically less complex areas, which are more easily exploited.

At present, the Afromontane forests have disappeared almost entirely from the Ethiopian highlands (estimated loss of up to of 35 % of originally forested areas) (EFAP, 1994; Kebede *et al.*, 2007). This situation is aggravated by the decrease in soil quality due to changes in water resources (Kloos and Legesse, 2010), which has caused a dramatic decline in these forests, and even their disappearance in large areas (Reusing, 2000; FAO, 2001).

In the Canary Islands, the area of laurel forests has recently declined to 12.5 % of their potential distribution

TABLE 4. *Average values of microsatellite diversity statistics for several species studied in the Canary Islands and Afromontane forests*

Species	*A*	H_e	Reference
Canary Islands			
Canarina canariensis	2.5	0.367	–
Bencomia exstipulata	3.50	0.440	González-Pérez *et al.*, 2009a
Bencomia caudata	5.13	0.620	González-Pérez *et al.*, 2009a
Myrica rivas-martinezii	6.50	0.560	González-Pérez *et al.*, 2009b
Neochamalea pulverulenta	7.85	0.425–0.812	Rigueiro *et al.*, 2009
Myrica faya	9.30	0.570	González-Pérez *et al.*, 2009b
Olea europaea subsp. *guanchica*	14.00	0.710	García-Verdugo *et al.*, 2010
Ilex perado subsp. *lopezlilloi*	2.13	0.480	Sosa *et al*., 2010a
Ilex perado subsp. *platyphylla*	5.37	0.600	Sosa *et al*., 2010a
Ilex canariensis	2.38	0.420	Sosa *et al*., 2010a
Sambucus palmensis	6.80	0.499	Sosa *et al*., 2010b
Silene nocteolens	15.83	0.780	Sosa *et al*., 2010a
Pinus canariensis	22.80	0.771	López de Heredia *et al.*, 2010
Parolinia ornata	4.38	0.515	González-Pérez and Caujapé-Castells, 2014
Ruta oreojasme	7.62	0.687	Meloni *et al.*, 2015
Phoenix canariensis	4.41	0.476	Saro *et al.*, 2015
Micromeria rivas-martinezii	9.80	0.500	Puppo *et al.*, 2016
Micromeria glomerata	4.27	0.310	Puppo *et al.*, 2016
Afromontane forests			
Canarina eminii	1.93	0.300	–
Coffea arabica	2–3.20	0.050–0.242	Silvestrini *et al.*, 2007
Acacia senegal	6.90	0.667	Omondi *et al.*, 2010
Senegalia mellifera	7.00	0.620	Ruiz Guajardo *et al.*, 2010
Prunus africana	2.0 – 4.00	0.105–0.728	Kadu *et al.*, 2011

A, allelic richness.

(Fernández-Palacios *et al.*, 2011); this includes the laurel forest contractions of Gran Canaria and Tenerife to 1 % and 10 %, respectively, of their original distributions (Santos Guerra, 1990; Del Arco *et al*., 2010; Fernández-Palacios *et al.*, 2011). The areas of adjacent suitable areas, such as thermophilous woodlands, have also declined considerably (de Nascimento *et al.*, 2015).

One of the biggest threats to rare widespread species is the sensitivity of their populations to demographic stochasticity (Kruckeberg and Rabinowitz, 1985; Barrett *et al.*, 1991). This could be a critical factor for the survival of *Canarina* in African and Macaronesian evergreen forest habitats, given the detected current decline in size of many populations (Table 3). Furthermore, the generally low levels of genetic diversity detected here make these taxa more susceptible to the effects of inbreeding (especially in small populations). Intraspecific genetic variation is the most fundamental level of biodiversity, providing the basis for evolutionary change, and is crucial for maintaining the capacity of species to adapt to new environmental conditions (Frankham *et al.*, 2002; Yannic *et al.*, 2013).

Although *C. canariensis* is locally abundant in some areas (e.g. Anaga, Teno and Tilos de Moya), greater efforts should be made to protect populations that contain exclusive genetic diversity. This is the case for populations in the three palaeo-islands of Tenerife (which also harbour pristine areas of laurisilva) and of Mount Elgon in the case of *C. eminii*. These areas act not only as reservoirs of unique genetic diversity, but also as cradles and sources of new allele migration events to nearby areas (Mairal *et al.*, 2015b, 2017b; Caujapé-Castells *et al.*, 2017). Though the results reported here refer only to two rare species (*C. eminii* and *C. canariensis*), similar patterns are expected (and have been observed) in other rare species, e.g. the Afroalpine plant *Lobelia rhynchopetalum* (Chala *et al.*, 2016).

At present, a large area covered by the endemic laurel forest (laurisilva) across the Canary Archipelago is under legal protection (Fernández-Palacios *et al.*, 2011), but this is not the case for the isolated patches of Afromontane forests in East Africa. Though some areas are preserved as National Parks (e.g. Harenna Forest, Rwenzori Mountains and Mount Elgon), conservation measures are not strictly enforced. Given the political and difficult economic situation of these countries, which makes it even more challenging to extend protection areas to preserve new populations, one crucial action in this current context would be *ex situ* conservation, via seed banks. Priority should be given to the conservation of the different geographically structured areas, covering in each of them those populations exhibiting the highest levels of ancient and endemic phylogenetic diversity. Genetic evidence gathered for *Canarina* (DNA, AFLPs and SSRs) further suggests that certain populations should be established as shrine areas in order to conserve the greater genetic variation *ex situ*: in the case of *C. eminii*, one population at each side of the Rift Valley and the population of Mount Elgon; for *C. canariensis*, one population in each palaeo-island in Tenerife (home to the best-preserved areas of laurisilva in Tenerife). *Ex situ* conservation plans should collect seeds covering these priority areas and avoid gene flow among them in order to favour potential speciation in the future. Thus, while the different areas should be managed separately (avoiding mixed reinforcements, i.e. without introducing propagules from different source areas), management within areas should stress the preservation of several populations, in order to safeguard a significant proportion of genetic variability of the species.

CONCLUSIONS

Genetic conservation studies often focus on rare species with a reduced distribution area ('narrow endemics'; Kruckeberg and Rabinowitz, 1985; Oostermeijer *et al.*, 2003; Fernández-Mazuecos *et al.*, 2014). However, geographically widespread species can also be considered rare (relict) and subject to the same status of threat, especially if habitat loss has led to isolation of populations and small population size ('large size and narrow distribution'; Kruckeberg and Rabinowitz, 1985). Here we show that this is the case for *C. eminii* and *C. canariensis*, two endemic bellflower species associated with the last remnants of subtropical forest in Africa and Macaronesia. Although both are locally widespread (*C. canariensis* is present in most of the Canarian islands and *C. eminii* occurs over a large part of East Africa), they feature the genetic signals of recent population decline due to habitat loss. The latter seems to be an effect of both historical climate change and rapid, human-induced, forest fragmentation in recent times. Especially for the Afromontane forests, loss of habitat due to increasing aridity and the alarming rates of forecast urban growth make these threatened ecosystems areas of priority interest for the conservation community (Seto *et al.*, 2012).

SUPPLEMENTARY DATA

Supplementary data are available online at https://academic.oup.com/aob and consist of the following. Table S1: primers used for PCR amplification for each locus. Table S2: geographic coordinates used in the ecological niche modelling of *Canarina canariensis*. Table S3: geographic coordinates used in the ecological niche modelling of *Canarina eminii*. Table S4: estimated probability of the likelihood function according to the Evanno method for the STRUCTURE analyses of *C. canariensis*. Table S5: estimated probability of the likelihood function according to the Evanno method for the STRUCTURE analyses of *C. eminii*. Table S6: hierarchical analysis of molecular variance (AMOVA). Table S7: probabilities for each scenario performed with DIYABC. Confidence intervals are shown in square brackets. Figure S1: graphical representation of the four scenarios simulating changes in population size with DIYABC. Supplementary references. Figure S2: Estimated probability of the likelihood function according to the Evanno method.

ACKNOWLEDGEMENTS

Field work could not have been conducted without the cooperation of the staff at the Jardín Botánico Canario 'Viera y Clavijo' – Unidad Asociada, CSIC (Cabildo de Gran Canaria). We thank the Cabildo de Tenerife, Jacinto Leralta, Ángel Fernández (La Gomera), Félix Manuel Medina from the Cabildo of La Palma and the Cabildo of El Hierro for help with accommodation and sampling logistics during field expeditions. We thank Juan Jose Aldasoro, Mª Luisa Alarcón and Andrea Sánchez-Meseguer for providing African samples, and Alicia Agudo, Carlos García-Verdugo, CNAE Members, Paloma Torroba and Victoria Culshaw for help during different stages of this work. This work was supported by a PhD research grant (BES-2010–037261) to M.M., including short research stays at JBCVC-CSIC (J.C.C.), University of Lausanne (N.A.) and University of Fribourg (L.P. and M.H.) funded by MINECO (grants CGL2009-1332-C03-01 (MICINN), CGL2012-40129-C02-01 and CGL2015-67849-P) (MINECO/FEDER) to I.S.; grant CGL2012-40129-C02-02 (MINECO) and a Marie Curie Intra-European Fellowship (PIEF-GA-2012-329088) to M.H. and project grant DEMIURGO (MAC/1/C20) to J.C.C.

LITERATURE CITED

Alemayehu T. 2006. *Diversity and ecology of vascular epiphytes in Harenna Afromontane forest, Bale, Ethiopia*. PhD thesis, Addis Ababa University.

Araújo MB, **New M. 2007**. Ensemble forecasting of species distributions. *Trends in Ecology and Evolution* **22**: 42–47.

del Arco Aguilar MJ, **González-González R**, **Garzón-Machado V**, **Pizarro-Hernández B. 2010**. Actual and potential natural vegetation on the Canary Islands and its conservation status. *Biodiversity and Conservation* **19**: 3089–3140.

Assefa A, **Ehrich D**, **Taberlet P**, **Nemomissa S**, **Brochmann C. 2007**. Pleistocene colonization of afro-alpine 'sky islands' by the arctic-alpine *Arabis alpina*. *Heredity* **99**: 1–10.

Axelrod DI, **Raven PH. 1978**. Late Cretaceous and Tertiary vegetation history of Africa. In: Werger MJA, ed. *Biogeography and ecology of Southern Africa*. Dordrecht: Springer, 77–130.

Barrett SC, **Kohn JR**, **Falk DA**, **Holsinger KE. 1991**. Genetic and evolutionary consequences of small population size in plants: implications for conservation. *Genetics and Conservation of Rare Plants* **29**: 3–30.

Beaumont MA. 2010. Approximate Bayesian computation in evolution and ecology. *Annual Review of Ecology, Evolution, and Systematics* **41**: 379–406.

Bonnefille R. 2011. Rainforest responses to past climate changes in tropical Africa. In: Bush M, Flenley J, Gosling W, eds. *Tropical rainforest responses to climate change*, 2nd edn. Berlin: Springer, 125–184.

Bramwell D. 1985. Contribución a la biogeografía de las Islas Canarias. *Botánica Macaronésica* **14**: 3–34.

Bramwell D, **Bramwell ZI. 2001**. *Flores silvestres de las Islas Canarias*. Madrid: Rueda.

Boutin-Ganache I, **Raposo M**, **Raymond M**, **Deschepper CF. 2001**. M13-tailed primers improve the readability and usability of microsatellite analyses performed with two different allele-sizing methods. *BioTechniques* **31**: 24–28.

Budde KB, **González-Martínez SC**, **Hardy OJ**, **Heuertz M. 2013**. The ancient tropical rainforest tree *Symphonia globulifera* L. f. (Clusiaceae) was not restricted to postulated Pleistocene refugia in Atlantic Equatorial Africa. *Heredity* **111**: 66–76.

Caujapé-Castells J. 2010. General GST and θ inflation due to biased intra-population sampling, and its consequences for the conservation of the Canarian flora. *Conservation Genetics* **11**: 709–720.

Caujapé-Castells J, **Sabbagh I**, **Castellano JJ**, ***et al.* 2013**. Transformer-4 version 2.0.1, a free multi-platform software to quickly reformat genotype matrices of any marker type, and archive them in the Demiurge information system. *Molecular Ecology Resources* **13**: 484–493.

Caujapé-Castells J, **García-Verdugo C**, **Marrero-Rodríguez A**, **Fernández-Palacios JM**, **Crawford DJ**, **Mort ME. 2017**. Island ontogenies, syngameons, and the origins of genetic diversity in the Canarian flora. *Perspectives in Plant Evolution, Ecology and Systematics* **27**: 9–22.

Chakraborty D, **Sinha A**, **Ramakrishnan U. 2014**. Mixed fortunes: ancient expansion and recent decline in population size of a subtropical montane primate, the Arunachal macaque *Macaca munzala*. *PLoS ONE* doi: 10.1371/journal.pone.0097061.

Chala D, **Brochmann C**, **Psomas A**, ***et al.* 2016**. Goodbye to tropical alpine plant giants under warmer climates? Loss of range and genetic diversity in *Lobelia rhynchopetalum*. *Ecology and Evolution* **6**: 8931–8941.

Chala D, **Zimmermann NE**, **Brochmann C**, **Bakkestuen V. 2017**. Migration corridors for alpine plants among the 'sky islands' of eastern Africa: do they, or did they exist? *Alpine Botany* **127**: 133–144.

Chybicki IJ, **Burczyk J. 2009**. Simultaneous estimation of null alleles and inbreeding coefficients. *Journal of Heredity* **100**: 106–113.

Cornuet JM, **Santos F**, **Beaumont MA**, ***et al.* 2008**. Inferring population history with DIYABC: a user-friendly approach to approximate Bayesian computation. *Bioinformatics* **24**: 2713–2719.

Cornuet JM, **Ravigné V**, **Estoup A. 2010**. Inference on population history and model checking using DNA sequence and microsatellite data with the software DIYABC (v1.0). *BMC Bioinformatics* **11**: 401.

Cronk QCB. 1992. Relict floras of Atlantic islands: patterns assessed. *Biological Journal of the Linnean Society* **46**: 91–103.

Dewoody J, **Nason JD**, **Hipkins VD. 2006**. Mitigating scoring errors in microsatellite data from wild populations. *Molecular Ecology Notes* **6**: 951–957.

Drummond AJ, **Rambaut A**, **Shapiro B**, **Pybus OG. 2005**. Bayesian coalescent inference of past population dynamics from molecular sequences. *Molecular Biology and Evolution* **22**: 1185–1192.

Earl DA. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* **4**: 359–361.

EFAP. **1994**. *Ethiopia Forestry Action Program*. Final report. Addis Ababa: Ministry of Natural Resources Development and Environmental Protection.

Espíndola A, **Pellissier L**, **Maiorano L**, **Hordijk W**, **Guisan A**, **Alvarez N. 2012**. Predicting present and future intra-specific genetic structure through niche hindcasting across 24 millennia. *Ecology Letters* **15**: 649–657.

Estoup A, **Jarne P**, **Cornuet JM. 2002**. Homoplasy and mutation model at microsatellite loci and their consequences for population genetics analysis. *Molecular Ecology* **11**: 1591–1604.

Evanno G, **Regnaut S**, **Goudet J. 2005**. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* **14**: 2611–2620.

Excoffier L, **Laval G**, **Schneider S. 2005**. Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics* **1**: 117693430500100003.

Fahrig L. 2003. Effects of habitat fragmentation on biodiversity. *Annual Review of Ecology, Evolution, and Systematics* **34**: 487–515.

Faith DP. 1992. Conservation evaluation and phylogenetic diversity. *Biological Conservation* **61**: 1–10.

FAO. **2001**. *Forest resource assessment 2000. Forestry Paper 140*. Rome: Food and Agricultural Organization of the United Nations.

Fernández-Mazuecos M, **Jiménez-Mejías P**, **Rotllan-Puig X**, **Vargas P. 2014**. Narrow endemics to Mediterranean islands: moderate genetic diversity but narrow climatic niche of the ancient, critically endangered *Naufraga* (Apiaceae). *Perspectives in Plant Ecology, Evolution and Systematics* **16**: 190–202.

Fernández-Palacios JM, **de Nascimento L**, **Otto R**, ***et al*****. 2011**. A reconstruction of Palaeo-Macaronesia, with particular reference to the long-term biogeography of the Atlantic island laurel forests. *Journal of Biogeography* **38**: 226–246.

Fjeldså J, **Lovett JC. 1997**. Geographical patterns of old and young species in African forest biota: the significance of specific montane areas as evolutionary centres. *Biodiversity and Conservation* **6**: 325–346.

Frankham R, **Briscoe D**, **McInnes K**, **Ballou J. 2002**. *Introduction to conservation genetics*. Cambridge, UK: Cambridge University Press.

García-Verdugo C, **Forrest AD**, **Fay MF**, **Vargas P. 2010**. The relevance of gene flow in metapopulation dynamics of an oceanic island endemic, *Olea europaea* subsp. *guanchica*. *Evolution* **64**: 3525–3536.

García-Verdugo C, **Sajeva M**, **La Mantia T**, **Harrouni C**, **Msanda F**, **Caujapé-Castells J. 2015**. Do island plant populations really have lower genetic variation than mainland populations? Effects of selection and distribution range on genetic diversity estimates. *Molecular Ecology* **24**: 726–741.

Garza JC, **Williamson EG. 2001**. Detection of reduction in population size using data from microsatellite loci. *Molecular Ecology* **10**: 305–318.

Gavin DG, **Fitzpatrick MC**, **Gugger PF**, ***et al.* 2014**. Climate refugia: joint inference from fossil records, species distribution models and phylogeography. *New Phytologist* **204**: 37–54.

Gitzendanner MA, **Soltis PS. 2000**. Patterns of genetic variation in rare and widespread plant congeners. *American Journal of Botany* **87**: 783–792.

González-Pérez MA, **Lledo MD**, **Lexer C**, ***et al.* 2009*a***. Genetic diversity and differentiation in natural and reintroduced populations of *Bencomia exstipulata* and comparisons with *B. caudata* (Rosaceae) in the Canary Islands: an analysis using microsatellites. *Botanical Journal of the Linnean Society* **160**: 429–441.

González-Pérez MA, **Sosa PA**, **Rivero E**, **González-González EA**, **Naranjo A. 2009*b***. Molecular markers reveal no genetic differentiation between *Myrica rivas-martinezii* and *M. faya* (Myricaceae). *Annals of Botany* **103**: 79–86.

González-Pérez MÁ, **Caujapé-Castells J. 2014**. Development and characterization of nuclear microsatellite markers for *Parolinia ornata* Webb (Brassicaceae), and cross-species amplification in all species described in the Canarian endemic genus *Parolinia*. *Conservation Genetics Resources* **6**: 705–706

Grandcolas P, **Nattier R**, **Trewick S. 2014**. Relict species: a relict concept? *Trends in Ecology & Evolution* **29**: 655–663.

Habel JC, **Assmann T, eds. 2009**. *Relict species: phylogeography and conservation biology*. Heidelberg, Dordrecht, London, New York: Springer Science & Business Media.

Hedberg O. 1961. Monograph of the genus *Canarina L.* (Campanulaceae). *Svensk Botanisk Tidskrift* **55**: 17–62.

Heald WF. 1951. Sky islands of Arizona. *Natural History* **60**: 95–96.

Hewitt GM. 2004. Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* **359**: 183–195.

Hijmans RJ, **Cameron SE**, **Parra JL**, **Jones PG**, **Jarvis A. 2005**. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* **25**: 1965–1978.

Juan C, **Emerson BC**, **Oromí P**, **Hewitt GM. 2000**. Colonization and diversification: towards a phylogeographic synthesis for the Canary Islands. *Trends in Ecology & Evolution* **15**: 104–109.

Kadu CAC, **Schueler S**, **Konrad H**, ***et al.* 2011**. Phylogeography of the Afromontane *Prunus africana* reveals a former migration corridor between East and West African highlands. *Molecular Ecology* **20**: 165–178.

Kalinowski ST. 2005. HP-Rare: a computer program for performing rarefaction on measures of allelic diversity. *Molecular Ecology Notes* **5**: 187–189.

Kebede M, **Ehrich D**, **Taberlet P**, **Nemomissa S**, **Brochmann C. 2007**. Phylogeography and conservation genetics of a giant lobelia (*Lobelia giberroa*) in Ethiopian and Tropical East African mountains. *Molecular Ecology* **16**: 1233–1243.

Kloos H, Legesse **W, eds. 2010**. *Water resources management in Ethiopia: implications for the Nile basin*. Amherst, NY: Cambria Press.

Kondraskov P, **Schuetz N**, **Schuessler C**, ***et al.* 2015**. Biogeography of Mediterranean hotspot biodiversity: reevaluating the 'Tertiary' relict hypothesis of Macaronesian laurel forests. *PLoS ONE* doi: 10.1371/journal.pone.0132091

Kruckeberg AR, **Rabinowitz D. 1985**. Biological aspects of endemism in higher plants. *Annual Review of Ecology and Systematics* **16**: 447–479.

Lande R. 1993. Risks of population extinction from demographic and environmental stochasticity and random catastrophes. *American Naturalist* **142**: 911–927.

Lavergne S, **Thompson JD**, **Garnier E**, **Debussche M. 2004**. The biology and ecology of narrow endemic and widespread plants: a comparative study of trait variation in 20 congeneric pairs. *Oikos* **107**: 505–518.

López de Heredia U, **Venturas M**, **López RA**, **Gil L. 2010**. High biogeographical and evolutionary value of Canary Island pine populations out of the elevational pine belt: the case of a relict coastal population. *Journal of Biogeography* **37**: 2371–2383.

Mairal M, **Pokorny L**, **Aldasoro JJ**, **Alarcón M**, **Sanmartín I. 2015*a***. Ancient vicariance and climate-driven extinction explain continental-wide disjunctions in Africa: the case of the Rand Flora genus *Canarina* (Campanulaceae). *Molecular Ecology* **24**: 1335–1354.

Mairal M, **Sanmartín I**, **Aldasoro JJ**, **Culshaw V**, **Manolopoulou I**, **Alarcón M. 2015*b***. Paleo-islands as refugia and sources of genetic diversity within volcanic archipelagos: the case of the widespread endemic *Canarina canariensis* (Campanulaceae). *Molecular Ecology* **24**: 3944–3963.

Mairal M, **Sanmartín I**, **Pellissier L. 2017*a***. Lineage-specific climatic niche drives the tempo of vicariance in the Rand Flora. *Journal of Biogeography* **44**: 911–923.

Mairal M, **Sanmartín I**, **Herrero A**, ***et al.* 2017*b***. Geographic barriers and Pleistocene climate change shaped patterns of genetic variation in the Eastern Afromontane biodiversity hotspot. *Scientific Reports* **7**: e45749.

Maley J. 2000. Last Glacial Maximum lacustrine and fluviatile formations in the Tibesti and other Saharan mountains, and large-scale climatic teleconnections linked to the activity of the Subtropical Jet Stream. *Global and Planetary Change* **26**: 121–136.

Masao CA, **Gizaw A**, **Piñeiro R**, ***et al.* 2013**. Phylogeographic history and taxonomy of some afro-alpine grasses assessed based on AFLPs and morphometry: *Deschampsia cespitosa, D. angusta* and *Koeleria capensis*. *Alpine Botany* **123**: 107–122.

McCormack JE, **Huang H**, **Knowles LL. 2009**. Sky islands. In: Gillespie RG, Clague DA, eds. *Encyclopedia of islands*. Berkeley: University of California Press, 841–843.

Meloni M, **Reid A**, **Caujapé-Castells J**, **Soto M**, **Fernández-Palacios JM**, **Conti E. 2015**. High genetic diversity and population structure in the endangered Canarian endemic *Ruta oreojasme* (Rutaceae). *Genetica* **143**: 571–580.

Mittermeier RA, **Robles Gil P**, **Hoffmann M**, ***et al.* 2004**. *Hotspots revisited*. Mexico City: CEMEX.

Myers N, **Mittermeier RA**, **Mittermeier CG**, **Da Fonseca GA**, **Kent J. 2000**. Biodiversity hotspots for conservation priorities. *Nature* **403**: 853–858.

de Nascimento L, **Nogué S**, **Criado C**, ***et al.* 2015**. Reconstructing Holocene vegetation on the island of Gran Canaria before and after human colonization. *Holocene* **26**: 113–125.

Nybom H. 2004. Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. *Molecular Ecology* **13**: 1143–1155.

Omondi SF, **Kireger E**, **Dangasuk OG**, ***et al.* 2010**. Genetic diversity and population structure of *Acacia senegal* (L) Willd. in Kenya. *Tropical Plant Biology* **3**: 59–70.

Van Oosterhout C, **Hutchinson WF**, **Wills DP**, **Shipley P. 2004**. MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* **4**: 535–538.

Oostermeijer JGB, **Luijten SH**, **Den Nijs JCM. 2003**. Integrating demographic and genetic approaches in plant conservation. *Biological Conservation* **113**: 389–398.

Osborne AH, **Vance D**, **Rohling EJ**, **Barton N**, **Rogerson M**, **Fello N. 2008**. A humid corridor across the Sahara for the migration of early modern humans out of Africa 120,000 years ago. *Proceedings of the National Academy of Sciences of the USA* **105**: 16444–16447.

Pérez de Paz J, **Caujapé-Castells J. 2013**. A review of the allozyme data set for the Canarian endemic flora: causes of the high genetic diversity levels and implications for conservation. *Annals of Botany* **111**: 1059–1073.

Petit RJ, **Aguinagalde I**, **de Beaulieu JL**, ***et al.* 2003**. Glacial refugia: hotspots but not melting pots of genetic diversity. *Science* **300**: 1563–1565.

Pokorny L, **Riina R**, **Mairal M**, ***et al.* 2015**. Living on the edge: timing of Rand Flora disjunctions congruent with ongoing aridification in Africa. *Frontiers in Genetics* **6**: 1–15.

Pritchard JK, **Stephens M**, **Donnelly P. 2000**. Inference of population structure using multilocus genotype data. *Genetics* **155**: 945–959.

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

Rabinowitz D. 1981. Seven forms of rarity. In: Synge H, ed. *The biological aspects of rare plant conservation*. Chichester: Wiley, 205–217.

Reusing M. 2000. Change detection of natural high forests in Ethiopia using remote sensing and GIS techniques. *International Archives of Photogrammetry and Remote Sensing* **33**: 1253–1258.

Rigueiro C, **Arroyo JM**, **Valido A**, **Jordano P. 2009**. Isolation and characterization of 13 microsatellite loci for *Neochamaelea pulverulenta* (Cneoraceae). *Molecular Ecology Resources* **9**: 1497–1500.

Rijsdijk KF, **Hengl T**, **Norder SJ**, **Otto R**, ***et al.* 2014**. Quantifying surface-area changes of volcanic islands driven by Pleistocene sea-level cycles: biogeographical implications for the Macaronesian archipelagos. *Journal of Biogeography* **41**: 1242–1254.

Roquet C, **Sáez L**, **Aldasoro JJ**, **Susanna A**, **Alarcón ML**, **Garcia-Jacas N. 2008**. Natural delineation, molecular phylogeny and floral evolution in Campanula. *Systematic Botany* **33**: 203–217.

Ruiz Guajardo JC, **Schnabel A**, **Ennos R**, **Preuss S**, **OteorArnaiz A**, **Stone G. 2010**. Landscape genetics of the key African acacia species *Senegalia mellifera* (Vahl) – the importance of the Kenyan Rift Valley. *Molecular Ecology* **19**: 5126–5139.

Sanmartín I, **Anderson CL**, **Alarcon M**, **Ronquist F**, **Aldasoro JJ. 2010**. Bayesian island biogeography in a continental setting: the Rand Flora case. *Biology Letters* **6**: 703– 707.

Santos-Guerra AS. 1990. *Bosques de laurisilva en la región macaronésica. Serie Naturaleza y Medio Ambiente, No. 49*. Strasbourg: Council of Europe.

Saro I, **González-Pérez MA**, **García-Verdugo C**, **Sosa PA. 2015**. Patterns of genetic diversity in *Phoenix canariensis*, a widespread oceanic palm (species) endemic from the Canarian archipelago. *Tree Genetics & Genomes* **11**: 1–13.

Senut B, **Pickford M**, **Ségalen L. 2009**. Neogene desertification of Africa. *Comptes Rendus Geoscience* **341**: 591–602.

Sepulchre P, **Ramstein G**, **Fluteau F**, **Schuster M**, **Tierceli JJ**, **Brunet M. 2006**. Tectonic uplift and Eastern Africa aridification. *Science* **313**: 1419–1423.

Seto K, **Guneralp B**, **Hutyra L. 2012**. Global forecasts of urban expansion to 2030 and direct impacts on biodiversity and carbon pools. *Proceedings of the National Academy of Sciences of the USA* **109**: 16083–16088.

Silvestrini M, **Junqueira MG**, **Favarin AC**, ***et al.* 2007**. Genetic diversity and structure of Ethiopian, Yemen and Brazilian *Coffea arabica* L. accessions using microsatellites markers. *Genetic Resources and Crop Evolution* **54**: 1367–1379.

Sosa PA, **González-Pérez MA**, **González-González E**, ***et al.* 2010*a***. Biología de la conservación de endemismos vegetales de los Parques Nacionales canarios: caracterización genética y demográfica. In: Ramírez L, Asensio B, eds. *Proyectos de investigación en parques nacionales 2006–2009*. Madrid: Organismo Autónomo Parques Nacionales, 225–248.

Sosa PA, **González-Pérez MÁ**, **Moreno C**, **Clarke JB. 2010*b***. Conservation genetics of the endangered endemic *Sambucus palmensis* Link (Sambucaceae) from the Canary Islands. *Conservation Genetics* **11**: 2357–2368.

Stebbins GL. 1980. Rarity of plant species: a synthetic viewpoint. *Rhodora* **82**: 77–86.

Stebbins GL, **Major J. 1965**. Endemism and speciation in the California flora. *Ecological Monographs* **35**: 1–35.

Tilman D. 1999. The ecological consequences of changes in biodiversity: a search for general principles. *Ecology* **80**: 1455–1474.

Tilman D, **May RM**, **Lehman CL**, **Nowak MA. 1994**. Habitat destruction and the extinction debt. *Nature* **371**: 65–66.

Triantis KA, **Borges PA**, **Ladle RJ**, ***et al.* 2010**. Extinction debt on oceanic islands. *Ecography* **33**: 285–294.

Tsutsui ND, **Suarez AV**, **Holway DA**, **Case TJ. 2000**. Reduced genetic variation and the success of an invasive species. *Proceedings of the National Academy of Sciences of the USA* **97**: 5948–5953.

Wang Q, **Zhou SL**, **Hong DY. 2013**. Molecular phylogeny of the platycodonoid group. (Campanulaceae s. str.) with special reference to the circumscription of *Codonopsis*. *Taxon* **62**: 498–504.

White F. 1983. *The vegetation of Africa: a descriptive memoir to accompany the UNESCO/AETFAT/UNSO vegetation map of Africa. Natural Resources Research Report No. 20*. Paris: UNESCO.

Williamson-Natesan EG. 2005. Comparison of methods for detecting bottlenecks from microsatellite loci. *Conservation Genetics* **6**: 551–562.

Yannic G, **Pellissier L**, **Ortego J**, ***et al.* 2013**. Genetic diversity in caribou linked to past and future climate change. *Nature Climate Change* **4**: 132–137.

Yeh FC, **Yang RC**, **Boyle TB**, **Ye ZH**, **Mao JX. 1997**. *POPGENE, the user-friendly shareware for population genetic analysis*. https://sites.ualberta.ca/~fyeh/popgene.html. Molecular Biology and Biotechnology Centre, University of Alberta, Canada.

CHAPTER 5

Living on the edge: timing of Rand Flora disjunctions congruent with ongoing aridification in Africa

Lisa Pokorny, Ricarda Riina, Mario Mairal, Andrea S. Meseguer,
Victoria Culshaw, Jon Cendoya, Miguel Serrano, Rodrigo
Carbajal, Santiago Ortiz, Myriam Heuertz and
Isabel Sanmartín

This chapter has been published in the peer-review journal
Frontiers in Genetics (2015) 6: 1-15

https://doi.org/10.3389/fgene.2015.00154 frontiers in Genetics

ORIGINAL RESEARCH
published: 01 May 2015
doi: 10.3389/fgene.2015.00154

Living on the edge: timing of Rand Flora disjunctions congruent with ongoing aridification in Africa

Lisa Pokorny [1]****, Ricarda Riina*** [1]***, Mario Mairal*** [1]***, Andrea S. Meseguer*** [2]***, Victoria Culshaw*** [1]***, Jon Cendoya*** [1]***, Miguel Serrano*** [3]***, Rodrigo Carbajal*** [3]***, Santiago Ortiz*** [3]***, Myriam Heuertz*** [4, 5, 6] ***and Isabel Sanmartín*** [1]*******

[1] Real Jardín Botánico (RJB-CSIC), Madrid, Spain, [2] INRA, UMR 1062, Centre de Biologie pour la Gestion des Populations (INRA, IRD, CIRAD, Montpellier SupAgro), Montferrier-sur-Lez, France, [3] Department of Botany, Pharmacy School, University of Santiago de Compostela, Santiago de Compostela, Spain, [4] Forest Research Centre (INIA-CIFOR), Madrid, Spain, [5] INRA, BIOGECO, UMR 1202, Cestas, France, [6] University of Bordeaux, BIOGECO, UMR 1202, Talence, France

Edited by:
James Edward Richardson,
Royal Botanic Garden Edinburgh, UK

Reviewed by:
Thomas L. P. Couvreur,
Institut de Recherche pour le Développement, Cameroon
Lars Chatrou,
Wageningen University, Netherlands

****Correspondence:***
Lisa Pokorny and Isabel Sanmartín,
Real Jardín Botánico (RJB-CSIC),
Plaza de Murillo 2, 28014 Madrid, Spain
pokorny@rjb.csic.es;
isanmartin@rjb.csic.es

Specialty section:
This article was submitted to Evolutionary and Population Genetics, a section of the journal Frontiers in Genetics

Received: *06 October 2014*
Accepted: *05 April 2015*
Published: *01 May 2015*

Citation:
Pokorny L, Riina R, Mairal M, Meseguer AS, Culshaw V, Cendoya J, Serrano M, Carbajal R, Ortiz S, Heuertz M and Sanmartín I (2015) Living on the edge: timing of Rand Flora disjunctions congruent with ongoing aridification in Africa. Front. Genet. 6:154. doi: 10.3389/fgene.2015.00154

The Rand Flora is a well-known floristic pattern in which unrelated plant lineages show similar disjunct distributions in the continental margins of Africa and adjacent islands—Macaronesia-northwest Africa, Horn of Africa-Southern Arabia, Eastern Africa, and Southern Africa. These lineages are now separated by environmental barriers such as the arid regions of the Sahara and Kalahari Deserts or the tropical lowlands of Central Africa. Alternative explanations for the Rand Flora pattern range from vicariance and climate-driven extinction of a widespread pan-African flora to independent dispersal events and speciation *in situ*. To provide a temporal framework for this pattern, we used published data from nuclear and chloroplast DNA to estimate the age of disjunction of 17 lineages that span 12 families and nine orders of angiosperms. We further used these estimates to infer diversification rates for Rand Flora disjunct clades in relation to their higher-level encompassing lineages. Our results indicate that most disjunctions fall within the Miocene and Pliocene periods, coinciding with the onset of a major aridification trend, still ongoing, in Africa. Age of disjunctions seemed to be related to the climatic affinities of each Rand Flora lineage, with sub-humid taxa dated earlier (e.g., *Sideroxylon*) and those with more xeric affinities (e.g., *Campylanthus*) diverging later. We did not find support for significant decreases in diversification rates in most groups, with the exception of older subtropical lineages (e.g., *Sideroxylon*, *Hypericum*, or *Canarina*), but some lineages (e.g., *Cicer*, *Campylanthus*) showed a long temporal gap between stem and crown ages, suggestive of extinction. In all, the Rand Flora pattern seems to fit the definition of biogeographic pseudocongruence, with the pattern arising at different times in response to the increasing aridity of the African continent, with interspersed periods of humidity allowing range expansions.

Keywords: Africa, historical biogeography, climate change, diversification rates, long-distance dispersal, Rand Flora, vicariance

Introduction

Large-scale biodiversity patterns have intrigued naturalists since the eighteenth century (Forster, 1778; von Humboldt and Bonpland, 1805; Wallace, 1878; Fischer, 1960; Stevens, 1989;

Lomolino et al., 2010). Recognizing that spatial variation in environmental variables such as temperature or precipitation is insufficient to explain such patterns, more integrative explanations that emphasize the role of both environmental and evolutionary factors have recently been advanced (Qian and Ricklefs, 2000; Wiens and Donoghue, 2004; Jablonski et al., 2006). As Wiens and Donoghue (2004) state "environmental variables cannot by themselves increase or decrease local or regional species richness"; only evolutionary processes such as dispersal, speciation and extinction can. Therefore, reconstructing rates of dispersal, speciation, and extinction across the component lineages of a biota might help us understand how assembly took place across space and through time (Pennington et al., 2004; Ricklefs, 2007; Wiens, 2011). Moreover, understanding patterns of biotic assembly is a pressing goal in biodiversity research at a time when nearly one tenth of species on Earth are projected to disappear in the next hundred years (Maclean and Wilson, 2011).

Africa is a continent especially interesting to study patterns of biotic assembly. On one hand, African tropical regions are comparatively species-poorer than regions situated in the same equatorial latitudes in the Neotropics and Southeast Asia (Lavin et al., 2001; Couvreur, 2015), which has led to the continent being referred to as the "odd man out" (Richards, 1973). On the other, Africa offers some extraordinary examples of continent-wide disjunctions. For example, tropical rainforests in Africa appear in two main blocks, the West-Central Guineo-Congolian region and the coastal and montane regions of East Africa, now separated by a 1000 Km-wide arid corridor (Couvreur et al., 2008). Another prime example is the so called *Rand Flora* (RF), a biogeographic pattern in which unrelated plant lineages show comparable disjunct distributions with sister taxa occurring on now distantly located regions in the continental margins of Africa: Macaronesia-northwest Africa, Western African mountains, Horn of Africa-South Arabia (including the Island of Socotra), Eastern Africa (incl. Madagascar), and Southern Africa (Christ, 1892; Lebrun, 1947, 1961; Quézel, 1978; Andrus et al., 2004; Sanmartín et al., 2010; **Figure 1**). All RF lineages share sub-humid to xerophilic affinities, so that the tropical lowlands of Central Africa and the large Sahara and Arabian deserts in the north or the Namib and Kalahari deserts in the south presumably constitute effective climatic barriers to their dispersal.

Swiss botanist K. H. H. Christ (1892) first referred to "cette flore *marginale* de l'Afrique," that is "this *marginal* African flora," in a note addressing the role the so called *ancient African flora* played on European floras, with emphasis on the Mediterranean biome. Later, in his "Die Geographie der Farne" (i.e., "The Geography of Ferns"; Christ, 1910), he very aptly named this geographic pattern "Randflora" (see pp. 259–275), where the Germanic word "Rand" stands for rim, edge, border, margin (see **Figure 1** inset), noting its similarities with Engler's "afrikanisch-makaronesische Element" (Engler, 1879, 1910; see pp. 76 in the former and pp. 983–984 and 1010 in the latter), that is, an "Afro-Macaronesian element" linking disjunct xerophilic taxa found in the continental margins of Africa and its adjacent islands (e.g., Canary Islands, Cape Verde, etc.).

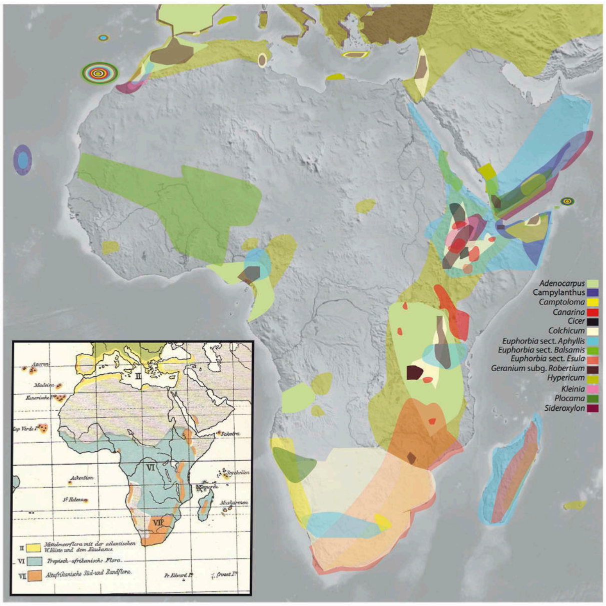


FIGURE 1 | Rand Flora disjunction pattern as evidenced by angiosperm plant lineages analyzed for this study. The inset shows K.H.H. Christ's (1910) depiction of "cette flore *marginale* de l'Afrique" or "Randflora" (in orange color), note their similar geographic limits. Taxa: *Adenocarpus* (Fabaceae), *Camptoloma* (Scrophulariaceae), *Campylanthus* (Plantaginaceae), *Canarina* (Platycodoneae, Campanulaceae), *Cicer* (Fabaceae), *Colchicum* (Colchicaceae), *Euphorbia* subgen. *Athymalus* (sects. *Anthacanthae* and *Balsamis*; Euphorbiaceae), *Euphorbia* subgen. *Esula* (sect. *Aphyllis*), *Euphorbia* subgen. *Esula* (African clade of sect. Esula), *Geranium* subgen. *Robertium* (Geraniaceae), *Hypericum* (Hypericaceae), *Kleinia* (Asteraceae), *Plocama* (Rubiaceae), and *Sideroxylon* (Sideroxyleae, Sapotaceae).

Historical explanations for this pattern and, in particular, its temporal framework, its exact boundaries, and the ecology of the plants involved have varied through these past two centuries. The early view (Engler, 1879, 1910; Christ, 1892, 1910) was one of a pan-African flora found throughout the continent that became restricted to its margins as a result of major climate changes (i.e., increasing aridification) throughout the *Tertiary* (i.e., the Cenozoic Period, 66.0–2.58 Ma). Lebrun (1947; see pp. 134–137), and later Monod (1971, p. 377) and Quézel (1978, p. 511), interpreted Christ's *ancient African flora* as a complex ensemble that had experienced alternating expansions and contractions through time, having had a chance to spread across northern Africa during favorable moments in the Miocene and needing to retract at the end of the Neogene (i.e., Pliocene): a further increase in aridity at the beginning of Pleistocene glaciations would have confined relictual or vicariant taxa to Macaronesia, northwest Africa and Arabia. Axelrod and Raven (1978) explained some of these disjunctions in relation to a more ancient, widespread Paleogene flora of subtropical origin that covered the entire African continent at the beginning of the Cenozoic, and that was decimated by successive events of aridification, of which the relict floras of Macaronesia, the Cape Region, and the Afromontane forests in eastern and western Africa would be remnants. Bramwell (1985) explains this pattern in terms of pan-biogeographic "general tracks" that connect what would be the remains of an ancient flora that extended across the Mediterranean and Northern Africa in the Miocene, and whose vestiges could be found in the Macaronesian laurisilva and a few enclaves in the island of Socotra, the Ethiopian Highlands and southern Yemen.

These authors share a vicariant perspective and presume RF lineages were part of a widespread pan-African *Tertiary* flora that became fragmented by the appearance of climatic barriers (i.e., aridification), leaving relictual lineages with reduced distributions at "refugia" in the margins of Africa (i.e., "continental" islands). This "refugium" idea rests on the fact that many of these RF regions—Macaronesia, the South African Cape region, and the semi-arid regions of Eastern Africa and Southern Arabia (e.g., Ethiopia, Yemen, Socotra)—harbor a large number of endemic species, when compared to neighboring areas. Moreover, the "fragmentation-refugium" hypothesis implies the disappearance, possibly by extinction, of RF lineages from part of their distributional range (e.g., across the Sahara in central Northern Africa), which is consonant with the "climatic vicariance" concept (Wiens, 2004): an environmental change creates conditions within a species' geographic range that are outside the ancestral climatic tolerances; individuals are unable to persist and the species' geographic range becomes fragmented.

The alternative explanation is one of independent dispersal (immigration) events among geographically isolated regions and subsequent speciation *in situ*. In this framework, divergence events need not be congruent across lineages, since long-distance dispersal (LDD) events are highly stochastic in nature (Nathan, 2006). Asides from transoceanic dispersal—which has been postulated in the case of *Aeonium* (Kim et al., 2008), *Geranium* (Fiz et al., 2008), and other RF lineages (Andrus et al., 2004) based on molecular phylogenetic evidence—, cross-continent LDD dispersal is also possible: published examples favoring cross-continent LDD include *Senecio*, with a disjunct distribution between Macaronesia-Northern Africa and South Africa (Coleman et al., 2003; Pelser et al., 2012). Moreover, dispersal does not necessarily imply long-distance migration events. In some cases, dispersal across intermediate areas that act as "stepping stones" or "land bridges" could have been possible. For example, the presence of isolated mountain ranges (offering suitable habitats) throughout the Sahara, such as the Tibesti and Hoggar massifs, could have allowed this short or medium-range dispersal in *Campanula* (Alarcón et al., pers. comm.). Correspondingly, some RF lineages might have used the Arabian Plate as a land bridge to reach East Africa (*Campanula*, Roquet et al., 2009; *Hypericum*, Meseguer et al., 2013), and others may have benefited from the new habitats offered by the Pliocene uplift of the Eastern Arc Mountains to migrate to or from South Africa (Meseguer et al., 2013).

Discriminating between climate-driven vicariance vs. independent dispersal events between geographically isolated regions requires framing the evolution of disjunct lineages on a temporal scale (Sanmartín, 2014). On the other hand, to unravel the origin of a biota or biome, a meta-analysis across dated phylogenies of multiple non-nested clades is needed (Pennington et al., 2010; Wiens, 2011; Couvreur, 2015). Sanmartín et al. (2010) carried out a meta-analysis of 13 lineages to infer relative rates of historical dispersal among RF regions (Macaronesia, Eastern Africa-Southern Arabia, and Southern Africa) and found the highest rate of biotic exchange between east and west Northern Africa, across the Sahara. However, they did not integrate absolute estimates of lineage divergences in their inference, since very few RF lineages (e.g., Roquet et al., 2009) had been dated at the time.

In this study, we estimate time divergences for up to 13 plant lineages (**Table 1**) displaying RF disjunct distributions (**Figure 1**), and use published divergence times for four other lineages (see Materials and Methods), in order to provide a much-needed temporal framework for this pattern. An extensive description of each of these lineages, geographic distributions and phylogenetic relationships is provided in Supplementary Materials. We also frame these disjunctions in the context of major climatic and geological events in the history of Africa (see summary below) and estimate net diversification rates in an attempt to address the role that evolutionary processes, such as climate-driven extinction, may have played in the formation of the African RF pattern.

Materials and Methods

Study Area: African Climate through Time

To understand biogeographic patterns in the African flora, it is necessary to briefly review the climatic and geological history that might have influenced the evolution of African plant lineages. Extensive reviews of African climatic and vegetation history can be found in Axelrod and Raven (1978); van Zinderen Bakker (1978); Maley (1996, 2000); Morley (2000); Jacobs et al. (2010), Plana (2004), and Bonnefille (2011), among others.

During the Late Mesozoic, Africa was part of the supercontinent Gondwana, located in the southern hemisphere, and enjoyed a relatively humid and temperate climate (Raven and Axelrod, 1974). After breaking up from South America ca. 95 Ma, Africa started moving northwards toward the equatorial zone (**Figure 2A**). The result was a general trend toward continental aridification in which different regions became arid or wet at alternative times (**Figure 2B**, Senut et al., 2009). Paleocene Africa (66–56 Ma) was mainly wet and warm, characterized by a major diversification in the West African flora (Plana, 2004). A global increase in temperatures in the Eocene (56–33.9 Ma) led to increased aridity in Central Africa, with a rainforest-savannah mosaic in the Congo region. This was followed by a global cooling event at the Eocene-Oligocene boundary (33.9 Ma), which led again to aridification and major extinction but did not change biome composition (Axelrod and Raven, 1978).

The Early Miocene (23–16 Ma) was warm and humid, with wide extension of rainforests, from the northern Sahara to parts of Southern Africa. The Mid Miocene (16–11.6 Ma) was a period of major changes in climate and topography. A combination of factors, including the gradual uplift of Eastern Africa, the successive closure of the Tethys seaway in the north, and the expansion of the East Antarctic ice sheet in the south (Trauth et al., 2009), led to a general intensification of the aridification process, though it was not homogeneous across the continent. Geological and paleontological evidence suggest that now arid regions (e.g., northern Africa, Horn of Africa, Namib Desert) were during

TABLE 1 | Rand Flora disjunctions, encompassing (higher level) lineages, recent molecular phylogenetic studies, and molecular markers used in here.

Order	Family	Tribe (or else)	Genus	Subgenus	Section (or else)	Disjunction name	Dataset reference	Molecular marker	
								Nuclear	Chloroplast
Fabales	Fabaceae	Genisteae	*Adenocarpus*			*Ad. manii*	Cubas et al., 2010	ETS, ITS	*trn*LF
Saxifragales	Crassulaceae	*Aeonium* alliance	*Aeonium*			*Ae. leucoblepharum*	Mort et al., 2002, 2007	ITS	–
Malpighiales	Euphorbiaceae		*Euphorbia*	*Athymalus*	*Anthacanthae* *Balsamis*	*Eu. omariana* *Eu. balsamifera*	Peirson et al., 2013	ITS	*ndh*F
Malpighiales	Euphorbiaceae			*Esula*	*Aphyllis Esula*	*Eu. tuckeyana* *Eu. usambarica* *Eu. schimperiana*	Barres et al., 2011; Riina et al., 2013	ITS	*ndh*F
Asterales	Campanulaceae		*Campanula*		*Azorina* (clade)	*Ca. jacobaea*	Alarcón et al., 2013	–	*trn*LF, *pet*BD, *rpl*32–*trn*L, *trn*SG
Lamiales	Scrophulariaceae	Buddlejoideae (subfamily)	*Camptoloma*			*Cm. canariense* *Cm. rotundifolium*	Kornhall et al., 2001; Oxelman et al., 2005	–	*trn*LF, *ndh*F, *rps*16
Lamiales	Plantaginaceae	Globularieae	*Campylanthus*			*Cy. salsoloides*	Thiv et al., 2010	ITS	*atp*B-*rbc*L
Asterales	Campanulaceae	Platycodoneae	*Canarina*			*Cn. canariensis*	Mairal et al., 2015	ITS	*pet*BD, *psb*J, *trn*LF, *trn*SG
Fabales	Fabaceae	Vicioids (clade)	*Cicer*			*Ci. canariense*	Javadi et al., 2007	ETS, ITS	*trn*SG, *mat*K, *trn*AH, *trn*A-Leu
Liliales	Colchicaceae	Colchiceae	*Colchicum*			*Co. schimperianum*	Manning et al., 2007; del Hoyo et al., 2009	–	*trn*LF, *atp*B-*rbc*L, *rps*16
Geraniales	Geraniaceae		*Geranium*	*Robertium*		*G. robertianum*	Fiz et al., 2008	ITS	–
Malpighiales	Hypericaceae	Hypericeae	*Hypericum*		*Androsaemum* *Campylosporus*	*H. scopulorum* *H. quartinianum*	Meseguer et al., 2013	–	*trn*LF, *trn*SG
Asterlaes	Asteraceae	Senecioneae	*Kleinia*			*K. neriifolia*	Pelser et al., 2007	ITS	*trn*LF
Gentianales	Rubiaceae	Putorieae	*Plocama*			*Pl. pendula* *Pl. crocyllis*	Backlund et al., 2007	–	*rps*16, *trn*TF, *atp*B-*rbc*L
Ericales	Sapotaceae	Sideroxyleae	*Sideroxylon*			*S. spinosus*	Smedmark et al., 2006; Smedmark and Anderberg, 2007	–	*ndh*F, *trn*H–*psb*A, *trn*CD

GenBank numbers can be found in the references listed under column "Dataset reference."

this period more humid than they are today, whereas other now humid regions (e.g., Congo Basin) were much drier (**Figure 2B**). Desertification started in the southwest (Namib Desert) around 17–16 Ma ago, and proceeded eastward and northward. In Southern Africa, tropical to subtropical vegetation was replaced by wooded savannah during the lower Mid-Miocene (Senut et al., 2009). In Northern Africa, the earliest evidence of aridity in the Sahara region is from the Late Miocene (11.6–5.3 Ma), ca. 7–6 Ma (Senut et al., 2009; **Figure 2B**). In Central Africa, a semi-arid desert ("Miocene Congo Desert," **Figure 2B**) occupied the region until the Mid Miocene, 13–12 Ma ago, when the Eastern African uplift and subsequent subsidence led to the establishment of the Congo River drainage and a general increase in humidity ("tropicalization"). Also in the Late Miocene, ca. 7–8 Ma, a new period of tectonic activity in Eastern Africa led to the uplift of the Eastern Arc Mountains and the uplands of West Central Africa (Cameroon volcanic line), which led to increasing aridity and the expansion of savannahs and grasslands in these regions (Sepulchre et al., 2006). Uplifting reached a maximum during the Plio-Pleistocene and led to the formation of the Ethiopian Highlands and the desertification of low-lying areas in the Horn of Africa (Senut et al., 2009). From the Late Pliocene to the Holocene, the alternation of glacial-and interglacial periods seems to have led to repeated contractions and expansions of distributional ranges across both subtropical and tropical taxa (Maley, 2000; Bonnefille, 2011). Some areas like the Saharan massifs of Tibesti and Hoggar or the Ennedi Mountains could have served as refuges during arid periods for subtropical taxa (Osborne et al., 2008), whereas the uplands of Upper and Lower Guinea and the east of the Congo Basin, the Albertine Rift, or the Eastern Arc Mountains could have played the same role for tropical plant taxa (Maley, 1996; **Figure 2B**).

Taxon Sampling

We retrieved sequences from GenBank from existing studies (**Table 1**) for the following 13 lineages exhibiting a distribution congruent with the RF pattern (Andrus et al., 2004; Sanmartín et al., 2010): *Adenocarpus*, *Aeonium*, *Camptoloma*, *Campylanthus*, *Cicer*, *Colchicum*, *Euphorbia* sects. *Antachanthae*, *Aphyllis*, *Balsamis*, and *Esula*, *Geranium*, *Kleinia*, and *Plocama* (**Figure 3**).

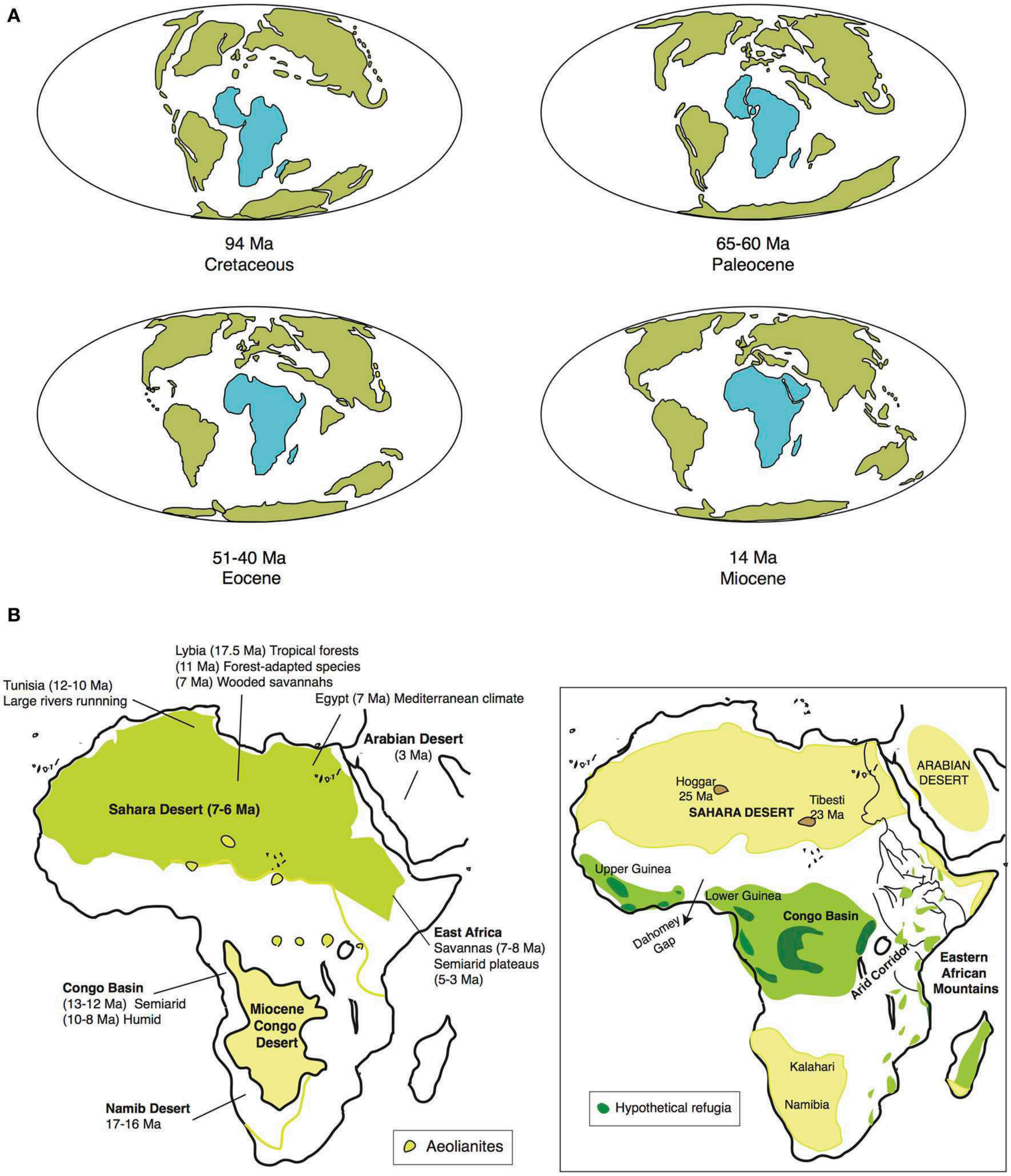


FIGURE 2 | (A) Tectonic fragmentation of the supercontinent Gondwana through time, showing Africa's drift northwards; and **(B)** main climatic events in Africa during Neogene (adapted from Senut et al., 2009): (B-left) Early Neogene Central Africa was more arid than North Africa, with a desert, semiarid region in the Congo Basin. Desertification started in southwest Africa in the Mid-Miocene, proceeding eastward and northward, and finalizing with the formation of the Sahara Desert. Conversely, Central Africa became tropical due to subsidence and Eastern African uplift. (B-right) Schematic representation of present-day vegetation belts, showing position of main deserts and rainforest refugia (Eastern Arc Mountains/Guineo-Congolian region (the latter fragmented into smaller refugia). Rand Flora lineages occupy the regions in the margin that are not deserts or rainforests, rarely some find refuge in mountain areas of North African Sahara (e.g., Tibesti and Hoggar Massifs).

We chose these lineages because sampling is nearly complete in most cases with very few to no missing taxa. Most of these RF taxa have been sequenced for several markers from the nuclear and chloroplast DNA regions. For each group we selected the markers with most sequences and tried representing both genomic compartments whenever possible. The sequences were aligned using the Opalescent package (Opal v2.1.0; Wheeler and Kececioglu, 2007) in Mesquite v3.01 (Maddison and Maddison, 2014) and

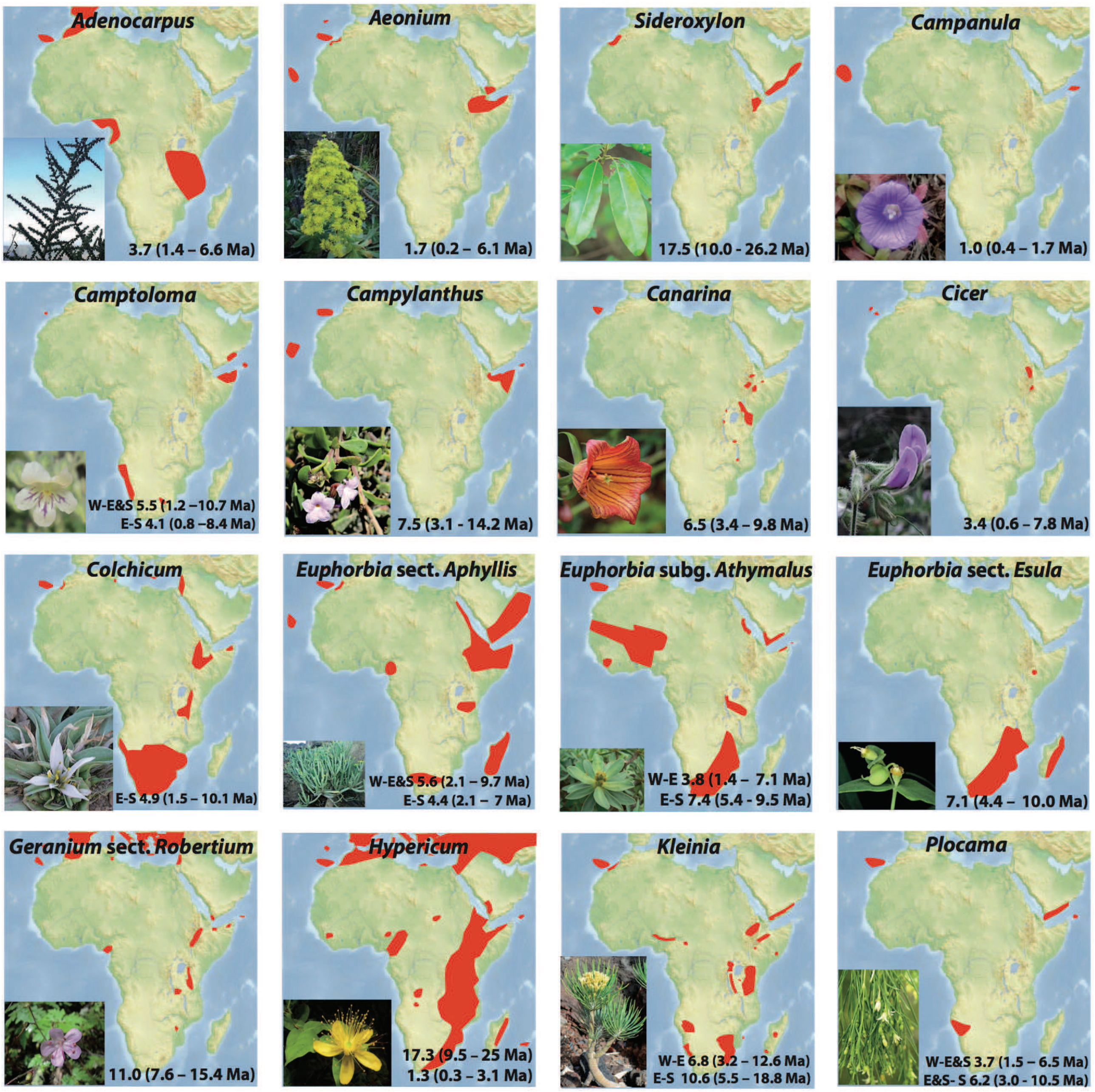


FIGURE 3 | Individual distributions and habit illustrations for 16 plant lineages exhibiting Rand Flora disjunctions. Estimated divergence times within each lineage correspond to the disjunctions represented in **Figures 4**, **5** and indicated in the MCC chronograms shown in **Figures S1–S16**. Taxa names correspond to those in **Table 1**.

manually adjusted in SE-AL v2.0a11 (Rambaut, 2002) using a similarity criterion, as recommended by Simmons (2004). For four other RF lineages —*Campanula* (Alarcón et al., 2013), *Canarina* (Mairal et al., 2015), *Hypericum* (Meseguer et al., 2013), and *Sideroxylon* (Stride et al., 2014)—we used recently published time estimates by our research team (except for *Sideroxylon*, which nonetheless used a dating approach similar to ours). Approximately 1600 sequences from ca. 675 taxa from 12 families and 9 orders of angiosperms were included in our study (**Table 1**).

Estimating Absolute Divergence Times

Divergence times were estimated under a Bayesian framework in BEAST v1.8 (Drummond et al., 2012). For each lineage, we constructed a dataset including the markers listed in **Table 1**, which were partitioned by genome (chloroplast vs. nuclear), whenever possible. The best-fitting substitution model for each partition was selected using the Akaike Information Criterion implemented in MrModeltest v2.2 (Nylander, 2004) and run in PAUP* v4.0b (Swofford, 2002). The relaxed uncorrelated lognormal clock model (UCLD, Drummond et al., 2006) and a Yule speciation process as tree model were selected for all datasets based on preliminary explorations. MCMC searches were run 5×10^7 generations and sampled and logged every 2500th generation. We used Tracer v1.6 (Rambaut et al., 2013) to determine stationarity of the Markov chain and to verify that all parameters had large enough effective sampling sizes (ESS>200). TreeAnnotator v1.8.0 (Drummond et al., 2012) and FigTree v. 1.4.2 (Rambaut, 2009) were used respectively to generate and visualize the resulting maximum clade credibility (MCC) chronograms.

Calibration points for obtaining absolute divergence times were based on either the fossil record or on published secondary calibration constraints (**Table 2**). The latter were obtained from published dated phylogenies of datasets including our study groups (e.g., the family to which the genus belongs), and were assigned normal distribution priors (Ho and Phillips, 2009) in the BEAST analysis that encompassed the mean and the 95% highest posterior density (HPD) confidence interval (CI) from these studies [except in the case of time constrains from Bell et al. (2010), for which a lognormal distribution was used, since posterior estimates for a normal prior were not available]. For fossil calibration points we used a lognormal prior, since this distribution better represents the stratigraphic uncertainty associated to the fossil record (Ho and Phillips, 2009). The offset of the lognormal distribution was set to the upper bound of the stratigraphic period where the fossil was found, and the standard deviation (SD) and mean were set so that the 95% CI encompassed the lower and upper bound of the period (e.g., for Late Eocene *Hypericum antiquum* a lognormal distribution offset at 33.9 Myr, with mean = 1.0 and SD = 0.7, was used to cover the length of the period where the fossil was found, that is 33.9–37.2 Ma). A summary of time constraints used for each dataset and their provenance can be found in **Table 2**.

Diversification Analyses

We used divergence times estimated above to calculate absolute diversification rates in the aforementioned lineages. There have been numerous developments in macroevolutionary birth-death models that allow a more accurate estimation of extinction and speciation rates from dated molecular phylogenies, including episodic time-variable models and trait-dependent diversification models (Stadler, 2013; Morlon, 2014; Rabosky et al., 2014).

TABLE 2 | Time constraints and prior probability distributions imposed on constrained nodes to estimate divergence times in RF lineages.

Taxon set	Node constrained	Time constraint (Myr)			Dating reference	Figure/Table/P.
		Distribution (offset)	Mean	SD		
Adenocarpus	ROOT: Genisteae	Normal	19.5	3.8	Lavin et al., 2005	Table 2, node 32
Aeonium alliance	ROOT: *Aeonium* alliance	Normal	18.83	1.0	Kim et al., 2008	Figure 2C
E. subg. *Athymalus*	*Athymalus* w/o *E. antso*	Normal	10.78	2.0	Horn et al., 2014	Figure 2
sect. *Anthacanthae*	CROWN: *Athymalus*	Normal	24.56	5.0		Table 1
and sect. *Balsamis*	*Anthacanthae*	Normal	18.22	3.4		Table 1
	MRCA *Anthacanthae-Balsamis*	Normal	7.56	1.4		Figure 2
E. subg. *Esula*	MRCA *Aphyllis-Exiguae* II	Normal	10.36	2.3	Horn et al., 2014	Figure 2
sect. *Aphyllis*	CROWN: *Aphyllis*	Normal	7.37	2.0		Figure 2
E. subg. *Esula*	MRCA *Arvales-Esula*	Normal	10.98	2.4	Horn et al., 2014	Figure 2
sect. *Esula*	CROWN: *Esula*	Normal	8.6	2.4		Figure 2, node 5
(African clade)	*E. virgata* clade	Normal	5.4	1.4		Figure S2
Camptoloma	MRCA *Buddlejeae-Camptoloma*	Normal	20.0	6.0	Navarro-Pérez et al., 2013	Figure 2
	Buddlejeae	Normal	7.5	3.0		Figure 2
Campylanthus	MRCA *Digitalis-Plantago*	Lognormal (0.0)	38.0	0.2	Bell et al., 2010	Figure S11
	MRCA *Plantago-Aragoa**	Lognormal (7.1)	1.5	1.0	Thiv et al., 2010	P. 610
Cicer	CROWN: *Cicer*	Normal	14.8	5.0	Lavin et al., 2005	Figure 3, node 80
Colchicum	MRCA *Gloriosa-Colchicum*	Normal	43.3	7.0	Chacón and Renner, 2014	Figure 3, node 128/Table 2
Geranium subg. *Robertium*	MRCA *Pelargonium-Geranium*	Normal	28.0	3.0	Fiz et al., 2008	Figure 3, node D
	CROWN: *Robertium*§	Lognormal (7.25)	1.0	1.0		P. 329
Kleinia	ROOT: Asteraceae†	Lognormal (47.5)	10.0	0.75	Barres et al., 2013	P. 872
	Lordhowea insularis	Lognormal (0.0)	7.0	1.0	Pelser et al., 2010	Table 1
Plocama	MRCA *Putorieae-Paederieae*	Normal	34.4	5.5	Bremer and Eriksson, 2009	Table 1

*At least one node (preferably toward the root) was constrained in each phylogeny (**Figures S1–S16** show resulting chronograms explicitly stating any constrained nodes).*

**Plantaginacearumpollis miocenicus (Late Miocene, 10.3 Ma; Nagy, 1963; Doláková et al., 2011).*

§Geranium cf. lucidum (Late Miocene, 7.246 Ma ± 0.005; Van Campo, 1989).

†Raiguenrayun cura (Middle Eocene, 47.5 Ma; Barreda et al., 2012).

However, these methods usually require both very large phylogenies (e.g., ≥100 tips) and a fairly complete sampling. We here chose a simpler approach, the "method-of-moments" estimator (Magallón and Sanderson, 2001), implemented in the R package GEIGER (Harmon et al., 2008). This method uses clade size (extant species number) and clade age (either crown or stem) to estimate net diversification rates (r = speciation minus extinction), under different values of background extinction or turnover rate (ε = extinction/speciation = 0.0, 0.5, and 0.9). Net diversification rates (*bd.ms* function in GEIGER) were here estimated for all RF disjunctions and for a series of successively encompassing clades (e.g., section, genus, tribe, subfamily, and so on) to detect possible rate shifts. Crown diversification rates could not be estimated for clades containing only two taxa because Magallón and Sanderson's formula ($r = [\log(n) - \log 2]/t$ in its simplest version, that is, with no extinction; for $\varepsilon > 0$ see formula number 7 in Magallón and Sanderson, 2001) results in zero in this case. In an attempt to counter this problem, clades containing two taxa were assigned a diversity value of 2.01, which permitted the estimation of net diversification rates (r).

Additionally, the probability of obtaining a clade with the same size and age as the RF disjunction, given the background diversification rate of the encompassing clade/s and at increasing extinction fractions ($\varepsilon = 0$, 0.5, and 0.9), was estimated with the *crown.p* function in GEIGER. We also estimated the 95% confidence interval of expected diversity through time (*crown.limits* function, GEIGER, $\varepsilon = 0$, 0.5, and 0.9) for a clade that diversifies with a rate equal to that of the family containing a RF disjunction with the highest diversification rate (i.e., Asteraceae); we then mapped RF lineages according to their crown or stem age and standing species diversity to assess which RF disjunct clades are significantly less diverse than expected given their stem and crown age in relation to the highest rate calculated for a RF family (Magallón and Sanderson, 2001; Warren and Hawkins, 2006).

Results

Divergence Times

Up to 21 disjunctions were identified and divergence times were estimated for 17 lineages exhibiting a geographic distribution consistent with the RF pattern (**Figures 3**, **4** and **Figures S1–S17**). These disjunctions represent two possible geographic splits: I) Eastern Africa (including the Eastern Arc Mountains, the Horn of Africa, and Southern Arabia) vs. Southern Africa (including southern Angola and Namibia and the Cape Flora region up to the Drakensberg Mountains), hereafter E-S, and II) Western Africa (including Macaronesia and NW Africa south to the Cameroon volcanic line) vs. Eastern Africa, (with or without S Africa), hereafter W-E(&S).

From youngest to oldest, E-S disjunctions (**Figure 4**) occur in *Plocama* (ca. 4 Ma between S African *Pl. crocyllis* on one side and, among other E African-S Arabian species, *Pl. yemenensis* and *Pl. tinctoria* on the other; **Figure 3** and **Figure S15**), *Camptoloma* (ca. 4 Ma between E African *Cm. lyperiiflorum* and S African *Cm. rotundifolium*; **Figure 3** and **Figure S4**), *Colchicum* (ca. 5 Ma between E African *Co. schimperianum* and S African *Co. albanense* and *Co. longipes*, **Figure 3** and **Figure S8**), the African clade of *Euphorbia* sect. *Esula* (ca. 7 Ma between S African and E African taxa; **Figure 3** and **Figure S10**), and *E.* sect. *Anthacanthae* (ca. 7.5 Ma separate subsects. *Platycephalae* and *Florispinae*; **Figure 3** and **Figure S11**).

Also from youngest to oldest, W-E disjunctions (**Figure 4**) can be found in the *Azorina* clade of *Campanula* (ca. 1 Ma between Cape Verdean *Ca. jacobaea* and Socotran *Ca. balfouri*; **Figure 3** and **Figure S3**), in *Hypericum* sect. *Campylosporus* (ca. 1.5 Ma within *H. quartinianum*; **Figure 3** and **Figure S13**), in *Aeonium* (1.7 Ma between E African *Ae. leucoblepharum* and a number of Macaronesian species; **Figure 3** and **Figure S2**), in *Cicer* (ca. 3.5 Ma between Canarian *Ci. canariense* and E African *Ci. cuneatum*; **Figure 3** and **Figure S7**), in *Adenocarpus* (ca. 4 Ma between E African *Ad. mannii* and a number of species in the *Ad. complicatus* complex; **Figure 3** and **Figure S1**), in *Euphorbia* sect. *Balsamis* (ca. 4 Ma between W African *Eu. balsamifera* subsp. *balsamifera* and E African-S Arabian *Eu. balsamifera* subsp. *adenensis*; **Figure 3** and **Figure S11**), in *Camptoloma* (ca. 5.5 Ma between Canarian *Cm. canariense*, on one hand, and E African *Cm. lyperiiflorum* and S African *Cm. rotundifolium*, on the other; **Figure 3** and **Figure S4**), *Eu.* sect. *Aphyllis* (ca. 5.5 Ma between Cape Verdean *Eu. tuckeyana* and all E African and S African species in this section; **Figure 3** and **Figure S9**), *Plocama* (ca. 6 Ma between Canarian *Pl. pendula* and S African *Pl. crocyllis* plus a number of E African/S Arabian *Plocama* species, **Figure 3** and **Figure S16**), in *Canarina* (6.5 Ma between Canarian *Cn. canariensis* and E African *Cn. eminii*; **Figure 3** and **Figure S6**), in *Kleinia* (ca. 7 Ma between the Macaronesian species, on one hand, and a clade of several E African species, on the other; **Figure 3** and **Figure S14**), in *Campylanthus* (ca. 7.5 Ma between the Macaronesian and the E African-S Arabian species in the genus; **Figure 3** and **Figure S5**), in *Geranium* subgen. *Robertium* (ca. 11 Ma between all E African species in this subgenus and a clade formed by W African taxa and a number of broadly distributed circum-Mediterranean and E Asian taxa; **Figure 3** and **Figure S12**), in the *Androsaemum* clade of *Hypericum* (ca. 17 Ma between Socotran *H. scopulorum*, *H. tortuosum* and Turkish *H. pamphylicum*, on one hand, and a number of Macaronesian and W Mediterranean species, on the other; **Figure 3** and **Figure S13**), and in *Sideroxylon* (ca. 17 Ma between Moroccan *S. spinosus* and E African *S. mascatense*; **Figure 3** and **Figure S16**).

Absolute Diversification Rates

Figure 5 and **Table S1** show results from net diversification rate analyses. Most lineages fall within the 95% CI of expected diversity under a no-extinction scenario ($\varepsilon = 0$) in the context of the RF family showing the highest rate of diversification (i.e., Asteraceae). However, some RF disjunct clades were significantly less diverse: W-E disjunctions in *Sideroxylon* (*S. spinosus* vs. *S. mascatense*), *Canarina* (*C. canariensis* vs. *C. eminii*), and *Hypericum* (*H. canariense* clade vs. *H. scopulorum* and *H. pamphylicum*). Other RF disjunct taxa were above the upper bound of the 95% CI: W-E(&S) disjunction in *Euphorbia* sect. *Aphyllis* (S), *Adenocarpus, Aeonium*, and *Campanula*; and E-S disjunction in *Plocama*. Otherwise, all taxa fell within the 95% CI with increasing ε values 0.5 and 0.9, except for *Sideroxylon*.

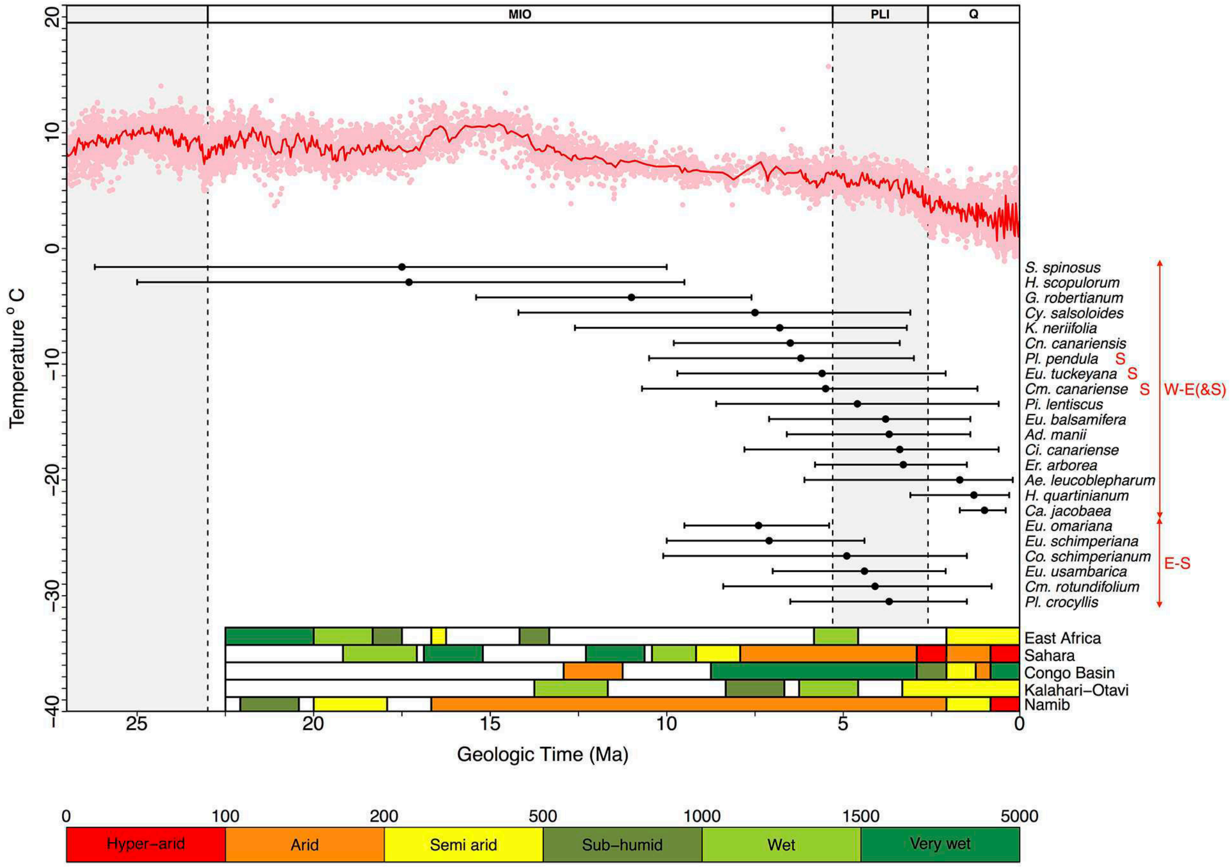


FIGURE 4 | Diagram showing estimated lineage divergence times (mean and 95% HPD confidence intervals) for Rand Flora disjunctions dated in this study and indicated in the MCC chronograms depicted in Figures S1–S17. W-E(&S): divergence times estimated between disjunct taxa distributed in Macaronesia-NW-W Africa vs. Eastern Africa (a red S indicates presence in Southern Africa); E-S: estimated divergence times between disjunct taxa distributed in southern Arabia-Eastern Africa vs. southern Africa. The red line above represents the change in global temperatures over the Cenozoic as reflected by global-deep-sea oxygen records compiled from Zachos et al. (2008); colored bars in the right bottom corner indicate climatic conditions in five regions that underwent major climate changes—either desertification or tropicalization—during the Neogene (adapted from Senut et al., 2009). Taxa names correspond to those in **Table 1**, plus two groups from the literature: *Pistacia lentiscus* and *Erica arborea* (see Discussion).

Interestingly these trends are generally repeated in the more encompassing lineages of the least diverse RF disjunct clades (e.g., *Canarina, Hypericum*, *Sideroxylon*). Notably, though *Camptoloma* has a low extant diversity given its age (three species diverging in the last 6 Myr), the subfamily it belongs to, that is Buddlejoideae, stands above the 95% CI for $\varepsilon = 0$ (**Figure 5**). Something similar can be observed in the case of *Kleinia*, which shows lower diversity than its encompassing lineage, tribe Senecioneae. Another example of potential diversification shift, though in the opposite direction, is that of *Euphorbia*, where the genus is significantly less diverse than expected given its age (for all ε values) but RF disjunct clades are species-richer than expected (i.e., *E.* sect. *Aphyllis*), except for those that fall within the 95% CI limits (e.g., *E.* sect. *Balsamis*, **Figure 5**).

When comparing crown vs. stem age it is noticeable that in some RF disjunct clades crown and stem ages are far apart: *Cicer canariensis* vs. *Ci. cuneatum* (crown age = 3.4 Ma, stem age = 12.2 Ma, with the stem age falling below the lower bound of 95% CIs when $\varepsilon = 0.0$ and 0.9; **Figure 5**). Other examples include, *Camptoloma* (crown age= 5.5 Ma, stem age = 10.2 Ma), *Campylanthus* (crown age = 7.5 Ma, stem age = 20.0 Ma), and most notably *Sideroxylon* (crown age = 17.4 Ma, stem age = 47.3 Ma, **Figure 5**).

Discussion

Rand Flora Disjunctions through Time

Engler's (1910) intuition on the *Tertiary* origins of the Afro-Macaronesian floristic element, aka Christ's (1910) Rand Flora, very much hit the mark on the timing of its assembly. Our divergence estimates for Rand Flora disjunctions span five successive time frames (**Figure 4**): Burdigalian, Tortonian, and Messinian Stages (within the Miocene), the Pliocene, and the Pleistocene. The two earliest disjunctions happen on genera *Sideroxylon* and *Hypericum* and date back to the Early Miocene (Burdigalian; 17.5 and 17.3 Ma, respectively), coinciding with the longest warming

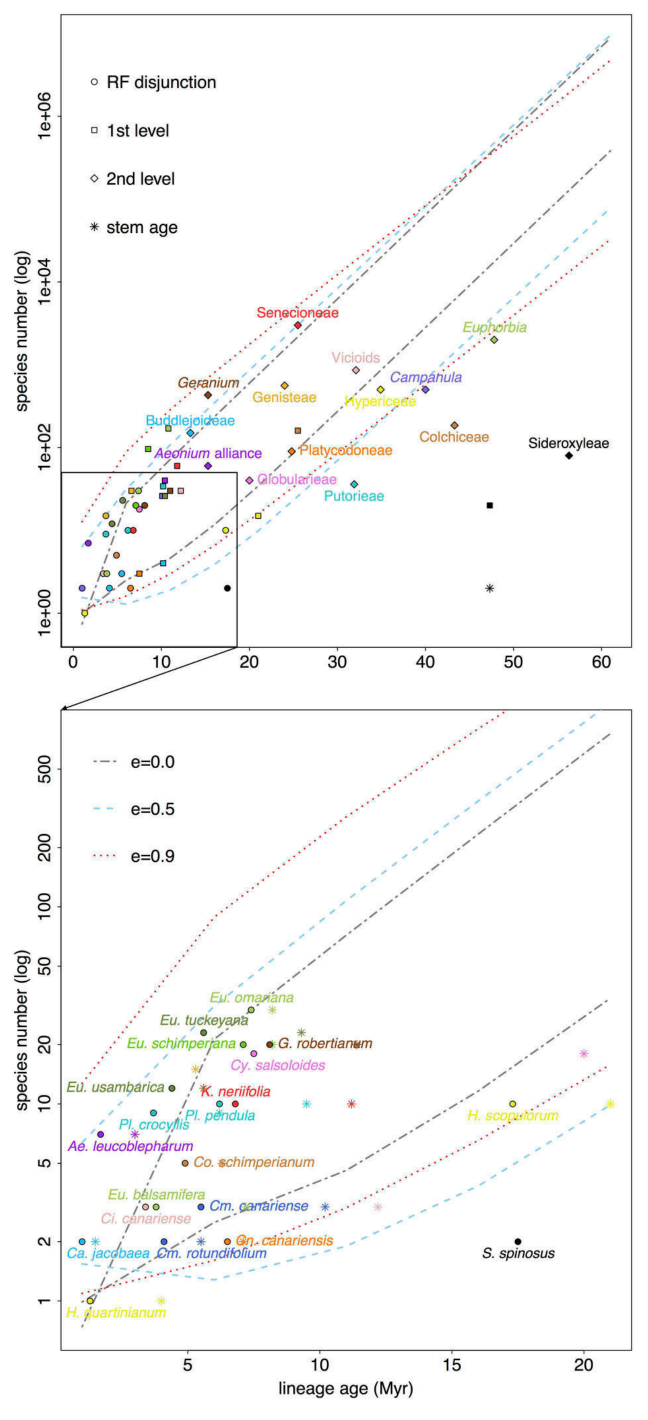


FIGURE 5 | RF lineages (names as in Table 1) are plotted according to their standing diversity (n) and age of the node (circle, crown; star, stem) corresponding to their disjunction (below). Successive encompassing lineages (above) also plotted (squares indicate the clade, section, subgenus the RF disjunct clade falls in; diamonds go one level above indicating genus, tribe, subfamily). Ninety five percent confidence intervals show expected diversity through time for a RF lineage that diversifies at the highest rate estimated (i.e., Asteraceae) given three possible scenarios: no extinction (ε = 0), turnover at equilibrium (ε = 0.5), and high extinction (ε = 0.9). See **Table S1** for associated net diversification rate estimates.

period of the Miocene (the Miocene Climatic Optimum; Zachos et al., 2008) and with the start of desertification in south-central Africa (Senut et al., 2009). Couvreur et al. (2008) also dated divergences in Annonaceae back to this time period and explained them in terms of a once-continuous Early Miocene rainforest that became fragmented by decreasing moisture brought by the closure of the Tethys Sea. The fact that *Sideroxylon* and *Hypericum* exhibit less xeric affinities than other RF lineages, and that their crown diversification dates back to the Paleogene (Meseguer et al., 2013; Stride et al., 2014), suggests these taxa could be relicts of an earlier megathermal flora (sensu Morley, 2000, 2003).

The next disjunction is that of *Geranium* subgen. *Robertium* and it dates back to the Late Miocene (Tortonian, 11.0 Ma). This disjunction follows a drastic decline in global temperatures (Late Miocene cooling, 11.6–5.3 Ma; Beerling et al., 2012) and coincides with the temporary closing of the Panama isthmus in America and a moist "washhouse" climate period in Europe (Böhme et al., 2008). This disjunction marks the separation of Macaronesian (e.g., *G. maderense*) and circum-Mediterranean taxa (e.g., *G. robertianum*), on one side, and E African species (e.g., *G. mascatense*), on the other, leaving open the possibility of a colonization of Macaronesia by a Mediterranean ancestor (**Figure 4** and **Figure S12**). Since the disjunction in *Geranium* subgen. *Robertium* is linked to a more humid period, rather than an increase on aridity, and because the possible Mediterranean origin of its Macaronesian taxa, this lineage does not exactly match the RF pattern.

Most other Neogene disjunctions seem to concentrate around the Miocene-Pliocene border (**Figure 4**). Messinian disjunctions can be observed in *Camptoloma, Campylanthus*, *Canarina*, *Euphorbia* sects. *Anthacanthae* and *Aphyllis*, *Kleinia*, and *Plocama*. Pliocene disjunctions are found in *Adenocarpus*, *Camptoloma*, *Cicer*, *Colchicum*, *Euphorbia.* sects. *Balsamis* and *Aphyllis*, and *Plocama*. These disjunctions follow two different geographic splits, W-E(&S) Africa and E-S Africa. W-E(&S) disjunctions present the widest temporal (as well as spatial) range. Besides the lineages dated here, other examples can be found in the literature of this W-E(&S) disjunction, e.g., according to Xie et al. (2014), in the Anacardiaceae *Pistacia lentiscus* and *P. aethiopica* diverged 4.55 Ma (see **Figure S17**). E-S disjunctions link South Africa and adjacent areas to the East African Rift Mountains, the Ethiopian Highlands, and the Arabian Peninsula. The timing of these E-S disjunctions (Mio-Pliocene) matches the uplift of the Eastern Arc Mountains (Sepulchre et al., 2006). The absence of W-S disjunctions is notable and probably results from African aridification having started in the early Miocene (some 17–16 Ma) in the region where the current Namib Desert stands. This aridification not only persisted through time in this area but also intensified and resulted in the formation of the Kalahari Desert (Senut et al., 2009), effectively limiting range expansions in this direction (W-S), in the absence of successful colonization following LDD. Even in the case of genus *Colchicum* (**Figure S8**), were S African species appear closely related to NW African ones, W Mediterranean species are always sister to E Mediterranean ones. These leaves open the possibility of a colonization of NW Africa (from S Africa) via E Africa and W Mediterranean populations with subsequent extinction in E Africa. An alternative colonization from Central-West Asia into South Africa and NW Africa seems unlikely given the phylogeny of this genus (**Figure S8**), though proper biogeographic inference to test either possibility remains to be done. Indeed, Sanmartín et al. (2010) found a higher frequency of biotic exchange between NW-E African elements than with either E-S African or W-S African ones, where the latter elements were hardly connected, if at all, confirming our observations. We further argue that the magnitude of observed biotic exchange follows the history of desertification in Africa.

In all, the sequential timing of Neogene disjunctions in RF lineages, which is nonetheless concentrated in certain time intervals (e.g., Late Miocene-Pliocene), is in agreement with a scenario of range expansions (dispersal) in favorable times (windows of opportunity) and range contractions (extinction) as aridification flared up. Extinction results in absence (of a population, species, clade, or lineage) and thus leaves hard to track traces in phylogenies in the absence of fossil data (Meseguer et al., 2015). If repeated cycles of speciation, dispersal, and extinction take place in the same area over time, only taxa that optimize any (or a combination) of these processes (e.g., increased speciation, higher dispersal, lower extinction rates) will persist. It is to be expected that more recent populations, species, clades, or lineages show traces of these processes when compared to ancient ones.

On the other hand, our net diversification rate estimates (**Figure 5**) do no fully support an extinction explanation since, in the context of the family with the highest diversification rate among RF lineages, i.e., Asteraceae, most of the taxa fall inside the 95% CI under a no-extinction scenario ($\varepsilon = 0.0$). However, the method chosen to estimate net diversification rates (Magallón and Sanderson, 2001), though more appropriate given phylogeny size and sampling effort, is still limited. Crown diversification rates cannot be estimated for clades with 2 terminal taxa (see Materials and Methods), which is the case for several RF lineages (e.g., *Sideroxylon*). Additionally, the "method-of-moments" estimator performs well detecting declining diversity for old groups in exceedingly species-poor clades (Magallón and Sanderson, 2001; Warren and Hawkins, 2006) or young groups notably species-rich (recent radiations, Magallón and Sanderson, 2001), but we observed that statistical power is low to detect declines in diversity for young species-poor groups (e.g., *Camptoloma*). Most RF disjunct clades dated comprise less than 10 species—e.g., *Aeonium*, *Campanula*, *Camptoloma*, *Cicer*, *Colchicum*, *Euphorbia* sect. *Balsamis*, *Kleinia*, and *Plocama*—, limiting our ability to effectively detect the effects of extinction.

Nonetheless, if we focus on crown ages, disjunct clades in *Canarina, Hypericum*, and *Sideroxylon* are less diverse than expected, and given that their encompassing lineages (**Table 1**, **Figure 5**) also follow this trend, it would be safe to assume these lineages have indeed experienced high levels of extinction through time. Likewise, if we were to focus on stem ages, a few other groups fall below the no-extinction scenario ($\varepsilon = 0.0$), notably, *Camptoloma*, *Campylanthus*, and *Cicer*. Moreover, these groups exhibit wide-spanning (often >10 Ma) stem-crown intervals (see *Sideroxylon* or *Cicer* in **Figure 5**), an observation that has been tied to historically high extinction rates in recent diversification studies (Antonelli and Sanmartín, 2011; Nagalingum et al., 2011). This would further support the hypothesis that lower diversification rates in RF lineages could be explained in terms of increased extinction rather than a decrease in speciation rates.

Additionally, and given the aforementioned limitations of our diversification method of choice, it would also be safe to conclude that, within *Euphorbia,* sects. *Anthacanthae* (sect. *Balsamis* included), sect. *Esula*, and sect. *Aphyllis*, present higher diversity than expected (above the CI for $\varepsilon = 0.0$ in all cases, and also above the CI for $\varepsilon = 0.5$ for the former two clades), which is exceptional in the context of the genus, since *Euphorbia* is significantly poorer than expected for all ε values. Horn et al. (2014) also detected increased diversification rates in these sections of *Euphorbia*. Desertification-tropicalization cycles in Africa (Senut et al., 2009) suggest repeated reconnections between now disjunct RF regions since the Neogene, which would have permitted biotic exchange in favorable periods, whereas the isolation of these regions at unfavorable times would have induced speciation through vicariance, enhancing endemicity in these sub-humid/sub-xeric lineages. Molecular dating in tropical trees from the genus *Acridocapus* (Malpighiaceae; Davis et al., 2002) and the Annonaceae family (Couvreur et al., 2008) shows a similar pattern of connection phases between East African and Guineo-Congolian rainforest regions since the Oligocene following major climate shifts.

The youngest disjunctions, those of *Aeonium*, *Campanula*, and *Hypericum* sect. *Campylosporus,* are Pleistocene in age (**Figure 4**) and far too recent to result from the Neogene aridification of the African continent. Either rare LDD (i.e., *Aeonium*; Kim et al., 2008) or stepping-stone dispersal events (i.e., *Campanula*, Alarcón et al., pers. comm.), perhaps favored by Pleistocene cool and drier glacial cycles, could explain these more recent disjunct geographic patterns, as previously observed in other African taxa, e.g., *Convolvulus* (Carine, 2005), *Moraea* (Galley et al., 2007), or the tree heath (*Erica arborea*). Désamoré et al. (2011) took notice of successive range expansions of *Er. arborea* from an Eastern African center of diversity toward Northwest Africa, Southwest Europe, and Macaronesia, first during the Late Pliocene (ca. 3 Ma; **Figure 4**) and subsequently in the Pleistocene (ca. 1 Ma).

Redefining the Rand Flora Pattern

In a recent review, Linder (2014) synthesized the individual histories of numerous African lineages by recognizing five different "floras," which he defined as "groups of clades, which: (a) are largely found in the same area, (b) have largely the same extra-African geographical affinities, (c) share a diversification history, and (d) have a common maximum age." The "Rand Flora" does not fit well this definition. This *flora* does group a number of lineages that share the same geographic range (even if discontinuous), but they have slightly different climatic tolerances, i.e., sub-humid to sub-xeric or xerophilic, and they do not necessarily share the same extra-African geographical affinities.

Some RF lineages fall within what Linder (2014) terms "tropic-montane flora" (e.g., *Hypericum*, *Canarina*), others within the "arid flora" (e.g., *Kleinia*, *Campylanthus*). Some RF lineages are better connected with the Mediterranean Region (e.g., *Adenocarpus*), others with Asia and the Indo-Pacific Region (e.g., *Plocama*). Moreover, RF taxa on either side of any given disjunction (i.e., W-E or E-S) do no longer share a "diversification history," though they do share the same fate as other RF lineages with similar distribution. In fact, the different ages estimated here for the various RF disjunctions agree well with what has been termed biogeographic *pseudocongruence* (Donoghue and Moore, 2003), a phenomenon whereby two or more lineages display the same biogeographic pattern but with different temporal origins (Sanmartín, 2014). What is shared by all RF lineages is the nature of the climatic (ecological) barriers separating the taxa at either side of any given disjunction: arid regions such as the Sahara, the Kalahari or the Namib deserts, or the tropical lowlands in Central Africa. The congruence between RF disjunction ages and successive major climatic events in Africa during the Neogene (**Figure 4**) suggest that the ongoing aridification of the continent (or the "tropicalization" of Central Africa) affected RF lineages according to their different physiological (climatic) tolerances: more sub-humid lineages diverged first (e.g., *Sideroxylon*), more xeric later (e.g., *Campylanthus*).

One point of contention in the literature has been the limits of the Rand Flora with respect to the "Arid Corridor" or "Arid Track" (hereafter AC), a path repeatedly connecting south-west to north-east arid regions in Africa (and henceforth to central and southwest Asia) first proposed by Winterbottom (1967) and later expanded by de Winter (1966, 1971) and Verdcourt (1969). Bellstedt et al. (2012) defined the AC pattern as the disjunction occurring between Southern Africa and Eastern African-Southern Arabian xeric floristic elements. Linder (2014) considered the RF as an expansion of the AC to the west, in agreement with Jürgens' (1997) view. However, we consider that the RF and AC patterns are different. AC elements have more xeric preferences than the sub-humid to sub-xeric ones exhibited by RF elements. AC elements often extend into deserts (e.g., Namib, Kalahari, Sahara)—see studies by Beier et al. (2004) on *Fagonia* (Zygophyllaceae), Bellstedt et al. (2012) on *Zygophyllum* (also Zygophyllaceae), Carlson et al. (2012) on *Scabiosa* (Dipsacaceae), or Bruyns et al. (2014) on Ceropegieae— and have broader, more continuous distributions, plus they tend to be younger in age (often Pleistocene, coincident with Quaternary glaciation cycles). Our understanding is that this younger xeric AC elements move in parallel to RF taxa webbing with them in areas favorable to either, and thus confusing their limits. Something similar could have happened with Afromontane elements migrating south to north as the Eastern African mountains rose through the Mio-Pliocene; these elements are not part of the RF (e.g., *Iris*, *Moraea*, Galley et al., 2007).

In this study, we have provided a temporal framework for the Rand Flora pattern and estimated net diversification rates for 17 RF lineages. Our results provide some support to the historical view of an ancient African flora, whose current disjunct distribution was probably modeled by the successive waves of aridification events that have affected the African continent starting in the Miocene, but whose origin predates the latest events of Pleistocene climate change. These patterns were probably formed by a combination of climate-driven extinction and vicariance within a formerly widespread distribution. Whether these lineages all had a continuous, never interrupted, distribution that occupied all the area that now lies in between the extremes of the disjunction, or they had a somewhat narrower distribution in the past and they expanded their range tracking their habitat across the landscape in response to changing climate (e.g., along a corridor), is difficult to say with the current evidence. Discerning between these hypotheses will require the integration of phylogenetic, biogeographic and ecological approaches to reconstruct the ancestral ranges and climatic preferences of ancestral lineages (Mairal et al., 2015; Meseguer et al., 2015). Compared to speciation, extinction has received far less attention in studies focusing on the assembly of tropical biotas. Disentangling extinction from other processes is particularly difficult because the biodiversity we observe today is only a small fraction of that of the past. The Rand Flora pattern might offer a prime study model to understand the effects of climate-driven extinction in the shaping of continent-wide biodiversity patterns.

Author Contributions

IS and LP conceived and designed the study. LP analyzed the data with help from IS, RR, and MM. LP and IS co-wrote the text, with contributions from MH, RR, MM, and AM. All authors contributed with data compilation, figure preparation, or text comments. MM has copyright of all plant pictures, except for *Cicer canariense*.

Acknowledgments

This study was funded by the Spanish Ministry of Economy and Competitiveness (MINECO): Project AFFLORA, CGL2012-40129-C02-01 to IS. LP was funded by CSIC postdoctoral contract within AFFLORA. MH was funded by CGL2012-40129-C02-02, the Research Council of Norway (203822/E40) and a Ramón y Cajal Fellowship (RYC2009-04537). RR was supported by a JAE-DOC postdoctoral fellowship (MINECO) and the European Social Fund. MM and VC were supported by MINECO FPI predoctoral fellowships (BES-2010-037261 and BES-2013-065389 respectively). We thank Virginia Valcárcel (Department of Biology, UAM, Spain) for help with data compilation and literature revision during the earlier stages of the project, Andrea Briega (Department of Ecology, UAH, Spain) for help with data compilation, and Manuel Gil for providing a *Cicer canariense* picture.

Supplementary Material

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fgene.2015.00154/abstract

Supplementary Materials include descriptions of study groups with references, **Table S1**, and **Figures S1–S17**.

Table S1 | Net diversification rates (bd.ms) for all RF disjunct clades and their encompassing lineages (bold = highest crown.p, red when $n \leq 2$) under three possible scenarios: no extinction ($\varepsilon = 0$), turnover at equilibrium ($\varepsilon = 0.5$), and high extinction ($\varepsilon = 0.9$). Probability (crown.p) of obtaining a clade with the same size and age as the RF disjunction, given the background diversification rate of the encompassing clade/s and at increasing extinction fractions (bold = highest crown.p, italics $p < 0.05$). Stem and Crown ages in Myr.

Figures S1–S17 | BEAST MCC chronograms showing mean estimates and 95% high posterior density (HPD) confidence intervals for those nodes receiving 50% support. Branch width is proportional to PP support. Red colored taxa indicate Eastern African provenance; Macaronesia/western African taxa and southern African taxa are colored in blue and green, respectively. Calibration points are indicated with stars; RF disjunctions within each lineage discussed in the text and represented in **Figures 3**–**5** are indicated with arrows.

References

Alarcón, M., Roquet, C., García-Fernández, A., Vargas, P., and Aldasoro, J. J. (2013). Phylogenetic and phylogeographic evidence for a Pleistocene disjunction between *Campanula jacobaea* (Cape Verde Islands) and *C. balfourii (Socotra). Mol. Phylogenet. Evol.* 69, 828–836. doi: 10.1016/j.ympev.2013.06.021

Andrus, N., Trusty, J., Santos-Guerra, A., Jansen, R. K., and Francisco-Ortega, J. (2004). Using molecular phylogenies to test phytogeographical links between East/South Africa–Southern Arabia and the Macaronesian islands—a review, and the case of *Vierea* and *Pulicaria* section Vieraeopsis (Asteraceae). *Taxon* 53, 333–333. doi: 10.2307/4135612

Antonelli, A., and Sanmartín, I. (2011). Mass extinction, gradual cooling, or rapid radiation? Reconstructing the spatiotemporal evolution of the ancient angiosperm genus *Hedyosmum* (Chloranthaceae) using empirical and simulated approaches. *Syst. Biol.* 60, 596–615. doi: 10.1093/sysbio/syr062

Axelrod, D. I., and Raven, P. H. (1978). "Late Cretaceous and Tertiary vegetation history of Africa," in *Biogeography and Ecology of southern Africa*, ed M. J. A. Werger (The Hague: Dr W. Junk bv Publishers), 77–130.

Backlund, M., Bremer, B., and Thulin, M. (2007). Paraphyly of Paederieae, recognition of Putorieae and expansion of *Plocama* (Rubiaceae-Rubioideae). *Taxon* 56, 315–328. Available online at: http://www.ingentaconnect.com/content/iapt/tax/2007/00000056/00000002/art00006

Barreda, V. D., Palazzesi, L., Katinas, L., Crisci, J. V., Tellería, M. C., Bremer, K., et al. (2012). An extinct Eocene taxon of the daisy family (Asteraceae): evolutionary, ecological and biogeographical implications. *Ann. Bot.* 109, 127–134. doi: 10.1093/aob/mcr240

Barres, L., Sanmartín, I., Anderson, C. L., Susanna, A., Buerki, S., Galbany-Casals, M., et al. (2013). Reconstructing the evolution and biogeographic history of tribe *Cardueae* (Compositae). *Am. J. Bot.* 100, 867–882. doi: 10.3732/ajb.1200058

Barres, L., Vilatersana, R., Molero, J., Susanna, A., and Galbany-Casals, M. (2011). Molecular phylogeny of *Euphorbia* subg. *Esula sect. Aphyllis* (Euphorbiaceae) inferred from nrDNA and cpDNA markers with biogeographic insights. *Taxon* 60, 705–720. Available online at: http://www.ingentaconnect.com/content/iapt/tax/2011/00000060/00000003/art00007

Beerling, D. J., Taylor, L. L., Bradshaw, C. D., Lunt, D. J., Valdes, P. J., Banwart, S. A., et al. (2012). Ecosystem CO2 starvation and terrestrial silicate weathering: mechanisms and global-scale quantification during the late Miocene. *J. Ecol.* 100, 31–41. doi: 10.1111/j.1365-2745.2011.01905.x

Beier, B. A., Nylander, J. A. A., Chase, M. W., and Thulin, M. (2004). Phylogenetic relationships and biogeography of the desert plant genus *Fagonia* (Zygophyllaceae), inferred by parsimony and Bayesian model averaging. *Mol. Phylogenet. Evol.* 33, 91–108. doi: 10.1016/j.ympev.2004.05.010

Bell, C. D., Soltis, D. E., and Soltis, P. S. (2010). The age and diversification of the angiosperms re-revisited. *Am. J. Bot.* 97, 1296–1303. doi: 10.3732/ajb.0900346

Bellstedt, D. U., Galley, C., Pirie, M. D., and Linder, H. P. (2012). The migration of the palaeotropical arid flora: zygophylloideae as an example. *Syst. Bot.* 37, 951–959. doi: 10.1600/036364412X656608

Böhme, M., Ilg, A., and Winklhofer, M. (2008). Late Miocene "washhouse" climate in Europe. *Earth Planet. Sci. Lett.* 275, 393–401. doi: 10.1016/j.epsl.2008.09.011

Bonnefille, R. (2011). "Rainforest responses to past climate changes in tropical Africa," in *Tropical Rainforest Responses to Climate Change, 2nd Edn*, eds M. Bush, J. Flenley, and W. Gosling (Berlin; Heidelberg: Springer-Verlag), 125–184.

Bramwell, D. (1985). Contribución a la biogeografía de las Islas Canarias. *Bot. Macaronésica* 14, 3–34.

Bremer, B., and Eriksson, T. (2009). Time tree of Rubiaceae: phylogeny and dating the family, subfamilies, and tribes. *Int. J. Plant Sci.* 170, 766–793. doi: 10.1086/599077

Bruyns, P. V., Klak, C., and Hanáček, P. (2014). Evolution of the stapeliads (Apocynaceae–Asclepiadoideae)—repeated major radiation across Africa in an Old World group. *Mol. Phylogenet. Evol.* 77, 251–263. doi: 10.1016/j.ympev.2014.03.022

Carine, M. A. (2005). Spatio-temporal relationships of the Macaronesian endemic flora: a relictual series or window of opportunity? *Taxon* 54, 895–903. doi: 10.2307/25065476

Carlson, S. E., Linder, H. P., and Donoghue, M. J. (2012). The historical biogeography of *Scabiosa* (Dipsacaceae): implications for Old World plant disjunctions. *J. Biogeogr.* 39, 1086–1100. doi: 10.1111/j.1365-2699.2011.02669.x

Chacón, J., and Renner, S. S. (2014). Assessing model sensitivity in ancestral area reconstruction using LAGRANGE: a case study using the Colchicaceae family. *J. Biogeogr.* 41, 1414–1427. doi: 10.1111/jbi.12301

Christ, H. (1892). Exposé sur le rôle que joue dans le domaine de nos flores la flore dite *ancienne africaine*. *Arch. Sci. Phys. Nat. Genève* 3, 369–374.

Christ, H. (1910). *Die Geographie der Farne*. Jena: Verlag von Gustav Fischer. 1–358.

Coleman, M., Liston, A., Kadereit, J. W., and Abbott, R. J. (2003). Repeat intercontinental dispersal and Pleistocene speciation in disjunct Mediterranean and desert *Senecio* (Asteraceae). *Am. J. Bot.* 90, 1446–1454. doi: 10.3732/ajb.90.10.1446

Couvreur, T. L. (2015). Odd man out: why are there fewer plant species in African rain forests? *Plant Syst. Evol.* 301, 1299–1313. doi: 10.1007/s00606-014-1180-z

Couvreur, T. L., Chatrou, L. W., Sosef, M. S., and Richardson, J. E. (2008). Molecular phylogenetics reveal multiple tertiary vicariance origins of the African rain forest trees. *BMC Biol.* 6:54. doi: 10.1186/1741-7007-6-54

Cubas, P., Pardo, C., Tahiri, H., and Castroviejo, S. (2010). Phylogeny and evolutionary diversification of *Adenocarpus* DC. *(Leguminosae). Taxon* 59, 720–732. Available online at: http://www.ingentaconnect.com/content/iapt/tax/2010/00000059/00000003/art00005

Davis, C. C., Bell, C. D., Fritsch, P. W., and Mathews, S. (2002). Phylogeny of *Acridocarpus-Brachylophon* (Malpighiaceae): implications for Tertiary tropical floras and Afroasian biogeography. *Evolution* 56, 2395–2405. doi: 10.1111/j.0014-3820.2002.tb00165.x

del Hoyo, A., García-Marín, J. L., and Pedrola-Monfort, J. (2009). Temporal and spatial diversification of the African disjunct genus *Androcymbium* (Colchicaceae). *Mol. Phylogenet. Evol.* 53, 848–861. doi: 10.1016/j.ympev.2009.08.005

Désamoré, A., Laenen, B., Devos, N., Popp, M., González-Mancebo, J. M., Carine, M. A., et al. (2011). Out of Africa: north-westwards Pleistocene expansions of the heather *Erica arborea*. *J. Biogeogr.* 38, 164–176. doi: 10.1111/j.1365-2699.2010.02387.x

de Winter, B. (1966). Remarks on the distribution of some desert plants in Africa. *Palaeoecol. Afr.* 1, 188–189.

de Winter, B. (1971). Floristic relationships between the northern and southern arid areas in Africa. *Mitt. Bot. Staatssamml. Munch.* 10, 424–437.

Doláková, N., Kovaìcovaì, M., and Basistovaì, P. (2011). Badenian (Langhian-Early Serravallian) palynoflora for the Carpathian Foredeep and Vienna Basin (Czech and Slovak Republics). *Acta Ent. Mus. Nat. Pra.* 67, 63–71. Available online at: http://www.muni.cz/research/publications/955017

Donoghue, M. J., and Moore, B. R. (2003). Toward an integrative historical biogeography. *Integr. Comp. Biol.* 43, 261–270. doi: 10.1093/icb/43.2.261

Drummond, A. J., Ho, S. Y., Phillips, M. J., and Rambaut, A. (2006). Relaxed phylogenetics and dating with confidence. *PLoS Biol.* 4:e88. doi: 10.1371/journal.pbio.0040088

Drummond, A. J., Suchard, M. A., Xie, D., and Rambaut, A. (2012). Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol. Biol. Evol.* 29, 1969–1973. doi: 10.1093/molbev/mss075

Engler, A. (1879). "Die extratropischen Gebiete der nördlichen Hemisphäre. Mit einer chromolithographischen Karte," in *Versuch einer Entwicklungsgeschichte der Pflanzenwelt: insbesondere der Florengebiete seit der Tertiärperiode. Theil 1*, ed A. Engler (Leipzig: Verlag von Wilhelm Engelmann), 1–202.

Engler, A. (1910). "Die Pflanzenwelt Afrikas inbesondere seiner tropischen Gebiete. Gründzuge der Pflanzenverbreitung in Afrika un die Characterpflanzen Afrikas," in *Die Vegetation der Erde*, eds A. Engler and O. Drude (Leipzig: Verlag von Wilhelm Engelmann), 1030.

Fischer, A. G. (1960). Latitudinal variations in organic diversity. *Evolution* 14, 64–81. doi: 10.2307/2405923

Fiz, O., Vargas, P., Alarcón, M., Aedo, C., García, J. L., and Aldasoro, J. J. (2008). Phylogeny and historical biogeography of Geraniaceae in relation to climate changes and pollination ecology. *Syst. Bot.* 33, 326–342. doi: 10.1600/036364408784571482

Forster, J. R. (1778). *Observations Made During a Voyage Round the World, on Physical Geography, Natural History, and Ethic Philosophy*. London: G. Robinson.

Galley, C., Bytebier, B., Bellstedt, D. U., and Linder, H. P. (2007). The Cape element in the Afrotemperate flora: from Cape to Cairo? *Proc. R. Soc. B* 274, 535–543. doi: 10.1098/rspb.2006.0046

Harmon, L. J., Weir, J. T., Brock, C. D., Glor, R. E., and Challenger, W. (2008). GEIGER: investigating evolutionary radiations. *Bioinformatics* 24, 129–131. doi: 10.1093/bioinformatics/btm538

Ho, S. Y., and Phillips, M. J. (2009). Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times. *Syst. Biol.* 58, 367–380. doi: 10.1093/sysbio/syp035

Horn, J. W., Xi, Z., Riina, R., Peirson, J. A., Yang, Y., Dorsey, B. L., et al. (2014). Evolutionary burst in *Euphorbia* (Euphorbiaceae) are linked with Photosynthetic pathway. *Evolution* 68, 3485–3504. doi: 10.1111/evo.12534

Jablonski, D., Roy, K., and Valentine, J. W. (2006). Out of the tropics: evolutionary dynamics of the latitudinal diversity gradient. *Science* 314, 102–106. doi: 10.1126/science.1130880

Jacobs, B. F., Pan, A. D., and Scotese, C. R. (2010). "A review of the Cenozoic vegetation history of Africa," in *Cenozoic Mammals of Africa*, eds L. Werdelin and W. J. Sanders (Berkeley, CA: University of California Press), 57–72.

Javadi, F., Wojciechowski, M. F., and Yamaguchi, H. (2007). Geographical diversification of the genus *Cicer* (Leguminosae: Papilionoideae) inferred from molecular phylogenetic analyses of chloroplast and nuclear DNA sequences. *Bot. J. Linn. Soc.* 154, 175–186. doi: 10.1111/j.1095-8339.2007.00649.x

Jürgens, N. (1997). Floristic biodiversity and history of African arid regions. *Biodiv. Conserv.* 6, 495–514. doi: 10.1023/A:1018325026863

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:e2139. doi: 10.1371/journal.pone.0002139

Kornhall, P., Heidari, N., and Bremer, B. (2001). Selagineae and Manuleeae, two tribes or one? Phylogenetic studies in the Scrophulariaceae. *Plant Syst. Evol.* 228, 199–218. doi: 10.1007/s006060170029

Lavin, M., Herendeen, P. S., and Wojciechowski, M. F. (2005). Evolutionary rates analysis of Leguminosae implicates a rapid diversification of lineages during the Tertiary. *Syst. Biol.* 54, 575–594. doi: 10.1080/10635150590947131

Lavin, M., Wojciechowski, M. F., Richman, A., Rotella, J., Sanderson, M. J., and Matos, A. B. (2001). Identifying Tertiary radiations of Fabaceae in the Greater Antilles: alternatives to cladistic vicariance analysis. *Int. J. Plant Sci.* 162, S53–S76. doi: 10.1086/323474

Lebrun, J. (1947). "Essai sur l'origine et le développement de la flore," in *Exploration du Parc National Albert. Mission J. Lebrun (1937–1938). La végétation de la plaine alluviale au sud du Lac Édouard*, ed J. Lebrun (Brussels: Inst. des Parcs Nationaux du Congo Belge). Fasc. 1. 2ème Part., 115–397.

Lebrun, J. (1961). Les deux flores d'Afrique tropicale. *Acad. Roy. Belg. Cl. Sci. Mém. (coll. 8. 2^{éme} sér.)* 32, 1–82.

Linder, H. P. (2014). The evolution of African plant diversity. *Front. Ecol. Evol.* 2, 1–14. doi: 10.3389/fevo.2014.00038

Lomolino, M. V., Riddle, B. R., Whittaker, R. J., and Brown, J. H. (2010). *Biogeography. 4th Edn*. Sunderland, MA: Sinauer Associates, Inc. 764.

Maclean, I. M. D., and Wilson, R. J. (2011). Recent ecological responses to climate change support predictions of high extinction risk. *Proc. Natl. Acad. Sci. U.S.A.* 108, 12337–12342. doi: 10.1073/pnas.1017352108

Maddison, W. P., and Maddison, D. R. (2014). *Mesquite: A Modular System for Evolutionary Analysis. Version 3.01*. Available online at: http://mesquiteproject.org

Magallón, S., and Sanderson, M. J. (2001). Absolute diversification rates in angiosperm clades. *Evolution* 55, 1762–1780. doi: 10.1111/j.0014-3820.2001.tb00826.x

Mairal, M., Pokorny, L., Aldasoro, J. J., Alarcón, M., and Sanmartín, I. (2015). Ancient vicariance and climate-driven extinction explain continental-wide disjunctions in Africa: the case of the Rand Flora genus *Canarina* (Campanulaceae). *Mol. Ecol.* 24, 1335–1354. doi: 10.1111/mec.13114

Maley, J. (1996). The African rain forest—main characteristics of changes in vegetation and climate from the Upper Cretaceous to the Quaternary. *Proc. Roy. Soc. Edinburgh Sect. B. Biol. Sci.* 104, 31–73. doi: 10.1017/S0269727000006114

Maley, J. (2000). Last Glacial Maximum lacustrine and fluviatile formations in the Tibesti and other Saharan mountains, and large-scale climatic teleconnections linked to the activity of the Subtropical Jet Stream. *Glob. Planet. Change* 26, 121–136. doi: 10.1016/S0921-8181(00)00039-4

Manning, J., Forest, F., and Vinnersten, A. (2007). The genus *Colchicum* L. redefined to include Androcymbium Willd. based on molecular evidence. *Taxon* 56, 872–882. doi: 10.2307/25065868

Meseguer, A. S., Aldasoro, J. J., and Sanmartín, I. (2013). Bayesian inference of phylogeny, morphology and range evolution reveals a complex evolutionary history in St. John's wort (Hypericum). *Mol. Phylogenet. Evol.* 67, 379–403. doi: 10.1016/j.ympev.2013.02.007

Meseguer, A. S., Lobo, J. M., Ree, R., Beerling, D. J., and Sanmartín, I. (2015). Integrating Fossils, Phylogenies, and Niche Models into Biogeography to reveal ancient evolutionary history: the Case of *Hypericum* (Hypericaceae). *Syst. Biol.* 64, 215–232. doi: 10.1093/sysbio/syu088

Monod, T. (1971). Remarques sur les symétries floristiques des zones sèches nord et sud en Afrique. *Mitt. Bot. Staatssamml. München* 10, 375–423.

Morley, R. J. (2000). *Origin and Evolution of Tropical Rain Forests*. Chinchester: John Wiley and Sons Ltd. 1–362.

Morley, R. J. (2003). Interplate dispersal paths for megathermal angiosperms. *Perspect. Plant Ecol. Evol. Syst.* 6, 5–20. doi: 10.1078/1433-8319-00039

Morlon, H. (2014). Phylogenetic approaches for studying diversification. *Ecol. Lett.* 17, 508–525. doi: 10.1111/ele.12251

Mort, M. E., Soltis, D. E., Soltis, P. S., Francisco-Ortega, J., and Santos-Guerra, A. (2002). Phylogenetics and evolution of the Macaronesian clade of Crassulaceae inferred from nuclear and chloroplast sequence data. *Syst. Bot.* 27, 271–288. doi: 10.1043/0363-6445-27.2.271

Mort, M. E., Soltis, D. E., Soltis, P. S., Santos-Guerra, A., and Francisco-Ortega, J. (2007). Physiological evolution and association between physiology and growth form in *Aeonium* (Crassulaceae). *Taxon* 56, 453–464.

Nagalingum, N. S., Marshall, C. R., Quental, T. B., Rai, H. S., Little, D. P., and Mathews, S. (2011). Recent synchronous radiation of a living fossil. *Science* 334, 796–799. doi: 10.1126/science.1209926

Nagy, E. (1963). Some new spore and pollen species from the Neogene of the Mecsek Mountain. *Acta Bot. Hung.* 9, 387–404.

Nathan, R. (2006). Long-distance dispersal of plants. *Science* 313, 786–788. doi: 10.1126/science.1124975

Navarro-Pérez, M. L., López, J., Fernández-Mazuecos, M., Rodríguez-Riaño, T., Vargas, P., and Ortega-Olivencia, A. (2013). The role of birds and insects in pollination shifts of *Scrophularia* (Scrophulariaceae). *Mol. Phylogenet. Evol.* 69, 239–254. doi: 10.1016/j.ympev.2013.05.027

Nylander, J. A. A. (2004). *MrModeltest v2.2 Program Distributed by the Author*. Uppsala: Evolutionary biology centre, Uppsala University, 2.

Osborne, A. H., Vance, D., Rohling, E. J., Barton, N., Rogerson, M., and Fello, N. (2008). A humid corridor across the Sahara for the migration of early modern humans out of Africa 120,000 years ago. *Proc. Natl. Acad. Sci. U.S.A.* 105, 16444–16447. doi: 10.1073/pnas.0804472105

Oxelman, B., Kornhall, P., Olmstead, R. G., and Bremer, B. (2005). Further disintegration of Scrophulariaceae. *Taxon* 54, 411–425. doi: 10.2307/25065369

Peirson, J. A., Bruyns, P. V., Riina, R., Morawetz, J. J., and Berry, P. E. (2013). A molecular phylogeny and classification of the largely succulent and mainly African *Euphorbia* subg. *Athymalus (Euphorbiaceae). Taxon* 62, 1178–1199. doi: 10.12705/626.12

Pelser, P. B., Abbott, R. J., Comes, H. P., Milton, J. J., Moeller, M., Looseley, M. E., et al. (2012). The genetic ghost of an invasion past: colonization and extinction revealed by historical hybridization in *Senecio*. *Mol. Ecol.* 21, 369–387. doi: 10.1111/j.1365-294X.2011.05399.x

Pelser, P. B., Kennedy, A. H., Tepe, E. J., Shidler, J. B., Nordenstam, B., Kadereit, J. W., et al. (2010). Patterns and causes of incongruence between plastid and nuclear Senecioneae (Asteraceae) phylogenies. *Am. J. Bot.* 97, 856–873. doi: 10.3732/ajb.0900287

Pelser, P. B., Nordenstam, B., Kadereit, J. W., and Watson, L. E. (2007). An ITS phylogeny of tribe Senecioneae (Asteraceae) and a new delimitation of Senecio L. *Taxon* 56, 1077–1077. doi: 10.2307/25065905

Pennington, R. T., Cronk, Q. C., and Richardson, J. A. (2004). Introduction and synthesis: plant phylogeny and the origin of major biomes. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 359, 1455–1464. doi: 10.1098/rstb.2004.1539

Pennington, R. T., Lavin, M., Särkinen, T., Lewis, G. P., Klitgaard, B. B., and Hughes, C. E. (2010). Contrasting plant diversification histories within the Andean biodiversity hotspot. *Proc. Natl. Acad. Sci. U.S.A.* 107, 13783–13787. doi: 10.1073/pnas.1001317107

Plana, V. (2004). Mechanisms and tempo of evolution in the African Guineo–Congolian rainforest. *Philos. Trans. R. Soc. Lond. B, Biol. Sci.* 359, 1585–1594. doi: 10.1098/rstb.2004.1535

Qian, H., and Ricklefs, R. E. (2000). Large-scale processes and the Asian bias in species diversity of temperate plants. *Nature* 407, 180–182. doi: 10.1038/35025052

Quézel, P. (1978). Analysis of the flora of mediterranean and saharan africa. *Ann. Missouri Bot. Gard.* 65, 479–534. doi: 10.2307/2398860

Rabosky, D. L., Grundler, M., Anderson, C., Shi, J. J., Brown, J. W., Huang, H., et al. (2014). BAMMtools: an R package for the analysis of evolutionary dynamics on phylogenetic trees. *Methods Ecol. Evol.* 5, 701–707. doi: 10.1111/2041-210X.12199

Rambaut, A. (2002). *Se-Al, Version 2.0 a11. Computer Program Distributed by the Author.* Available online at: http://tree.bio.ed.ac.uk/software/seal/

Rambaut, A. (2009). *FigTree, Version 1.4.2. Computer Program Distributed by the Author.* http://tree.bio.ed.ac.uk/software/figtree/

Rambaut, A., Suchard, M. A., Xie, W., and Drummond, A.J. (2013). *Tracer v1.6*. Available online at: http://tree.bio.ed.ac.uk/soft-ware/tracer

Raven, P. H., and Axelrod, D. I. (1974). Angiosperm biogeography and past continental movements. *Ann. Missouri Bot. Gard.* 61, 539–673. doi: 10.2307/2395021

Richards, P. W. (1973). "Africa, the 'Odd man out,'" in *Tropical Forest Ecosystems of Africa and South America: A Comparative Review*, eds B. J. Meggers, E. S. Ayensu, and W. D. Duckworth (Washington, DC: Smithsonian Institution Press), 21–26.

Ricklefs, R. E. (2007). Estimating diversification rates from phylogenetic information. *Trends Ecol. Evol.* 22, 601–610. doi: 10.1016/j.tree.2007.06.013

Riina, R., Peirson, J. A., Geltman, D. V., Molero, J., Frajman, B., Pahlevani, A., et al. (2013). A worldwide molecular phylogeny and classification of the leafy spurges, *Euphorbia* subgenus *Esula* (Euphorbiaceae). *Taxon* 62, 316–342. doi: 10.12705/622.3

Roquet, C., Sanmartín, I., García-Jacas, N., Sáez, L., Susanna, A., Wikström, N., et al. (2009). Reconstructing the history of Campanulaceae with a Bayesian approach to molecular dating and dispersal–vicariance analyses. *Mol. Phylogenet. Evol.* 52, 575–587. doi: 10.1016/j.ympev.2009.05.014

Sanmartín, I. (2014). "Biogeography," in *The Tree of Life*, eds P. Vargas and R. Zardoya (Sunderland, MA: Sinauer Associates, Inc.), 555–576.

Sanmartín, I., Anderson, C. L., Alarcon, M., Ronquist, F., and Aldasoro, J. J. (2010). Bayesian island biogeography in a continental setting: the Rand Flora case. *Biol. Lett.* 6, 703–707. doi: 10.1098/rsbl.2010.0095

Senut, B., Pickford, M., and Ségalen, L. (2009). Neogene desertification of Africa. *C. R. Geoscience* 341, 591–602. doi: 10.1016/j.crte.2009.03.008

Sepulchre, P., Ramstein, G., Fluteau, F., Schuster, M., Tiercelin, J. J., and Brunet, M. (2006). Tectonic uplift and Eastern Africa aridification. *Science* 313, 1419–1423. doi: 10.1126/science.1129158

Simmons, M. P. (2004). Independence of alignment and tree search. *Mol. Phylogenet. Evol.* 31, 874–879. doi: 10.1016/j.ympev.2003.10.008

Smedmark, J. E., and Anderberg, A. A. (2007). Boreotropical migration explains hybridization between geographically distant lineages in the pantropical clade Sideroxyleae (Sapotaceae). *Am. J. Bot.* 94, 1491–1505. doi: 10.3732/ajb.94.9.1491

Smedmark, J. E., Swenson, U., and Anderberg, A. A. (2006). Accounting for variation of substitution rates through time in Bayesian phylogeny reconstruction of Sapotoideae (Sapotaceae). *Mol. Phylogenet. Evol.* 39, 706–721. doi: 10.1016/j.ympev.2006.01.018

Stadler, T. (2013). Recovering speciation and extinction dynamics based on phylogenies. *J. Evol. Biol.* 26, 1203–1219. doi: 10.1111/jeb.12139

Stevens, G. C. (1989). The latitudinal gradient in geographical range: how so many species coexist in the tropics. *Am. Nat.* 133, 240–256. doi: 10.1086/284913

Stride, G., Nylinder, S., and Swenson, U. (2014). Revisiting the biogeography of *Sideroxylon* (Sapotaceae) and an evaluation of the taxonomic status of *Argania* and *Spiniluma*. *Austral. Syst. Bot.* 27, 104–118. doi: 10.1071/SB14010

Swofford, D. (2002). *Phylogenetic Analysis Using Parsimony (PAUP* v4.0b)*. Sunderland, MA: Sinauer Associates.

Thiv, M., Thulin, M., Hjertson, M., Kropf, M., and Linder, H. P. (2010). Evidence for a vicariant origin of Macaronesian–Eritreo/Arabian disjunctions in *Campylanthus* Roth (Plantaginaceae). *Mol. Phylogenet. Evol.* 54, 607–616. doi: 10.1016/j.ympev.2009.10.009

Trauth, M. H., Larrasoaña, J. C., and Mudelsee, M. (2009). Trends, rhythms and events in Plio-Pleistocene African climate. *Quat. Sci. Rev.* 28, 399–411. doi: 10.1016/j.quascirev.2008.11.003

Van Campo, E. (1989). Flore pollinique du Miocene superieur de Venta del Moro (Espagne). *Acta Palynol.* 1, 9–32.

van Zinderen Bakker, E. M. Sr. (1978). "Quaternary vegetation changes in southern Africa," in *Biogeography and Ecology of southern Africa*, ed M. J. A. Werger (The Hague: Dr W. Junk bv Publishers), 131–143.

Verdcourt, B. (1969). The arid corridor between the northeast and southwest areas of Africa. *Palaeoecol. Afr.* 4, 140–144.

von Humboldt, A., and Bonpland, A. (1805). *Essai sur la Geìographie des Plantes; Accompagneì d'un Tableau Physique des Reìgions Eìquinoxiales, Fondeì sur des Mesures exeìcuteìes, Depuis le DixieÌme Degreì de Latitude Boreìale Jusqu'au DixieÌme Degreì de Latitude Australe, Pendant les Anneìes 1799, 1800, 1801, 1802 et 1803*. Paris: chez Levrault, Schoell et compagnie, libraries, 155.

Wallace, A. R. (1878). *Tropical Nature, and Other Essays*. London: Macmillan and co. 356.

Warren, B. H., and Hawkins, J. A. (2006). The distribution of species diversity across a flora's component lineages: dating the Cape's 'relicts'. *Proc. R. Soc. B* 273, 2149–2158. doi: 10.1098/rspb.2006.3560

Wheeler, T. J., and Kececioglu, J. D. (2007). Multiple alignment by aligning alignments. *Bioinformatics* 23, i559–i568. doi: 10.1093/bioinformatics/btm226

Wiens, J. J. (2004). What is speciation and how should we study it? *Am. Nat.* 163, 914–922. doi: 10.1086/386552

Wiens, J. J. (2011). The causes of species richness patterns across space, time, and clades and the role of "ecological limits." *Q. Rev. Biol.* 86, 75–96. doi: 10.1086/659883

Wiens, J. J., and Donoghue, M. J. (2004). Historical biogeography, ecology and species richness. *Trends Ecol. Evol.* 19, 639–644. doi: 10.1016/j.tree.2004.09.011

Winterbottom, J. M. (1967). Climatological implications of avifaunal resemblances between South Western Africa and Somaliland. *Palaeoecol. Afr.* 2, 77–79.

Xie, L., Yang, Z. Y., Wen, J., Li, D. Z., and Yi, T. S. (2014). Biogeographic history of *Pistacia* (Anacardiaceae), emphasizing the evolution of the Madrean-Tethyan and the eastern Asian-Tethyan disjunctions. *Mol. Phylogenet. Evol.* 77, 136–146. doi: 10.1016/j.ympev.2014.04.006

Zachos, J. C., Dickens, G. R., and Zeebe, R. E. (2008). An early Cenozoic perspective on greenhouse warming and carbon-cycle dynamics. *Nature* 451, 279–283. doi: 10.1038/nature06588

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

CHAPTER 6

Lineage-specific climatic niche drives the tempo of vicariance in the Rand Flora

Mario Mairal, Isabel Sanmartín, Loïc Pellissier

This chapter has been has been published in the peer-review journal *Journal of Biogeography*. DOI: 10.1111/jbi.12930

Journal of Biogeography (J. Biogeogr.) (2017)

Journal of Biogeography

ORIGINAL ARTICLE

Lineage-specific climatic niche drives the tempo of vicariance in the Rand Flora

Mario Mairal[1]*, Isabel Sanmartín[1]† and Loïc Pellissier[2,3]†

[1]Department of Biodiversity and Conservation, Real Jardín Botánico (RJB-CSIC), Plaza de Murillo 2, 28014 Madrid, Spain, [2]Landscape Ecology, Institute of Terrestrial Ecosystems, ETH Zürich, Zürich, Switzerland, [3]Swiss Federal Research Institute WSL, 8903 Birmensdorf, Switzerland

ABSTRACT

Aim The disjunct distribution patterns of sister taxa can arise when previously continuous distribution ranges are fragmented by environmental changes such as major climatic events. Populations become isolated on either side of the newly established environmental barrier, and absence of gene flow promotes allopatric speciation, in a process that is known as ecological vicariance. If climate change altered the ancestral range gradually, such as along temporal temperature or moisture gradients, the age of divergence of disjunct species should be related to the lineage tolerance to climatic conditions. Here, we investigate this hypothesis using as a study model the African Rand Flora, a continental-scale floristic pattern that relates sister taxa distributed on either side of the Saharan Desert.

Location Africa, Macaronesia, Mediterranean Basin and the Middle East.

Methods We estimated the extant climatic tolerances of 14 Rand Flora lineages based on present occurrence data, and correlated the phylogenetic age of divergence between vicariant clades. We tested whether the tempo of the vicariance in the Rand Flora lineages was associated with the average values of their climatic niches in agreement with niche-driven divergence. We hindcasted species ranges using species distribution models combined with palaeoclimate simulations to infer the potential distribution of each lineage's ancestors.

Results We found a positive relationship between the lineage temperature niche and the age of the Rand Flora disjunction: lineages with subtropical affinities diverged first, whereas those with a higher tolerance to drier conditions (temperate or sub-xeric adaptations) exhibited younger disjunctions. The range reconstructions showed the existence of climatic corridors south of the Sahara in the wetter Late Miocene, which became interrupted during the mid-Pliocene warming event.

Main conclusions Our results suggest that climate change leading to the formation of the Sahara Desert drove Rand Flora lineages divergences along a temporal sequence that matched the climatic niche of species.

Keywords

continental disjunctions, extinction, niche conservatism, Rand Flora, refuges, vicariance

*Correspondence: Mario Mairal, Real Jardín Botánico (RJB-CSIC), 28014 Madrid, Spain.
E-mail: mariomairal@gmail.com
†Equal contributions.

INTRODUCTION

The concept of vicariance dates back to the origins of modern historical biogeography (Croizat, 1952; Brundin, 1967): vicariance arises when the ancestral range of a species becomes divided by the appearance of a new biogeographical barrier, followed by allopatric speciation (Humphries & Parenti, 1999; Sanmartín, 2012). Although vicariance has often been associated with geological events, such as the formation of a mountain chain or an ocean basin between two tectonic

http://wileyonlinelibrary.com/journal/jbi
doi:10.1111/jbi.12930

plates, this process refers to any environmental change that impacts a species geographical range. For example, a dramatic event of climate change might create new environmentally unsuitable area for a given species; individuals are unable to persist or disperse through these areas of unsuitable conditions, and the species distribution becomes fragmented (Wiens, 2004). The concept of ecological vicariance is strongly associated with the idea of niche conservatism, the tendency of species to retain their ancestral environmental preferences (Harvey & Pagel, 1991; Peterson *et al.*, 1999; Wiens & Graham, 2005; Wiens *et al.*, 2010; Crisp & Cook, 2012). In ecological vicariance, the rate of environmental change is faster than the species rate of adaptation, compelling the species to migrate or persist in residual suitable areas (Wiens, 2004). Vicariance is often associated with a general pattern of geographical fragmentation that is not unique to a single species but shared across unrelated lineages (Sanmartín, 2012). For ecological vicariance, this is the equivalent of an abrupt change in climate so rapid, intense and spatially widespread, that its effects are felt clade-wide across different lineages that present similar ecological affinities within the same geographical region.

Disjunct distribution patterns may result from the breaking of connectivity between previously continuous ranges. In comparison with transoceanic disjunctions (Wen, 1999; Xiang *et al.*, 2000; Sanmartín *et al.*, 2001; Donoghue & Smith, 2004), large-scale intra-continental disjunctions have received less attention in biogeography but they are especially relevant in the context of ecological vicariance. They are often explained by major environmental changes that fragment a continuously distributed lineage range, resulting in disconnected relict populations with independent evolutionary trajectories at each side of the disjunction (Axelrod & Raven, 1978; Crisp & Cook, 2007; Mairal *et al.*, 2015). In that situation, it is assumed that species migrate and conserve their climatic niche rather than evolve *in situ*. Recent phylogenetic studies show that phylogenetic niche conservatism (PNC) is a prevalent pattern in species diversification (Prinzing *et al.*, 2001; Ackerly, 2003; Crisp *et al.*, 2009; Wiens *et al.*, 2010) and might play a major role into the process of disjunction (Wiens, 2004; Kozak & Wiens, 2010). In the context of climate-driven disjunction, under the paradigm of climatic niche conservatism (Wiens & Graham, 2005), species are expected to shift their range to track suitable conditions at a pace determined by their climatic tolerances.

One of the best examples of intra-continental disjunctions is the African Rand Flora (Christ, 1892). The name refers to an enigmatic biogeographical pattern in which distantly related plant lineages share a similar disjunct distribution at the margins of the African continent and adjacent islands (between Macaronesia–north-west Africa, Horn of Africa–South Arabia–Socotra Island, Eastern Africa–Madagascar and South Africa (Bramwell, 1985; Andrus *et al.*, 2004; Sanmartín *et al.*, 2010; Pokorny *et al.*, 2015). The word 'Rand' originates from the German for edge or rim and refers to the fact that these lineages now occupy geographically isolated regions in the continental margins of Africa, forming a ring and leaving the centre of the continent 'hollow' (Pokorny *et al.*, 2015). Although Rand Flora lineages differ in their ecological affinities – some are adapted to subtropical conditions, others exhibit more xeric affinities, and some can be found in the Afrotemperate mountain regions – none of them are adapted to desert or rain forest habitats. Therefore, the tropical lowlands of Central Africa (Guineo-Congolian region) and the Sahara and Sino-Arabic Deserts in the north or the Kalahari in the south act for them as effective ecological barriers to dispersal.

The Rand Flora pattern has traditionally been explained by vicariance of a once-continuous flora across North and Central Africa by climatic aridification events (vicariance–refugium hypothesis), which left geographically disjunct sister clades, for example, between Eastern Africa/South Arabia and north-west Africa/Macaronesia (Axelrod & Raven, 1978; Bramwell, 1985; Thiv *et al.*, 2010). At the start of the Neogene (25 Ma), Africa had a warm and humid climate and was covered by tropical forests and woodlands. Global climatic cooling after the mid-Miocene Climatic Optimum (MMCO, 17–14 Ma) – and a combination of factors, including the closing of the Tethys Seaway in the north and the tectonic uplift of Eastern Africa – triggered a gradual aridification process that proceeded from south-west to north, and which replaced the former forest by grasslands and savanna (Coetzee, 1993; Plana, 2004; Micheels *et al.*, 2009; Senut *et al.*, 2009). The onset of aridity in the Sahara did not start until the end of the Miocene, 7 Ma, but during the rapid climatic oscillations of the Plio-Pleistocene, the alternation of more humid periods with other drier periods might have favoured vegetation range expansion and fragmentation across North and Central Africa (Trauth *et al.*, 2009; see Pokorny *et al.*, 2015 for a more detailed description). The vicariance–refugium hypothesis has its counterpart in the idea that these disjunct distributions were formed in more recent times by independent long-distance dispersal (LDD) events, from north to south or south to north, followed by speciation *in situ* (del Hoyo *et al.*, 2009; Sanmartín *et al.*, 2010). LDD has been used to explain other examples of wide-range disjunctions in Africa (Coleman *et al.*, 2003; Calviño *et al.*, 2006; Pelser *et al.*, 2012).

Testing the ecological vicariance–refugium hypothesis requires an accurate estimation of divergence times among the disjunct Rand Flora clades as well as of the climatic preferences of each lineage: When did they diverge? Under which climatic conditions could they maintain viable populations (Wiens, 2004)? Recently, Pokorny *et al.* (2015) answered one of these questions. Using time-calibrated molecular phylogenies, they estimated the age of the Rand Flora disjunction across 16 angiosperm lineages. Their results showed that many of these disjunctions fell within the Late Miocene-Pliocene period, after the onset of aridification in Africa, although some lineages exhibited older (mid-Miocene) and younger (Pleistocene) divergences.

Here, we investigated the relationship between the climatic tolerances of species and their age of divergence, using 14 lineages with disjunctions across the Saharo-Arabic Desert (Fig. 1, Table 1). First, we tested whether the tempo of vicariance between the disjunct western and eastern clades within each Rand Flora lineage is associated with the average values of the pooled climatic niche of the lineage. In a situation of niche conservatism under shifting climatic conditions, the climatic preferences of lineages should have constrained the temporal sequence of divergence, with the most sensitive taxa diverging earlier and those more tolerant to the new climatic conditions diverging later. In contrast, if the climate niches are labile under the considered time frame, or if climate is changing along axes that are not

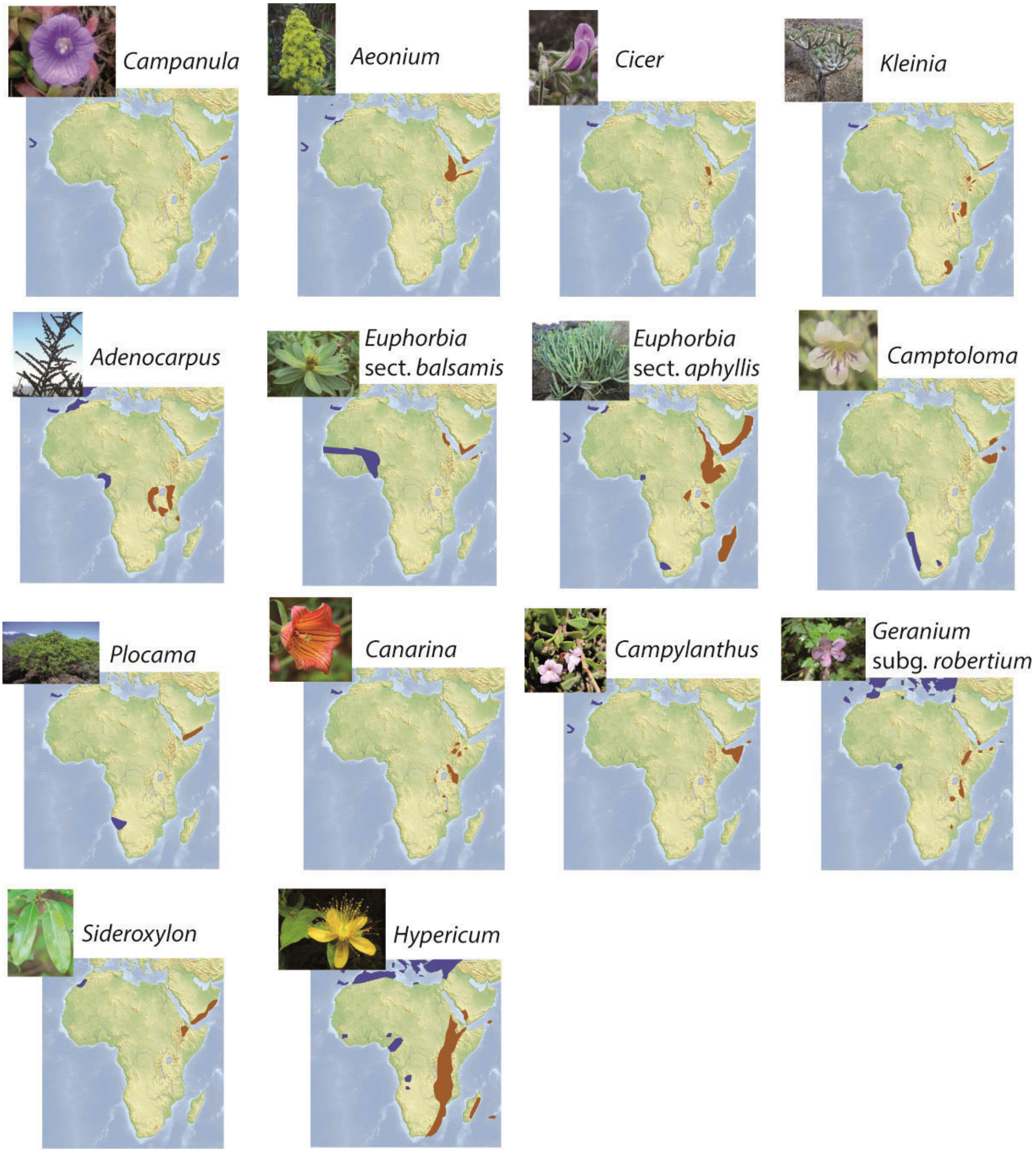


Figure 1 Current geographical distribution of the 14 Rand Flora lineages used in this study. Distribution colours represent the disjunct clades at each side of the Sahara and whose climatic niches were modelled here: east (brown) and west (blue).

Table 1 Selected Rand Flora lineages (from youngest to oldest) used in this study, exhibiting an east/west disjunction across the Saharo-Arabic Desert. For each lineage, the western and eastern disjunct clades, their species composition and age of divergence (mean and 95% HPD confidence interval, from Pokorny *et al.*, 2015) are given.

Lineage	Western clade	Western geographical distribution	Eastern clade	Eastern geographical distribution	Age (Ma)	HPD (Ma)
Campanula	*C. jacobaea*	Cape Verde	*C. balfourii*	Socotra	1	0.4–1.7
Aeonium	*A. balsamiferum*	Canary Islands	*A. leucoblepharum*	East Africa	1.65	0.2–4.7
Cicer	*C. canariense*	Canary Islands	*C. cuneatum*	East Africa–South Arabia	3.43	0.6–7.8
Kleinia	*K. neriifolia–K. anteuphorbium*	Macaronesia	*K. anteuphorbium*	East Africa	3.57	2–6.9
Adenocarpus	3 Macaronesian species (e.g. *A. ombriosus*) + *A. complicatus*	Macaronesia	*A. mannii*	Subtropical Africa	3.66	1.4–6.6
Euphorbia sect. *balsamis*	*E. balsamifera* ssp *balsamifera*	Macaronesia–W. Africa	*E. balsamifera* ssp. *adenensis*	East Africa	3.8	1.4–7.1
Camptoloma	*C. canariense*	Canary Islands	*C. lyperiiflorum*	Horn of Africa–South Africa	5.45	1.2–10.7
Euphorbia sect. *aphyllis*	12 sps. subsect. *macaronesicae* (*e.g. E. aphylla*)	Macaronesia	14 sps. subsect. *africanae* (e.g. *E. nubica*)	Central, East and South Africa	5.5	3.7–9.7
Plocama	*P. pendula*	Macaronesia	*P. tinctoria–P. yemenensis–P. crocyllis*	East Africa–South Arabia–South Africa	6.19	3–10.5
Canarina	*C. canariensis*	Macaronesia	*C. eminii*	East Africa	6.5	3.4–9.8
Campylanthus	*C. salsoloides–C. glaber*	Macaronesia	Remaining *Campylanthus* species	East Africa–Southern Arabia–Pakistan	7.5	3.1–14.2
Geranium	Nine species (e.g. *G. robertianum*)	Macaronesia–Morocco–Iberian Peninsula	Three species (e.g. *G. ocellatum*)	East Africa	11.04	7.6–15.4
Sideroxylon	*S. spinosa*	Argan sector (Morocco)	*S. mascatense*	East Africa–South Arabia	17.5	10–26.2
Hypericum	*H. androsaemum–H. grandifolium–H. hircinum–H. foliosum*	Macaronesia	*H. pamphylicum*	Turkey	17.3	9.5–25

relevant to the species, no relationship should be observed between the species climatic niche and the age of divergence. Second, to understand the spatiotemporal evolution of the Rand Flora disjunction, we modelled the climatic niche for each lineage, and used a combination of spatial distribution modelling techniques and palaeoenvironmental data to map their potential distribution across the past 17 Myr.

MATERIALS AND METHODS

Rand Flora lineages

We focused on 14 angiosperm plant lineages exhibiting a disjunction, at the level of species or clades of species (Fig. 1, Table 1), between both sides of the Saharo-Arabic desert: *Adenocarpus* DC. (Fabaceae), *Aeonium* Webb & Berth, *Campanula* L. (Crassulaceae), *Camptoloma* Benth. (Scrophulariaceae), *Campylanthus* Roth. (Scrophulariaceae), *Canarina* L. (Campanulaceae), *Cicer* L. (Fabaceae), *Euphorbia* L. sect. *balsamis, Euphorbia* L. sect. *aphyllis* (Euphorbiaceae), *Geranium* (Picard) Rouy (Geraniaceae), *Hypericum* L. (Hypericaceae), *Kleinia* Mill. (Astraceae), *Plocama* W. Aiton (Rubiaceae) and *Sideroxyon* L. (Sapotaceae). The selected lineages represent a geographical split between Eastern Africa (including the Eastern Arc Mountains, the Horn of Africa, the Socotra archipelago and southern Arabia) and Western Africa (including Macaronesia, the Atlas Mountains and region around the Cameroon volcanic line). All these lineages were included in the study by Pokorny *et al.* (2015) together with published molecular phylogenetic data that included nearly complete taxon sampling. Divergence times were estimated using a Bayesian relaxed clock, with absolute ages calibrated with the fossil record or secondary age constraints obtained from a more inclusive phylogenetic analysis; see Pokorny *et al.* (2015) for more details. Current distributions, dated chronograms of all 14 lineages indicating their distribution, phylogenetic relationships, dating uncertainty and the disjunct node of interest (i.e. the divergence between the north-western/eastern clades) are shown in Fig. 1 and in Figure S1.1 (see Appendix S1 in Supporting Information).

Occurrences and environmental data

We collected species occurrence data covering the entire distributional range of the disjunct species/clades within each Rand Flora lineage. These were obtained from published monographs and inventories, online databases (www.jardincanario.org/flora-de-gran-canaria; www.gbif.org, www.anthos.es) and georeferenced populations collected during several fieldtrips between 2010 and 2014 (Madeira, Canary Islands, Cape Verde, Morocco, Iberian Peninsula, Ethiopia, Kenya, Uganda and Madagascar). In total, 4286 georeferenced localities were collated. Coordinates were checked by projecting them on the appropriate geographical map and then eliminating localities of questionable reliability. The availability of thousands of georeferenced extant occurrences allowed us to quantify the climatic variables that limit the geographical distribution of the disjunct Rand Flora clades. We used the 19 bioclimatic variables available from WorldClim (at a resolution of 30 arc-seconds; www.worldclim.org; Hijmans *et al.*, 2005) and extracted the values of these variables for each occurrence of the disjunct species in each Rand Flora lineage. We did this separately for each bioclimatic variable and computed the mean for each species.

Climatic niche and age of divergence

We investigated whether past climate change drove population divergence in Rand Flora lineages along a temporal sequence that was correlated to their climatic niche: if climate change is gradual over time, the age of the divergence tempo of vicariance should be correlated with the lineage climatic tolerance. Using a linear model, we estimated the relationship between the age of divergence of the eastern and western clades in each Rand Flora lineage and the mean climatic niche of the lineage (i.e. the two clades pooled) from the bioclimatic variables. We also evaluated whether a quadratic term provided a better fit of the relationship to the data than a linear model using the Akaike information criterion (AIC). Ecological impact of climate change over time may begin with mild changes and accelerate over time, which might generate a nonlinear relationship between the climatic niche and age of divergence. To control for false discovery rate, we applied corrections to the *P*-values associated to our regression linear model and the quadratic term, using the R package ‘fdrtool’ (Strimmer, 2008).

While a significant relationship between age of divergence and climatic niche would provide evidence consistent with niche conservatism, we performed an additional analysis to evaluate this assumption. For each bioclimatic variable, we evaluated whether the climatic niche variation was lower among species belonging to the same lineage (e.g. between *H. pamphylicum* and *H. grandifolium*) than among species belonging to different lineages irrespective of the side of the disjunction (e.g. between *H. grandifolium* and *Plocama pendula*). For those variables showing a significant association with age of divergence in the previous analysis, we computed the average niche value for each disjunct Rand Flora species. We then computed the difference in the average niche value between all sister taxa and tested whether the difference is lower within than between lineages using a Mantel test. For each climatic variable, we plotted the niche difference within the same Rand Flora lineage and across different Rand Flora lineages; if the second is consistently larger than the first across variables, this would support PNC. Finally, we performed a principal component analysis (PCA) with all climatic variables to quantify the degree of climatic niche overlap between the east/west disjunct sister taxa in relation to all other species in the genus/lineage.

Species distribution modelling

To model the present and past distribution of clades, we combined the available occurrences for the species within each lineage with a set of six bioclimatic variables for which estimates were available for the past: total annual precipitation, maximum and minimum monthly precipitation, annual mean temperature, and maximum and minimum monthly temperature. We ran the analyses considering 2- to 4-month periods that cover the two seasons with more accentuated differences in precipitation: November to February and June to September. Pseudoabsences were generated by selecting 5000 random points across Africa, as suggested in Wisz & Guisan (2009). We used ensemble modelling to generate our predictions. This procedure integrates the results from multiple modelling techniques within an ensemble hindcasting framework to achieve more robust hindcasts (Araújo & New, 2007). Four modelling techniques, generalized linear models, generalized additive models (GAM), general boosting method (GBM) and random forests, were run and summarized using R packages (R Core Team, 2014): 'biomod2' (Thuiller *et al.*, 2013), 'foreign', 'raster' (Hijmans & van Etten, 2016), 'SDMTools' (VanDerWal *et al.*, 2011), 'rms' (Harrell, 2016), 'gbm' (Ridgeway, 2015), 'gam' (Hastie, 2016), 'rJava' (Urbanek, 2010), 'dismo' (Hijmans *et al.*, 2016) and 'randomForest' (Liaw & Wiener, 2002) with default settings. For past climate simulations, we used four global Hadley Centre general circulation models (HadCM3L) that incorporate the effect of changes in atmospheric CO_2 and represent major climate (warming or cooling) events worldwide in Africa (Meseguer *et al.*, 2015): a 400 p.p.m. CO_2 Late Miocene simulation representing the early warm period of the MMCO, a 280 p.p.m. CO_2 Late Miocene simulation representing the cold and dry conditions prevalent after the Late Miocene Cooling (LMC) event (*c.* 11.6–5.3 Ma) and a 560 p.p.m. CO_2 Pliocene simulation representing the conditions at the Mid-Pliocene Warming Event (MPWE; 3.6 Ma) and a simulation of the Preindustrial World with 280 p.p.m. CO_2 to provide a baseline HadCM3L climate before the industrial revolution. Details of the HadCM3L model setup for the palaeoclimate simulations are given in Meseguer *et al.* (2015).

RESULTS

Climatic niche and age of divergence

We found significant relationships between the climatic niche of the lineages (i.e. the eastern and western clades pooled) and the age of the divergence between the two clades. The group of bioclimatic variables related to temperature showed the highest coefficient of determination. Maximum temperature of the warmest month showed the best fit (Fig. 2: $R^2 = 0.883$; slope = 0.165), followed by other variables related to temperature, such as mean diurnal range ($R^2 = 0.8081$; slope = 0.001) and mean temperature of the warmest quarter ($R^2 = 0.768$; slope = 0.239) (Table 2). In contrast, precipitation variables had low values of coefficient of determination. These results were corroborated by the models adding a quadratic term, which showed a smaller AIC value for the variables with the highest coefficient of determination (maximum temperature of the warmest month $55.56 < 60.61$; mean diurnal range $63.58 < 67.54$; mean temperature of the warmest quarter $61.44 < 70.2$; see Table 2). Supplemental plots for all climate variables are shown in Figure S1.2. Overall, the oldest lineages showed climatic tolerances to warmer climates (e.g. *Sideroxylon*, *Hypericum*), while younger divergences were correlated to more

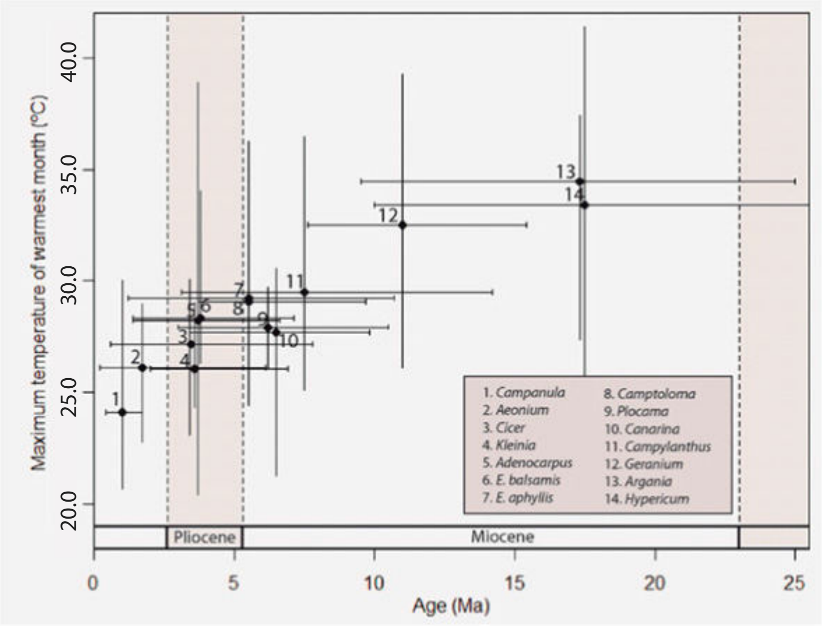


Figure 2 For each Rand Flora lineage analysed here, the age of divergence (in Ma) between the western and eastern disjunct clades is plotted against their climatic optimum for the maximum temperature of the warmest month (in °C). Bars in the *x*-axis show the 95% HPD confidence intervals for the age estimates (full circles represent the mean estimates). Bars in the *y*-axis show the 5% and 95% quantiles for the temperature range.

Table 2 Statistical relationships between the age of divergence of the eastern and western clades (Table 1) and the mean climatic niche of the lineage (the two clades pooled); R^2: coefficient of determination obtained using a linear model. Also shown are the results of the models including both linear and quadratic terms. AIC, Akaike information criterion.

Bioclim variable	R^2	Linear term	AIC	Linear term	Quadratic term	AIC
Maximum temperature of warmest month	0.8831	0.165	60.61	0.9165	4.68E-07	55.56
Mean diurnal range	0.8081	0.0001	67.54	0.8519	1.09E-05	63.58
Mean temperature of warmest quarter	0.7681	0.2395	70.2	0.8729	4.72E-06	61.44
Temperature annual range	0.7442	0.0015	71.57	0.7029	0.0005032	73.33
Annual mean temperature	0.3203	0.034	85.25	0.2455	0.0853	86.39
Temperature seasonality	0.2921	0.046	85.82	0.3444	0.09808	86.75
Mean temperature of wettest quarter	0.2691	0.0573	86.27	0.2164	0.1044	86.9
Mean temperature of driest quarter	0.1703	0.1425	90.63	−0.116	0.7296	91.85
Precipitation of driest month	0.165	0.1496	88.13	0.1953	0.1208	87.28
Mean temperature of coldest quarter	0.155	0.1637	90.55	0.6825	0.0007259	74.26
Precipitation of coldest quarter	0.1451	0.179	88.46	0.03738	0.3236	89.78
Precipitation of wettest quarter	0.1164	0.2326	88.92	0.124	0.1926	88.46
Precipitation of wettest month	0.1005	0.2694	89.17	0.101	0.2221	88.83
Minimum temperature of coldest month	0.0765	−0.0497	88.3	0.2705	0.07041	85.9
Precipitation of warmest quarter	0.0685	0.366	89.66	0.1784	0.1354	87.57
Annual precipitation	0.042	0.2339	88.93	−0.04497	0.5082	90.93
Precipitation seasonality	0.0177	0.064	90.41	−0.1227	0.7542	91.94
Isothermality	0.017	0.6565	90.42	0.1194	0.1983	88.54
Precipitation of driest quarter	0.0016	0.8915	88.04	0.1389	0.1753	88.22

temperate or arid climatic preferences (Fig. 2). The Mantel test indicated a statistically significant positive relationship between the climatic and phylogenetic matrices for the majority of variables (Table 3), although this number decreased after applying the Bonferroni correction. Similarly, plotting the median and quartile ranges for these variables showed that the within-lineage difference for niche values was lower than the difference across lineages (Figure S1.3). Although the Bonferroni correction is often considered to be too conservative, we noted that the variables showing the lowest *P*-value (Table 3) did not always show the largest statistical correlation to the age of divergence (Table 2, 3). Finally, the PCA analysis (Figure S1.4) showed that the degree of overlap in climatic niche space between the east/west disjunct sister taxa is comparable to the one found in the remaining species in the lineage/genus; even those with less overlap between east and west (e.g. *Campylanthus*) showed values falling within the genus climatic niche space.

Table 3 Mantel test results to assess whether within-lineage difference in climatic values is lower than the difference across lineages for each bioclimatic variable. *Variables that remain significant after Bonferroni test correction. *Q*-values are the false discovery rate-adjusted *P*-values.

Bioclim variable	*Z*-statistics	*P*-value	*Q*-value
Annual mean temperature	530.15	0.001*	0.016
Mean diurnal range	38.56	0.009	0.035
Isothermality	174.76	0.001*	0.007
Temperature seasonality	268.83	0.071	0.004
Maximum temperature of warmest month	588.57	0.099	0.034
Minimum temperature of coldest month	841.12	0.015*	0.18
Temperature annual range	102.32	0.637	0.045
Mean temperature of wettest quarter	609.543	0.001*	0.025
Mean temperature of driest quarter	510.66	0.003	0.062
Mean temperature of warmest quarter	459.09	0.005	0.188
Mean temperature of coldest quarter	778.52	0.001*	0.002
Annual precipitation	777.96	0.135	0.011
Precipitation of wettest month	142.78	0.051	0.18
Precipitation of driest month	218.35	0.481	0.067
Precipitation seasonality	412.16	0.781	0.002
Precipitation of wettest quarter	362.86	0.043	0.013
Precipitation of driest quarter	797.67	0.433	0.034
Precipitation of warmest quarter	217.26	0.169	0.04
Precipitation of coldest quarter	263.48	0.509	0.142

Species distribution modelling

Species distribution models and palaeoenvironmental data indicated a general decrease in climatic suitability through time for most species since the mid-Miocene (Fig. 3). Interestingly the MMCO showed suitable projections for mid-Miocene groups (*Hypericum* and *Sideroxylon*; Fig. 3a), perhaps signalling that these groups are more ancient and paratropical. Projections in the Late Miocene period, representing a global cooling event, revealed wide distribution ranges with connections joining the east and west edges of Africa. Several clades, whose divergences date back to the Late Miocene-Zanclean period (11.67–3.6 Ma; Fig. 3b), are shown to have suitable ranges connecting Eastern and Western Africa with an environmentally suitable corridor south of the Sahel (e.g. *Geranium* subgenus *Robertium*, *Plocama*, *Canarina*, *Campylanthus*, *Euphorbia* sect. *balsamis*, *Euphorbia* sect. *aphyllis*,

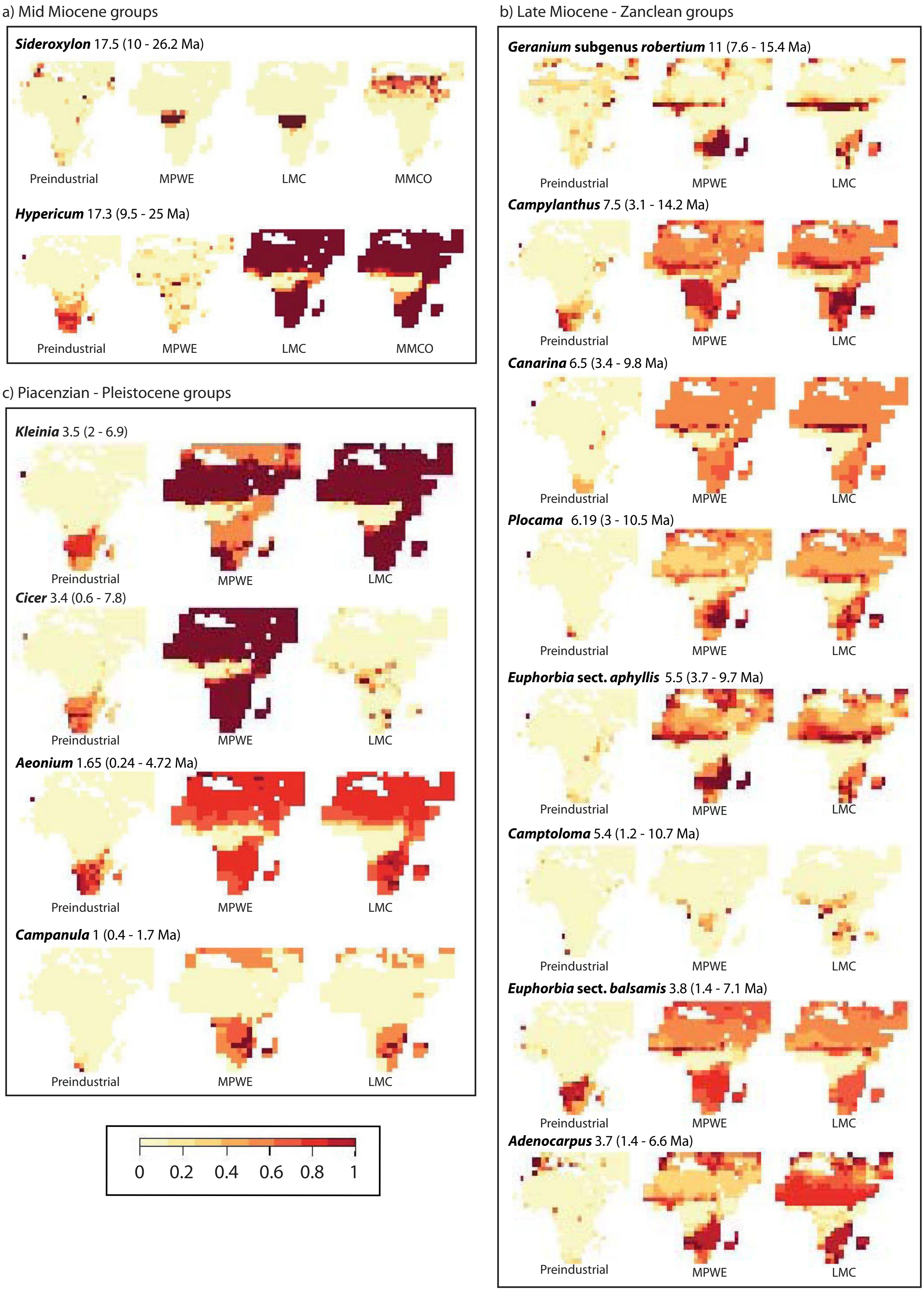

a) Mid Miocene groups
Sideroxylon 17.5 (10 - 26.2 Ma)
Preindustrial
MPWE
LMC
MMCO
Hypericum 17.3 (9.5 - 25 Ma)
Preindustrial
MPWE
LMC
MMCO
c) Piacenzian - Pleistocene groups
Kleinia 3.5 (2 - 6.9)
Preindustrial
MPWE
LMC
Cicer 3.4 (0.6 - 7.8)
Preindustrial
MPWE
LMC
Aeonium 1.65 (0.24 - 4.72 Ma)
Preindustrial
MPWE
LMC
Campanula 1 (0.4 - 1.7 Ma)
Preindustrial
MPWE
LMC
0
0.2
0.4
0.6
0.8
1
b) Late Miocene - Zanclean groups
Geranium subgenus robertium 11 (7.6 - 15.4 Ma)
Preindustrial
MPWE
LMC
Campylanthus 7.5 (3.1 - 14.2 Ma)
Preindustrial
MPWE
LMC
Canarina 6.5 (3.4 - 9.8 Ma)
Preindustrial
MPWE
LMC
Plocama 6.19 (3 - 10.5 Ma)
Preindustrial
MPWE
LMC
Euphorbia sect. aphyllis 5.5 (3.7 - 9.7 Ma)
Preindustrial
MPWE
LMC
Camptoloma 5.4 (1.2 - 10.7 Ma)
Preindustrial
MPWE
LMC
Euphorbia sect. balsamis 3.8 (1.4 - 7.1 Ma)
Preindustrial
MPWE
LMC
Adenocarpus 3.7 (1.4 - 6.6 Ma)
Preindustrial
MPWE
LMC

Figure 3 Geographical projection of the climatic niche of Rand Flora lineages over four different palaeoclimate simulations representing major warming or cooling events in the Late Cenozoic Earth history (Meseguer *et al.*, 2015). Reconstructions are grouped into three geological periods according to the estimated age of the disjunction (Table 1): (a) mid-Miocene groups (17–14 Ma); (b) Late Miocene-Zanclean groups (11.6–3.6 Ma) and (c) Piacenzian-Pleistocene groups (3.6–0.12 Ma). For each lineage, mean age of divergence between the western and eastern clades is shown in million years (Ma) with the 95% HPD confidence interval in parentheses (Table 1). Abbreviations for palaeoclimatic layers: MMCO (Mid-Miocene Climatic Optimum: 17–14 Ma), LMC (Late Miocene Climate cooling: 11.6–5.3 Ma), MPWE (Mid-Pliocene Warming Event: 3.6 Ma) and Preindustrial (Preindustrial world, < 1900). Colours in the bar at the bottom left indicate the climatic suitability likelihood for the presence of each clade and according to its climatic tolerance range: soft yellow colours indicate low climatic suitability values, while dark red indicates high suitability regions.

Adenocarpus). In the mid-Pliocene simulation (representing a major warming event), the predicted potential distribution shows wider disjunctions that split the suitable range for these taxa on either side of the continent. Other lineages with younger divergences show broad disjunctions and high suitability values during the mid-Pliocene (*Kleinia, Cicer, Aeonium*). In general, hindcasted models showed wider suitable ranges for Rand Flora species in North and Central Africa in the past before the aridification of the Sahara as compared to their current disjunct distributions.

DISCUSSION

Under the PNC hypothesis (Donoghue, 2008), it has been proposed that species should retain similar environmental niches over evolutionary time scales (niche conservatism, Wiens & Graham, 2005; Wiens *et al.*, 2010) and that this constraint should modulate the rate of speciation of clades (Wiens, 2004). Species failure to adapt to new environmental conditions should isolate populations on peripheral locations, which would trigger the formation of new lineages. A significant relationship between the time-since-speciation and the environmental niche of sister taxa is expected to provide clues on the role of niche conservatism in driving speciation (Pyron *et al.*, 2014).

Tempo of vicariance and climatic niche

We found a strong positive correlation between the time of divergence of sister-species/clades distributed in Macaronesia–NW Africa and Eastern Africa–Arabia and the mean climatic niche of the Rand Flora lineages (Table 2). This suggests that niche conservatism, manifested by the failure of a species to adapt to increasingly arid conditions in North Africa, starting around the mid-Late Miocene (Pokorny *et al.*, 2015), likely drove the initial splitting of allopatric lineages. In addition to the lineages studied here, other examples of biotic disjunction between the edges of the African continent have been described in vascular plants (Sanmartín *et al.*, 2010), vertebrates (Clouet & Wink, 2000; Lorenzen *et al.*, 2012) and invertebrates (Opatova & Arnedo, 2014). Therefore, in the context of Saharan desertification, our results support the hypothesis of large-scale niche conservatism (Crisp *et al.*, 2009), in which species generally retained their ancestral environmental constraints and shifted their range at a pace matching their climatic niche, thus initiating vicariance and speciation.

Those Rand Flora lineages characterized by the highest values of maximum temperature of the warmest month (Fig. 2), and therefore more subtropical climatic niches, had the earliest split between eastern and western African clades (e.g. *Sideroxylon*, *Hypericum*). This fits with the evolutionary history of these lineages, which have been postulated to have originated from 'boreotropical' ancestors, that is, Holarctic origins from tropical ancestors in Early Cenozoic times when tropical climates were widespread across North America and Europe (Smedmark & Anderberg, 2007; Meseguer *et al.*, 2013). The next clades with ancient age of divergence are characterized by more temperate or subxeric affinities (e.g. *Canarina*, *Geranium*, *Campylanthus*, *Plocama*). In contrast, groups such as *Kleinia* (Fabaceae) or *Aeonium* (Crassulaceae) show younger splits and adaptations to drier habitats (Fig. 2). Thus, the sequence of splits in Rand Flora lineages seems to match the lineage climatic niches, following the gradual aridification of North Central Africa, with younger divergences corresponding to arid-adapted lineages. The fact that we found a slightly better fit with a quadratic model for the most informative variables suggests that the effect of climate change on species divergence might have accelerated through time.

These results corroborate the links between niche conservatism and speciation during environmental changes, in the context of ecological vicariance (Wiens, 2004). If climate change is fast and widespread, so that species have no time to adapt to new conditions, populations will either go extinct or be constrained to move and track suitable areas (Futuyma, 2010; Waldron, 2010). The species geographical range becomes fragmented into smaller refugia where populations persist, followed by local speciation (Wiens, 2004). This results in a common signature of vicariance across several unrelated lineages. The Australian flora offers such a case, where the formation of the arid Nullarbor Plain produced a congruent molecular signature of geographical division across multiple plant clades (Crisp & Cook, 2007). In a different situation, Pellissier *et al.* (2013) showed that Arctic-Alpine disjunct species retained their climatic tolerance after the Quaternary glacial period, which may lead to local vicariant speciation. On the other hand, if climate change is labile, with oscillations in different directions, or follows a spatial gradient, one would expect to see a pattern of niche evolution with the two clades at each extreme side of the disjunction adapted to different climate optima (adaptation). We did not see this pattern here: the climatic preferences of the eastern and western clades did not show a significant degree of divergence compared with the rest of species in the genus. We also found differences in climatic affinities between genera larger than within each genus, which provides support to our assumption of PNC. Finally, if climate change is gradual or accelerating in one direction as along a temporal temperature gradient, one would expect to see a pattern, where species will conserve their ancestral climatic preferences rather than adapt to the new conditions. This would result in a predictable pattern (Donoghue & Edwards, 2014) of biogeographical 'pseudocongruence' (Pokorny *et al.*, 2015): the timing of vicariance events should match the lineage climatic tolerance, with a positive relationship between the age of the disjunction event and the species climatic niche as found in this study.

Is vicariance linked to niche conservatism and persistence in refuges?

The species distribution modelling hindcasted to the reconstructed past climates showed generally broad geographical ranges for Rand Flora lineages in the Late Miocene, but a much more restricted current distribution. The colder and wetter climate of the Late Miocene resulted in favourable conditions for species, probably allowing continuous gene flow throughout North Africa. For many lineages, a climatic corridor of suitable environment can be observed south of the Sahel, connecting Eastern and Western Africa. The existence of this past climatic corridor agrees well with the study by Sanmartín *et al.* (2010), which demonstrated historically high rates of dispersal or biotic connection between north-west Africa–Macaronesia and Eastern Africa–Southern Arabia. As the climate became drier with more extreme temperature maxima towards the mid-Pliocene – which also coincided with a period of intense uplift in Eastern Africa (Sepulchre *et al.*, 2006) – species ranges contracted (Fig. 3) and central North Africa became unsuitable for most lineages. Interestingly, the opposite pattern can be observed for Holarctic groups, where the LMC event was responsible for large geographical disjunctions across North American and European temperate forests, whereas the subsequent increase in temperature at the mid-Pliocene warming interval allowed some of these forest taxa to regain a more widespread distribution (Sanmartín *et al.*, 2001; Meseguer *et al.*, 2015).

The gradual aridification resulted in lineages tracking suitable climatic conditions to have distributions restricted to climatic refugia in the east and western edges of Africa (e.g. the Afromontane region) and adjacent oceanic and continental islands (Macaronesia, Socotra). These areas conserved more subtropical or temperate habitats, due to topographic features (high altitude) or the mild influence of oceanic climates. Indeed, these regions have long been considered to represent the actual refugia of the last remnants of subtropical flora that covered North Central Africa (Axelrod & Raven, 1978), in the Afromontane archipelago of East Africa and the laurisilva forest in Macaronesian archipelagos (though the latter has been challenged by recent phylogenetic work, Kondraskov *et al.*, 2015).

Species distribution modelling hindcasted to past climate is based on the premise that extant species share the same climatic preferences as their extinct ancestors. Although this might be true for short time scales, it is less likely for geological time scales of millions of years, especially under a scenario of rapid climatic oscillations as the Cenozoic (Stigall, 2012). In the absence of fossil data to reconstruct ancestral climatic preferences (see Meseguer *et al.*, 2015), this seems an unavoidable assumption. On the other hand, large-scale PNC has been argued for angiosperms (Crisp *et al.*, 2009), based on evidence that biome shifts are rare both within continents and in transoceanic colonizations – probably because these transitions often require the development of complex physiological adaptations (e.g. Donoghue, 2008). In our study, niche conservatism is supported by the finding of lower variance within than across lineages and by the strong positive relationship between time from divergence and the lineage mean climatic niche. Thus, our study supports the hypothesis that the variation in lineage tolerance to environmental conditions played a key role in the processes that created the Rand Flora biogeographical pattern.

Overall, our results provide support to the vicariance–refugium hypothesis as an explanation of the Rand Flora disjunction: the ancestors of Rand Flora lineages had in the past a more widespread distribution across North Africa, which became fragmented as a result of the increasing aridification of the continent from the Miocene onwards. The fact that these widespread distributions are not all of the same age (Fig. 2) lends support to the idea that the Rand Flora pattern was not formed by a single vicariance event but originated at different times during episodes of global drought during the Miocene-Pliocene (Pokorny *et al.*, 2015). Our projections (Fig. 3) do not allow us to discard the possibility that some Rand Flora lineages never actually had a continuous distribution from west to east North Africa. The presence of patches of climatic suitability in MPWE palaeoclimate reconstructions suggests the possibility of stepping stones or climatic refuges facilitating migration across the Sahara during climatically unfavourable periods for at least some of these lineages.

ACKNOWLEDGEMENTS

We thank Victoria Culshaw, Lisa Pokorny and Andrea Briega for helping to collect the data. We also thank Nadir Álvarez for comments during the preparation of the work. This study was funded by the Spanish government through grants CGL2009-1332-C03-01, CGL2012-40129-C02-01 and CGL2015-67849-P (MINECO/FEDER), and a PhD fellowship (BES-2010-037261) to M.M. supervised by I.S.

REFERENCES

Ackerly, D.D. (2003) Community assembly, niche conservatism, and adaptive evolution in changing environments. *International Journal of Plant Sciences*, **164**, 165–184.

Andrus, N., Trusty, J., Santos-Guerra, A., Jansen, R.K. & Francisco-Ortega, J. (2004) Using molecular phylogenies to test phytogeographical links between East/South Africa-Southern Arabia and the Macaronesian islands: a review, and the case of *Vierea* and *Pulicaria* Section *Vieraeopsis* (Asteraceae). *Taxon*, **53**, 333–346.

Araújo, M.B. & New, M. (2007) Ensemble forecasting of species distributions. *Trends in Ecology and Evolution*, **22**, 42–47.

Axelrod, D.I. & Raven, P.H. (1978) Late Cretaceous and Tertiary vegetation history of Africa. *Biogeography and ecology of southern Africa* (ed. M.J.A. Werger), pp. 77–130. Springer, The Hague.

Bramwell, D. (1985) Contribución a la biogeografía de las Islas Canarias. *Botánica Macaronésica*, **14**, 3–34.

Brundin, L. (1967) Insects and the problem of Austral disjunctive distribution. *Annual Review of Entomology*, **12**, 149–168.

Calviño, C.I., Tilney, P.M., van Wyk, B.E. & Downie, S.R. (2006) A molecular phylogenetic study of southern African apiaceae. *American Journal of Botany*, **93**, 1828–1847.

Christ, H. (1892) Exposé sur le rôle que joue dans le domaine de nos flores la flore dite ancienne africaine. *Archives des Sciences Physiques et Naturelles, Genève*, **3**, 369–374.

Clouet, M. & Wink, M. (2000) The buzzards of Cape Verde *Buteo* (*buteo*) *bannermani* and Socotra *Buteo* (*buteo*) spp.: first results of a genetic analysis based on nucleotide sequences of the cytochrome b gene. *Alauda*, **68**, 55–58.

Coetzee, J.A. (1993) African flora since the terminal Jurassic. *Biological relationships between Africa and South America* (ed. by P. Goldblatt), pp. 37–61. Yale University Press, New Haven, CT.

Coleman, M., Liston, A., Kadereit, J.W. & Abbott, R.J. (2003) Repeat intercontinental dispersal and Pleistocene speciation in disjunct Mediterranean and desert *Senecio* (Asteraceae). *American Journal of Botany*, **90**, 1446–1454.

Crisp, M.D. & Cook, L.G. (2007) A congruent molecular signature of vicariance across multiple plant lineages. *Molecular Phylogenetics and Evolution*, **43**, 1106–1117.

Crisp, M.D. & Cook, L.G. (2012) Phylogenetic niche conservatism: what are the underlying evolutionary and ecological causes? *New Phytologist*, **196**, 681–694.

Crisp, M.D., Arroyo, M.T., Cook, L.G., Gandolfo, M.A., Jordan, G.J., McGlone, M.S., Weston, P.H., Westoby, M., Wilf, P. & Linder, H.P. (2009) Phylogenetic biome conservatism on a global scale. *Nature*, **458**, 754–756.

Croizat, L. (1952) *Manual of phytogeography*. Junk, The Hague.

del Hoyo, A., García-Marín, J.L. & Pedrola-Monfort, J. (2009) Temporal and spatial diversification of the African disjunct genus *Androcymbium* (Colchicaceae). *Molecular Phylogenetics and Evolution*, **53**, 848–861.

Donoghue, M.J. (2008) A phylogenetic perspective on the distribution of plant diversity. *Proceedings of the National Academy of Sciences USA*, **105**, 11549–11555.

Donoghue, M.J. & Edwards, E.J. (2014) Biome shifts and niche evolution in plants. *Annual Review of Ecology, Evolution, and Systematics*, **45**, 547–572.

Donoghue, M.J. & Smith, S.A. (2004) Patterns in the assembly of temperate forests around the Northern Hemisphere. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **359**, 1633–1644.

Futuyma, D.J. (2010) Evolutionary constraint and ecological consequences. *Evolution*, **64**, 1865–1884.

Harrell, J.F.E. (2016) *rms: Regression modeling strategies*. R package version 3.4-0. https://cran.r-project.org/web/packages/rms/rms.pdf

Harvey, P.H. & Pagel, M.D. (1991) *The comparative method in evolutionary biology*, Vol. **239**. Oxford University Press, Oxford.

Hastie, T. (2016) *gam: Generalized additive models*. R package version, 1. https://cran.r-project.org/web/packages/gam/gam.pdf.

Hijmans, R.J. & van Etten, J. (2016) *raster: raster: Geographic data analysis and modeling*. R package version, 2–15. https://cran.r-project.org/web/packages/raster/raster.pdf

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**, 1965–1978.

Hijmans, R.J., Phillips, S., Leathwick, J. & Elith, J. (2016) *dismo: Species distribution modeling*. R package version 0.8-17. https://cran.r-project.org/web/packages/dismo/dismo.pdf

Humphries, C.J. & Parenti, L.R. (1999) *Cladistic biogeography*. OUP, Oxford.

Kondraskov, P., Schütz, N., Schüßler, C., de Sequeira, M.M., Guerra, A.S., Caujapé-Castells, J., Jaén-Molina, R., Marrero-Rodríguez, A., Koch, M.A., Linder, H.P., Kovar-Eder, J. & Thiv, M. (2015) Biogeography of Mediterranean hotspot biodiversity: re-Evaluating the'Tertiary relict 'hypothesis of Macaronesian laurel forests. *PLoS ONE*, **10**, e0132091.

Kozak, K.H. & Wiens, J.J. (2010) Accelerated rates of climatic-niche evolution underlie rapid species diversification. *Ecology Letters*, **13**, 1378–1389.

Liaw, A. & Wiener, M. (2002) Classification and regression by randomForest. *R News*, **2**, 18–22.

Lorenzen, E.D., Heller, R. & Siegismund, H.R. (2012) Comparative phylogeography of African savannah ungulates. *Molecular Ecology*, **21**, 3656–3670.

Mairal, M., Pokorny, L., Aldasoro, J.J., Alarcón, M. & Sanmartín, I. (2015) Ancient vicariance and climate-driven extinction continental-wide disjunctions in Africa: the case of the Rand Flora genus *Canarina* (Campanulaceae). *Molecular Ecology*, **24**, 1335–1354.

Meseguer, A.S., Aldasoro, J.J. & Sanmartín, I. (2013) Bayesian inference of phylogeny, morphology and range evolution reveals a complex evolutionary history in St. John's wort (*Hypericum*). *Molecular Phylogenetics and Evolution*, **67**, 379–403.

Meseguer, A.S., Lobo, J.M., Ree, R., Beerling, D.J. & Sanmartín, I. (2015) Integrating fossils, phylogenies, and niche models into biogeography to reveal ancient evolutionary history: the case of *Hypericum* (Hypericaceae). *Systematic Biology*, **64**, 215–232.

Micheels, A., Eronen, J. & Mosbrugger, V. (2009) The Late Miocene climate response to a modern Sahara desert. *Global and Planetary Change*, **67**, 193–204.

Opatova, V. & Arnedo, M.A. (2014) Spiders on a hot volcanic roof: colonisation pathways and phylogeography of the Canary Islands endemic trap-door spider *Titanidiops canariensis* (Araneae, Idiopidae). *PLoS ONE*, **9**, e115078.

Pellissier, L., Bråthen, K.A., Vittoz, P., Yoccoz, N.G., Dubuis, A., Meier, E., Zimmermann, N.E., Randin, C.F., Thuiller, W., Garraud, L., Es, J.V. & Guisan, A. (2013) Thermal niches are more conserved at cold than warm limits in arctic-alpine plant species. *Global Ecology and Biogeography*, **22**, 933–941.

Pelser, P.B., Abbott, R.J., Comes, H.P., Milton, J.J., Möller, M., Looseley, M.E., Cron, G.V., Barcelona, J.F., Kennedy, A.H., Watson, L.E., Barone, R., Hernández, F. & Kadereit, J.W. (2012) The genetic ghost of an invasion past: colonization and extinction revealed by historical hybridization in *Senecio*. *Molecular Ecology*, **21**, 369–387.

Peterson, A.T., Soberón, J. & Sánchez-Cordero, V. (1999) Conservatism of ecological niches in evolutionary time. *Science*, **285**, 1265–1267.

Plana, V. (2004) Mechanisms and tempo of evolution in the African Guineo-Congolian rainforest. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **359**, 1585–1594.

Pokorny, L., Riina, R., Mairal, M., Meseguer, A.S., Culshaw, V., Cendoya, J., Serrano, M., Carbajal, R., Ortiz, S., Heuertz, M. & Pokorny, L. (2015) Living on the edge: timing of Rand Flora disjunctions congruent with ongoing aridification in Africa. *Frontiers in Genetics*, **6**, 1–15.

Prinzing, A., Durka, W., Klotz, S. & Brandl, R. (2001) The niche of higher plants: evidence for phylogenetic conservatism. *Proceedings of the Royal Society B: Biological Sciences*, **268**, 2383–2389.

Pyron, R.A., Costa, G.C., Patten, M.A. & Burbrink, F.T. (2014) Phylogenetic niche conservatism and the evolutionary basis of ecological speciation. *Biological Reviews*, **90**, 1248–1262.

R Core Team (2014) *R: a language and environment for statistical computing*. Version 3.1.1. Available at: http://www.r-project.org.

Ridgeway, G. (2015) *Generalized boosted regression models*. Documentation on the R Package 'gbm', version 1· 5, 7. https://cran.r-project.org/web/packages/gbm/gbm.pdf

Sanmartín, I. (2012) Historical biogeography: evolution in time and space. *Evolution: Education and Outreach*, **5**, 555–568.

Sanmartín, I., Enghoff, H. & Ronquist, F. (2001) Patterns of animal dispersal, vicariance and diversification in the Holarctic. *Biological Journal of the Linnean Society*, **73**, 345–390.

Sanmartín, I., Anderson, C.L., Alarcón, M., Ronquist, F. & Aldasoro, J.J. (2010) Bayesian island biogeography in a continental setting: the Rand Flora case. *Biology Letters*, **6**, 703–707.

Senut, B., Pickford, M. & Ségalen, L. (2009) Neogene desertification of Africa. *Comptes Rendus Geoscience*, **341**, 591–602.

Sepulchre, P., Ramstein, G., Fluteau, F., Schuster, M., Tiercelin, J.J. & Brunet, M. (2006) Tectonic uplift and Eastern Africa aridification. *Science*, **313**, 1419–1423.

Smedmark, J.E. & Anderberg, A.A. (2007) Boreotropical migration explains hybridization between geographically distant lineages in the pantropical clade Sideroxyleae (Sapotaceae). *American Journal of Botany*, **94**, 1491–1505.

Stigall, A.L. (2012) Using ecological niche modelling to evaluate niche stability in deep time. *Journal of Biogeography*, **39**, 772–781.

Strimmer, K. (2008) fdrtool: a versatile R package for estimating local and tail area-based false discovery rates. *Bioinformatics*, **24**, 1461–1462.

Thiv, M., Thulin, M., Hjertson, M., Kropf, M. & Linder, H.P. (2010) Evidence for a vicariant origin of Macaronesian–Eritreo/Arabian disjunctions in *Campylanthus* Roth (Plantaginaceae). *Molecular Phylogenetics and Evolution*, **54**, 607–616.

Thuiller, W., Georges, D. & Engler, R. (2013) *biomod2: Ensemble platform for species distribution modeling*. R package version, 2, r560. http://www2.uaem.mx/r-mirror/web/packages/biomod2/biomod2.pdf

Trauth, M.H., Larrasoaña, J.C. & Mudelsee, M. (2009) Trends, rhythms and events in Plio-Pleistocene African climate. *Quaternary Science Reviews*, **28**, 399–411.

Urbanek, S. (2010) *rJava: Low-level R to Java interface*. https://cran.r-project.org/web/packages/rJava/rJava.pdf

VanDerWal, J., Falconi, L., Januchowski, S., Shoo, L. & Storlie, C. (2011). *SDMTools: Species distribution modelling tools: Tools for processing data associated with species distribution modelling exercises*. R package version, 1.

Waldron, A. (2010) Lineages that cheat death: surviving the squeeze on range size. *Evolution*, **64**, 2278–2292.

Wen, J. (1999) Evolution of Eastern Asian and Eastern North American disjunct distributions in flowering plants. *Annual Review of Ecology, Evolution, and Systematics*, **30**, 421–455.

Wiens, J.J. (2004) Speciation and ecology revisited: phylogenetic niche conservatism and the origin of species. *Evolution*, **58**, 193–197.

Wiens, J.J. & Graham, C.H. (2005) Niche conservatism: integrating evolution, ecology and conservation biology. *Annual Review of Ecology, Evolution, and Systematics*, **36**, 519–539.

Wiens, J.J., Ackerly, D.D., Allen, A.P., Anacker, B.L., Buckley, L.B., Cornell, H.V., Damschen, E.I., Davies, T.J., Grytnes, J.A., Harrison, S.P., Hawkins, B.A., Holt, R.D., McCain, C.M. & Stephens, P.R. (2010) Niche conservatism as an emerging principle in ecology and conservation biology. *Ecology Letters*, **13**, 1310–1324.

Wisz, M.S. & Guisan, A. (2009) Do pseudo-absence selection strategies influence species distribution models and their predictions? An information-theoretic approach based on simulated data. *BMC Ecology*, **9**, 8.

Xiang, Q.Y., Soltis, D.E., Soltis, P.S., Manchester, S.R. & Crawford, D.J. (2000) Timing the Eastern Asian-Eastern North American floristic disjunction: molecular clock corroborates palaeontological estimates. *Molecular Phylogenetics and Evolution*, **15**, 462–472.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Appendix S1 Supplementary figures.

BIOSKETCHES

Mario Mairal's research is focused on macroevolution and microevolution in island systems, with special focus in the biogeography and ecology of the Rand Flora pattern.

Isabel Sanmartín is a senior researcher at the Real Jardín Botánico, CSIC (Madrid, Spain). Her main research interests include the study of large-scale biogeographical patterns and the development of analytical inference methods with special reference to model-based approaches.

Loïc Pellissier is Assistant Professor in Landscape Ecology and investigates how past landscape changes have influenced species distributions to guide future forecast of biodiversity responses to global changes.

Author contributions: M.M., I.S. and L.P. designed the study; M.M. and I.S. gathered the data; M.M. and L.P. analysed the data; M.M., I.S. and L.P. led the writing.

Editor: Richard Ree

ANEXOS /

ANNEXES

SUPPLEMENTARY DATA
CHAPTER 1

Ancient vicariance and climate-driven extinction explain continental-wide disjunctions in Africa: the case of the Rand Flora genus *Canarina* (Campanulaceae)

SUPPLEMENTARY TEXT

1) *Study-group*

All species of *Canarina* are herbs with thick perennial roots, seasonal herbaceous stems and large, bell-shaped solitary flowers. They have fleshy and indehiscent fruits and tricolporate or tricolporoidate pollen grains. The basic chromosome number is n=17, and the three species are diploids with 2n = 34 (Hedberg 1961). Eastern African *Canarina* species have been shown to be pollinated by nectarivorous birds (Olesen et al. 2012), meanwhile *C. canariensis* is pollinated by opportunistic nectar-feeding birds (Rodríguez-Rodríguez & Valido 2011).

Canarina abyssinica Engl. (1902) is a terrestrial climber herb up to 2m long, occasionally longer. Its leaves are triangular to pentagonal and have characteristic long coiled petioles that they use to climb up other plants. The species occurs in a few sites in the uplands of East Africa; southern Sudan, the Ethiopian highlands, Uganda, western Kenya and Northern Tanganyika (Fig 1b, Hedberg 1961), although this distribution could be at present more restricted (see below). Its habitats are rocky outcrop zones and opened wooded grasslands and savannas; it does not occur in closed forests. Its altitudinal range spans between 1400 and 2500 m.

Canarima eminii Aschers. ex Schweinf. (1892) is a terrestrial or mostly epiphytic herb up to 2m long. Its leaves are triangular to ovate. Intimately associated to the afromontane forests that form vegetation belts in the East African mountain systems, its distribution is very fragmented, occurring in southern Sudan, the Ethiopian highlands, Uganda, Rwanda, East Democratic Republic of Congo, Kenya, and the north of Lake Malawi (Tanzania) as its southernmost distribution (Fig 1b). Its habitats are afromontane forests, growing frequently in the canopy as epiphyte, between 1500 and 3200 m.

Canarina canariensis (L.) Vatke. (1874) is a terrestrial herb with prostrate or climbing erect stem, up to 3m long. Its leaves are hastate. It is often found in the scrubs, ravines, shaded rocks and opened habitats surrounding the laurisilva forest (between 200 and 1000 m of altitude) of the westernmost islands: Gran Canaria, Tenerife, La Gomera, La Palma and El Hierro, but it is only abundant in the first two islands.

2) Expanded Material and Methods

2.1 DNA amplification, sequencing and alignment

DNA was extracted using the DNeasy Plant Mini Kit (QUIAGEN Inc., California) following the manufacturers' instructions from silica gel dried leaves obtained from specimens collected in several field expeditions to the Canary Islands and East Africa between 2009 and 2012. A pilot study was carried out to select regions within the chloroplast genome that provide adequate variation at species and population levels.

The seven cpDNA loci selected, together with the nuclear ribosomal internal transcribed spacer (ITS), were successfully amplified using the primers listed in Table S2. Amplification was achieved in a 25μl reaction volume using the PCR mix BioMix (Bioline, Germany) with 1μl of DNA and 1μl of bovine serum albumin. PCR conditions were: initialization at 95ºC for 1–4 min; 35 cycles of denaturation at 95ºC for 1 min, annealing at various temperatures —52ºC for *pet*B-*pet*D, *trn*G intron, *rpl*32-*trn*L, *3´trn*V-*ndh*C, *psb*J-*pet*A, and *trn*L-*trn*F, 52ºC-60ºC for *trn*S-*trn*G, and 54ºC for ITS— for 2 min, and elongation at 72ºC for 2 min; and a final extension at 72ºC for 10 min. PCR products were checked on 1% agarose gels and purified using ExoSap® or the Zymx purification kit. PCR products were sent to Macrogen (Korea) and StabVida (Lisboa, Portugal) for sequencing, using the aforementioned PCR primers (Table S2). Amplification of ITS revealed no double bands and clear single peaks in sequences, so no cloning strategy was necessary. Sequences were edited in Geneious Pro 5.4.4, and aligned using MAFFT v. 6.814b with the E-INS-I algorithm and manually adjusted when necessary following alignment rules described by Kelchner (2000).

2.2 Phylogenetic Inference

Phylogenetic relationships were estimated for each marker separately using Bayesian Inference, implemented in MrBayes 3.2.2 (Ronquist *et al*. 2012). Choice of substitution models was based on the Akaike Information Criterion implemented in MrModelTest 2.2 (Nylander 2004) and run in PAUP* v4.0b10 (Swofford 2002): the General Time Reversible model (GTR) was selected for all regions except ITS and *rpl*32, which added a gamma prior distribution to model among-site rate variation (GTR + G). Two independent analyses of four chains each were run for 10 million generations, sampling every 1000th. Convergence was assessed by monitoring cumulative split frequencies. After discarding the first 25% samples as burnin, we pooled the remaining trees to construct a 50% majority rule consensus tree. Additionally, Maximum Likelihood analyses were run in the software RAxML (Stamatakis *et al*. 2008) using the online tool (http://embnet.vital-it.ch/raxml-bb/); clade support was assessed by bootstrap analysis using 100 replicates.

On our sampling effort and the decreasing distribution of Canarina abyssinica.

In an effort to collect new fresh material of the East African species of *Canarina*, we carried out several field expeditions to Ethiopia (2009) and Kenya-Uganda (2010), led by J.J. Aldasoro. These failed to find *Canarina abyssinica* in any of the visited localities recorded in the literature (Hedberg, 1961), despite exhaustive search attempts (e.g., Kotob, Cabanne Valley, Zuqualla Abo Monastery, Agere Maryan, Mt Elgon, Manu escarpment, Aberdares, Timderet, Lumbwa). Similarly, a recent field expedition to Ethiopia (16-30 November 2014), led by L. Pokorny, failed to find the species in additional recorded localities (i.e., Harar; Shashemene). Moreover, we noted that much of the forest in these localities was highly degraded, having either completely disappeared or being threatened by the advancement of extensive agriculture (i.e., wheat and corn, rather than the native *Eragrostis tef* or *Sorghum bicolor*), forestry (i.e., native forests are being replaced by *Cupressus lusitanica* stands and fast-growing *Eucalyptus globulus* plantations), and other human activities (i.e., cattle raising).

According to EFAP (1994), up to 35% of Ethiopian Highlands were covered by montane forests a couple hundred years ago (Kebede *et al*. 2007), however due to population pressure and changes in the use of land, forested areas in Africa are now in decline (even faster in Ethiopia), as a FAO report (2001) highlights. *Canarina* *abyssinica* is uncommon in wet forests (Agnew & Agnew 1994) and mostly lives in open mountain forests (Hedberg 1961), which are in a "downward spiral in soil degradation" according to Kloos & Legesse (2010). Thus, it is likely that the species has been seriously affected by the deforestation and land clearing that afflicts Eastern Africa (see Forest of Sheka, page 80 in. http:// www.melcaethiopia.org/images/ stories/Publication/ Forests%20of%20Sheka.pdf

Interestingly, *C. abyssinica* is reported to be abundant in the Sheka Forest Biosphere Reserve, in southwestern Ethiopia (http:// www.cepf.net/SiteCollectionDocuments/ eastern_afromontane/62562_Safeguard_ShekaForestBiosphereReserveNominationForm .pdf). However, a careful study of the pictures depicted in this report showed that these actually corresponds to *C. eminii*, which is indeed an epiphyte as described in page 6, whereas *C. abyssinica* is not. This confusion between the two species does not seem not to be uncommon, as they are often misidentified in herbaria such as in the collections of the University of Uppsala (J.J. Aldasoro, pers. comm.). The confusion might also stem from the fact that the types of *C. abyssinica* and *C. eminii* were burned in a fire in Berlin during World War II (c.f. Hedberg 1961, page 40). It might also be attributed to the old age ot the specimens; most of the occurrences recorded in the GBIF dataset are more than 40 years old.

In sum, although we admit *that Canarina abyssinica* could still be found in other areas not sampled by us (for example, we did not sample all the mountains of Kenya, South Sudan or the border of Ethiopia with Somalia, due to security concerns), we believe that the species is less abundant now that it was in the past (Hedberg, 1961) and can be considered as in retreat or under threat from dissappearing from its current habitat.

Additional references:

Agnew ADQ, Agnew S (1994) Kenya upland wild flowers. East African Natural History, Nairobi, Kenya.

Harrell JFE (2012) rms: Regression modeling strategies. R package version 3.4-0.

Hastie T (2008) gam: Generalized additive models. *R package version*, **1**.

Hijmans RJ, van Etten J (2014) raster: raster: Geographic data analysis and modeling. *R package version*, 2-2.

Hijmans RJ, Phillips S, Leathwick J, Elith J (2012) dismo: Species distribution modeling. *R package version* 0.7-17.

Kelchner SA (2000) The evolution of non-coding chloroplast DNA and its application in plant systematics. *Annals of the Missouri Botanical Gard*en, **87**, 482–498.

Kloos H, Legesse W (Eds.) (2010) Water resources management in Ethiopia: implications for the Nile basin. Cambria Press.

Liaw A, Wiener M (2002) Classification and Regression by randomForest. *R news*, **2**, 18-22.

Nylander JAA (2004) MrModeltest v2. Program distributed by the author. *Evolutionary Biology Centre, Uppsala University*, 2.

Olesen JM, Alarcon M, Ehlers BK, Aldasoro JJ, Roquet C (2012) Pollination, biogeography and phylogeny of oceanic island bellflowers (Campanulaceae). *Perspectives in Plant Ecology, Evolution and Systematics*, **14**, 169-182.

Ridgeway G (2006) Generalized boosted regression models. Documentation on the R Package ‘gbm’, version 1· 5, 7.

Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard, Huelsenbeck JP (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic biology*, **61**, 539–542.

Rodríguez-Rodríguez MC, Valido A (2011) Consequences of plant-pollinator and floral-herbivore interactions on the reproductive success of the Canary Islands endemic *Canarina canariensis* (Campanulaceae). *American journal of botany*, **98**, 1465-1474.

Team RC (2012) R: A language and environment for statistical computing.

Thuiller W, Georges D, Engler R (2013) biomod2: Ensemble platform for species distribution modeling. R package version, **2**, r560.

Urbanek S (2010) rJava: Low-level R to Java interface.

VanDerWal J, Falconi L, Januchowski S, Shoo L, Storlie C (2011). SDMTools: Species distribution modelling tools: Tools for processing data associated with species distribution modelling exercises. R package version, **1**.

SUPPLEMENTARY TABLES

Table S1. Voucher information and GenBank accession numbers for all taxa included in this study. Newly generated sequences are printed in bold. Abbreviations for the main geographical areas discussed in the text: WR, West Rift; ER, East Rift; UG, Uganda; GC, Gran Canaria; TFE, Tenerife East; TFW, Tenerife West; GO, Gomera; LPA, La Palma; EH, El Hierro.

Population # / Code	**Taxon**	**Locality (Country)**	**Collector and Coll. # (Herbarium)**	**Geographic coordinates**	***Rpl*32-*trn*L**	**3´*trn*V-*ndh*C**	***psb*J-*pet*A**	***pet*B-*pet*D**	***trn*L-*trn*F**	***trn*S-*trn*G**	***trn*G intron**	**ITS**
1WR- Gifta	*C. eminii*	Debre Markos to Bahir Da Road, Gifta, N of Bure (Ethiopia)	*Aldasoro & Alarcón 9982* (BC)	10° 35' 13.0626" N, 37° 29' 4.4838" E	**KP761548**	**KP761681**	**KP761514**	**KP761488**	**KP761613**	**KP761632**	**KP761579**	**KP761452**
2 WR- Dembecha	*C. eminii*	Debre Markos to Bahir Dar Road, Dembecha to Bure (Ethiopia)	*Aldasoro & Alarcón 9979* (BC)	10° 36' 53.5926" N, 37° 30' 9.6798" E	**KP761549**	**KP761683**	**KP761515**	**KP761489**	**KP761614**	**KP761633**	**KP761580**	**KP761451**
3 ER- Harenna	*C. eminii*	Harenna Forest, Negele to Goba Road (Ethiopia)	*Aldasoro & Alarcón 10322* (BC)	6º 42' 03" N, 39º 43' 35" E	**KP761553**	**KP761685**	**KP761518**	**KP761485**	**KP761611**	**KP761638**	**KP761583**	**KP761450**
4 ER- Yirga	*C. eminii*	Agere Maryan, Yirga Chefé Road	*Aldasoro & Alarcón 10060* (BC)	6º 04' 54" N, 38º 14' 33" E	**KP761552**	**KP761682**	**KP761516**	**KP761484**	**KP761610**	**KP761637**	**KP761582**	**KP761448**
5 UG-Elgon	*C. eminii*	Mount Elgon (Uganda)	*Sánchez-Meseguer & Aldasoro 69* (BC)	6º 39' 3"N, 39º 43' 57"E	**KP761550**	**KP761684**	**KP761517**	**KP761483**	**KP761615**	**KP761634**	**KP761584**	**KP761447**
6 ER- Aberdares	*C. eminii*	Aberdare Mts. (Kenia)	*Sánchez-Meseguer & Aldasoro 103* (BC)	0º 45' 52" N, 36º 44' 35" E,	**KP761554**	**KP761686**	**KP761520**	**KP761486**	**KP761612**	**KP761636**	**KP761585**	**KP761449**
7 UG- Rwenzori	*C. eminii*	Rwenzori Mountains (Uganda)	*Sánchez-Meseguer & Aldasoro 18* (BC)	0º 21' 35" N, 29º 58' 20'E'	**KP761551**	**KP761687**	**KP761519**	**KP761487**	**KP761616**	**KP761635**	**KP761581**	**KP761453**
8 ER- Birbisa	*C. abyssinica*	Birbisa, Gara Muleta Valley. (Ethiopia)	*Aldasoro & Alarcón 10424* (BC)	9° 16' 2.964" N, 41° 40' 45.7746" E	**KP761526**	**KP761659**	**KP761491**	**KP761460**	**KP761587**	**KP761630**	KP761556	**KP761424**
10 GC- Corcho	*C. canariensis*	Lanzarote-Cueva Corcho (Gran Canaria, Canary Islands, C.I.,)	*Aldasoro & Alarcón* 10355bis	28º 01' 31" N, 15º 35' 46" W	**KP761534**	**KP761672**	**KP761498**	**KP761465**	**KP761606**	**KP761642**	**KP761564**	**KP761430**
11 GC- Virgen	*C. canariensis*	Barranco de La Virgen (Gran Canaria, C. I.)	Caujapé J. 7325 & Mairal M. MM318	28° 3' 39.7296" N, 15° 35' 18.387" W	**KP761540**	**KP761676**	**KP761511**	**KP761471**	**KP761594**	**KP761654**	**KP761562**	**KP761438**
12 GC- Azuaje	*C. canariensis*	Barranco de Azuaje (Gran Canaria, C. I.)	Soto M. 7126	28° 5' 55.9032" N, 15° 34' 31.5768" W	**KP761542**	**KP761670**	**KP761500**	**KP761473**	**KP761597**	**KP761648**	**KP761560**	**KP761445**
13 GC- Moya	*C. canariensis*	Tilos de Moya– (Gran Canaria, C. I.)	Mairal M. MM246 & Caujapé J. 7202	28° 5' 15.3342" N, 15° 35' 34.3428" W	**KP761531**	**KP761674**	**KP761510**	**KP761478**	**KP761592**	**KP761655**	**KP761574**	**KP761436**
14 GC- El Sao	*C. canariensis*	Camino de El Sao a Presa de Los Pérez– (Gran Canaria, C. I.)	Caujapé J. 7173 & Mairal M. MM48	28° 3' 27.324" N, 15° 39' 35.157" W	**KP761539**	**KP761673**	**KP761509**	**KP761472**	**KP761598**	**KP761651**	**KP761566**	**KP761437**
15 GC- Guayedra	*C. canariensis*	Andenes de Guayedra– (Gran Canaria, C. I.)	Caujapé J. 7404	28° 3' 50.5008" N, 15° 41' 47.2524" W	**KP761532**	**KP761671**	**KP761503**	**KP761467**	**KP761608**	**KP761653**	**KP761559**	**KP761441**
16 TFE- Chamuscadas	*C. canariensis*	Monte Chamuscadas (Tenerife, C. I.)	*Aldasoro & Alarcón* 10321bis	28°32'13.68"N, 16°15'8.45"W	**KP761547**	**KP761665**	**KP761494**	**KP761468**	**KP761605**	**KP761641**	**KP761571**	**KP761426**
17 TFE- Bailadero	*C. canariensis*	El Bailadero (Tenerife, C. I.)	Mairal M. MM238	28° 33' 1.9908" N, 16° 12' 12.4956" W	**KP761528**	**KP761663**	**KP761508**	**KP761463**	**KP761596**	**KP761652**	**KP761573**	**KP761439**
18 TFE- Carnero	*C. canariensis*	Tope del Carnero (Tenerife, C. I.)	González A. 1000 & Mairal M. MM259	28° 33' 26.4702" N, 16° 18' 56.631" W	**KP761527**	**KP761661**	**KP761493**	**KP761462**	**KP761589**	**KP761639**	**KP761558**	**KP761444**
19 TFE- Badajoz	*C. canariensis*	Barranco de Badajoz– (Tenerife, C. I.)	Caujapé J. 9959	28° 18' 12.6036" N, 16° 25' 42.459" W	**KP761535**	**KP761679**	**KP761513**	**KP761480**	**KP761595**	–	**KP761578**	**KP761446**
20 TF- Ruiz	*C. canariensis*	Barranco Ruiz– (Tenerife, C. I.)	Caujapé J. 4672 & Mairal M. MM209 &	28° 23' 7.8252" N, 16° 37' 35.886" W	**KP761533**	–	**KP761512**	**KP761464**	**KP761604**	–	**KP761576**	**KP761442**

21 TFW- Adeje	*C. canariensis*	Adeje- Barranco Infierno– (Tenerife, C. I.)	Caujapé J. 9198 & Mairal M .MM 270	28° 7' 2.3844" N, 16° 42' 20.7828" W	**KP761536**	**KP761680**	**KP761506**	**KP761479**	**KP761609**	**KP761656**	**KP761577**	**KP761443**
22 TFW- Palmar	*C. canariensis*	El Palmar-Monte del Agua (Tenerife, C. I.)	*Aldasoro & Alarcón 10339* (BC)	28° 20' 30.681" N, 16° 49' 28.3152" W	**KP761544**	**KP761662**	**KP761504**	**KP761470**	**KP761607**	**KP761650**	**KP761567**	**KP761427**
23 TFW- Teno Alto	*C. canariensis*	Teno Alto (Tenerife, C. I.)	González A. 1001 & Mairal M. MM 283	28° 20' 35.7" N, 16° 52' 30.3024" W	**KP761543**	**KP761663**	**KP761495**	**KP761476**	**KP761590**	**KP761649**	**KP761572**	**KP761440**
24 GO- Palmita	*C. canariensis*	Barranco de la Palmita (La Gomera, C. I.)	Mairal M. MM025	28° 10' 1.9596" N, 17° 13' 7.197" W	**KP761546**	**KP761668**	**KP761505**	**KP761481**	**KP761600**	**KP761640**	**KP761563**	**KP761429**
25 GO- Tamargada	*C. canariensis*	Tamargada (La Gomera, C. I.)	Mairal M. MM009	28° 11' 4.7466" N, 17° 14' 22.8222" W	**KP761545**	**KP761677**	**KP761501**	**KP761482**	**KP761599**	**KP761643**	**KP761569**	**KP761428**
26 LPA- Tilos	*C. canariensis*	Los Tilos (La Palma, C. I.)	Mairal M. MM014	28° 47' 23.568" N, 17° 48' 6.8646" W	**KP761541**	**KP761678**	**KP761496**	**KP761469**	**KP761591**	**KP761657**	**KP761575**	**KP761432**
27 LPA- Agua	*C. canariensis*	Barranco del agua– (La Palma, C. I.)	Mairal M. MM013	28° 47' 37.3092" N, 17° 47' 33.7848" W	**KP761537**	**KP761669**	**KP761497**	**KP761474**	**KP761602**	**KP761644**	**KP761565**	**KP761434**
28 LPA- Galga	*C. canariensis*	Cubo de La Galga– (La Palma, C. I.)	Mairal M. MM026	28° 45' 49.7592" N, 17° 46' 29.7654" W	**KP761538**	**KP761667**	**KP761499**	**KP761477**	**KP761601**	**KP761646**	**KP761570**	**KP761435**
29 LPA- Barata	*C. canariensis*	Barranco La Barata– (La Palma, C. I.)	Mairal M. MM024	28° 48' 10.5798" N, 17° 47' 36.8154" W	**KP761530**	**KP761666**	**KP761507**	**KP761475**	**KP761603**	**XKP761647**	**KP761568**	**KP761431**
30 EH- El Hierro	*C. canariensis*	Hoya del Pino (El Hierro, C. I.)	Mairal M. MM017	27° 43' 50.1702" N, 18° 1' 35.7234" W	**KP761529**	**KP761675**	**KP761502**	**KP761466**	**KP761593**	**KP761645**	**KP761561**	**KP761433**

Platycodoneae	**Taxon**	**Locality (Country)**	**Coll. & Coll. # (Herbarium)**	**Geographic coordinates**	***psb*J-*pet*A**	***pet*B-*pet*D**	***trn*L-*trn*F**	***trn*S-*trn*G**	**ITS**
Outgroup	*Lobelia nana*	Sierra de Achala, Condola (argentina)	J.J. Aldasoro 6780		**KP761521**	FN397074.1	**KP761617**	**KP761622**	AY350629.1
Outgroup	*Cyphia subtubulata*				–	FN397061.1	–	–	–
Outgroup	*Campanula jacobaea*	*Santo Antâo, Cova*– (Cape Verde)	Mairal M. MM183	17º06' 0" N 25º04'02" W	–	JX914762.1	–	**KP761623**	**KP761454**
Outgroup	*Codonopsis benthamii*	Cultivated in Edrom Nursery, from– E Sikkim (Scotland)	GWJ9352		**KP761523**	**KP761458**	**KP761619**	**KP761627**	**KP761455**
Outgroup	*Campanumoea javanica*	Sichuan province, Emei Mountain (China)	4028 X.J. LI		**KP761522**	**KP761457**	**KP761618**	**KP761626**	DQ889459.1
Outgroup	*Cyananthus lobatus*	Cultivated in Edrom Nursery (Scotland)	018907–71386_ CC4634		–	FN397058.1	JN851188.1	**KP761628**	**KP761456**
Outgroup	*Cyclocodon lancifolium*	Sichuan province,– Emei Mountain (China)	X.J. LI s.n.		**KP761525**	FN397059.1	**KP761621**	**KP761625**	EF206701.1
Outgroup	*Platycodon grandiflorum*	Cultivated in Real Jardín Botánico de Madrid (Spain)			**KP761524**	FN397087.1	**KP761620**	**KP761624**	AB699584
Outgroup	*Ostrowskia magnifica*	Kondara Gorge, Botanical Zoologic & Genetic Science Center of Varzob District (Tajikistan)	J.J. Aldasoro 23000	38º 48' 37.9"N, 68º 48' 45.2" E	**KP761490**	**KP761459**	**KP761586**	**KP761629**	**KP761423**

Table S2. Primers used for PCR amplification and sequencing.

Nº primer	**Primer name**	**Sequence 5´-3´**	**Marker (reference study)**
1	*rpl*32-F	CAG TTC CAA AAA AAC GTA CTT C	*rpl*32-*trn*L (Shaw *et al*. 2007)
2	*trn*L	CTG CTT CCT AAG AGC AGC GT	*rpl*32-*trn*L (Shaw *et al*. 2007)
3	*3´trn*V	GTC TAC GGT TCG ART CCG TA	*3´trn*V-*ndh*C (Shaw *et al*. 2007)
4	*ndh*C	TAT TAT TAG AAA TGY CCA RAA AAT ATCATA TTC	*3´trn*V-*ndh*C (Shaw *et al*. 2007)
5	*psb*J	ATA GGT ACT GTA RCY GGT ATT	*psb*J-*pet*A (Shaw *et al*. 2007)
6	*pet*A	AAC ART TYG ARA AGG TTC AAT T	*psb*J-*pet*A (Shaw *et a*l. 2007)
7	*pet*B-*1365*F	TTG ACY CGT TTT TAT AGT TTA C	*pet*B-*pet*D (Borsch *et al*. 2009)
8	*pet*D-*738R*	AAT TTA GCY CTT AAT ACA GG	*pet*B-*pet*D (Borsch *et al*. 2009)
9	*trn*G	GAA CGA ATC ACA CTT TTA CCA C	*trn*S-*trn*G-IGS (Shaw *et al*. 2005)
10	*trn*S	GCC GCT TTA GTC CAC TCA GC C	*trn*S-*trn*G-IGS (Shaw *et al*. 2005)
11	*trn*G2G	GCG GGT ATA GTT TAG TGG TAA AA	*trn*G (Shaw *et al*. 2007)
12	*trn*G(UUC)	GTA GCG GGA ATC GAA CCC GCA TC	*trn*G (Shaw *et al*. 2007)
13	*trn*LF-c	CGA AAT CGG TAG ACG CTA CG	*trn*L-*trn*F (Shaw *et al*. 2005)
14	*trn*LF-f	ATT TGA ACT GGT GAC ACG AG	*trn*L-*trn*F (Shaw *et al*. 2005)
15	ITS1A	GGA AGG AGA AGT CGT AAC AAG G	ITS (White *et al*. 1990)
16	ITS4	TCC TCC GCT TAT TGA TAT GC	ITS (White *et al*. 1990)

Table S3. Summary of results from the congruence analysis among chloroplast markers (see text). Above: alternative topologies supported by BUCKy with or without *rpl32* using the *Platycodoneae* cpDNA dataset; numbers in bold represent taxa as 1 *Lobelia*, 2 *C. abyssinica*, 3 *C.eminii*, 4 *Codonopsis*, 5 *C. canariensis*, 6 *Platycodon,* 7 *Cyclocodon*, and 8 *Ostrowskia*), followed by the sample concordance factors. Below: rate multipliers and tree length (TL) estimates for each gene as obtained in a partitioned analysis in MrBayes of the *Canarina* cpDNA dataset.

	Markers	**Topologies**
Bucky	cpDNA complete (including *rpl*32)	((((**1**:1,**4**:1):0.998,**6**:1):0.977,**7**:1):1,((**2**:1,**3**:1):0.991,**5**:1):1,**8**:1)
	cpDNA (excluding *rpl*32)	((((**1**:1,**4**:1):1,**6**:1):0.815,**7**:1):1,(**2**:1,(**3**:1,**5**:1):0.896):1,**8**:1)

	_rpl_32-_trn_L	**_3´trn_V-_ndh_C**	**_psb_J-_pet_A**	**_pet_B-_pet_D**	**_Trn_L-_trn_F**	**_Trn_S-_trn_G**	**_trn_G intron**
Rate multipliers	3.19	0.855	0.813	0.697	0.672	0.681	0.333
TL (tree length)	1.374	–	–	0.448	–	0.668	–

Table S4. Model likelihood estimators obtained using the Path sampling (PS) and Stepping-Stone (SS) sampling methods implemented in BEAST.

	Clock	**Prior**	**Pathlikelihood.delta**
Platycodoneae	Strict	Yule	-9219
	Strict	Birth-Death	-9452
	UCLD	Yule	-9282
	UCLD	Birth-Death	-9252
Canarina	Strict	Yule	-11193
	Strict	Birth-Death	-11196
	UCLD	Yule	-11171*
	UCLD	Birth-Death	-11328*

*these runs could not reach convergence (poor mixing)

Table S5. Geographical coordinates used in the Ecological Niche Modelling of *Canarina*.

Canarina canariensis		***Canarina eminii***	
Latitude	Longitude	Latitude	Longitude
28.05685282	-15.58776829	-2.25	28.68333
28.08216502	-15.59768136	0.21	29.5
28.0510065	-15.589861	7.56	39.09
28.01069557	-15.52732573	-9.13	33.71
28.04484808	-15.58348836	9.02	36.29
28.0717	-15.5383	-8.98	33.6
28.09024162	-15.60081615	-9.17	33.88
28.55015111	-16.20409333	-2.13333	28.83333
28.33081506	-16.8188953	1.1	34.24
28.5167	-16.4	-9.18	33.65
28.54805939	-16.21842849	8.58333	38.01667
28.04431533	-15.48392558	-9	33.7
28.45153473	-16.41565582	-2.23333	28.78333
28.55	-16.22	7.15	38.71667
28.53295921	-16.26375579	9.03333	36.63333
28.355173	-16.812436	7.55	39.08
28.5463837	-16.18735977	7.5	36.52
28.1836473	-17.271695	-3.14	29.34
28.1826325	-17.2417145	9.05	38.56667
28.1581688	-17.1934095	-3.18	29.6
28.1560741	-17.2062944	1.084138889	37.04888889
28.16269789	-17.22366673	1.068055556	37.24805556
28.1536398	-17.1709597	6.700833333	39.72638889
28.150631	-17.1918828	6.081666667	38.2425
28.1494649	-17.1919593	1.328888889	34.41611111
28.1536398	-171709597	0.764444444	36.74305556
28.556337	-16.264902	0.359722222	29.97222222
28.1430218	-17.216242	9.018375135	38.772202
28.82516335	-17.8866558	9.080695092	38.772187
28.01944	-15.55306	1.912099093	30.617505
28.61919366	-17.78626204	0.232984815	29.839975
27.77990979	-17.98717071	0.332078311	29.824007
27.70821803	-17.96491689	9.083332999	38.583333
28.67504035	-17.79049071	-2.7290055	28.751193
28.51267699	-16.35197031	3.778690614	32.917296
28.55	-16.26667	-1.60159591	27.978259
28.02527	-15.5961	1.753087361	34.686349
28.061036	-15.58844	1.246711426	34.381467
28.098862	-15.575438	1.245984013	34.3974488
28.087592	-15.592873	1.24671151	34.381466
28.05759	-15.65976	1.196913283	34.670038

28.064028	-15.696459	2.515355305	34.765731
28.562944	-16.166788	0.203156266	34.907933
28.550553	-16.203471	-9.111647597	33.5294294
28.55735	-16.31573	0.615923341	35.5084578
28.303501	-16.42846	-1.212228546	36.8253984
28.385507	-16.626635	-0.141257009	37.3078994
28.117329	-16.705773	6.858020949	36.1962449
28.341855	-16.824532	-0.440115688	36.6739351
28.343249	-16.875084	6.04625072	37.602662
28.167211	-17.2186658	-0.630463274	35.8288148
28.184651	-17.2396727	1.211568243	35.5559462
28.78988	-17.8019068	1.10287324	37.0550245
28.793697	-17.792718	1.098176303	37.7105833
28.76382	-17.77493		
28.8029388	-17.7935598		
27.7306028	-18.0265898		
28.30796375	-16.76491462		
28.56456788	-16.16696242		
28.56026612	-16.2841235		
28.16746387	-17.21712289		
28.32851682	-16.70319362		
28.35624536	-16.79834381		
28.37539544	-16.62823676		
28.12817544	-16.702372		
28.3527617	-16.80702		
28.11533631	-15.5949431		

Table S6. Mean ages and 95% HPD confidence intervals for the different BEAST analyses represented in Fig. 3a-b-c: a) *Platycodoneae* dataset; b) *Canarina* dataset, standard dating; c) nested dating of the *C. eminii* and *C. canariensis* datasets. Node numbers correspond to those shown in Fig. S4a-b-c.

a)

Node	Mean age	Confidence Interval
1	52.97	40.13-66.73
2	40.09	20.59 - 57.13
3	41.89	28.56 - 54.67
4	29.13	18.17 - 41.95
5	16.50	7.45 - 27.39
6	8.83	2.93 - 17.09
7	22.14	12.72 - 33.30
8	20.82	11.80 - 31.67
9	13.76	6.57 - 21.74
10	8.19	3.31 - 14.13
11	6.47	2.11 - 11.88

b)

Node	Mean age	Confidence Interval
1	14.13	7.79 - 20.73
2	11.64	6.22 - 17.33
3	7.18	3.78 - 10.85
4	6.48	3.44 - 9.82
5	1.28	0.605 - 2.092
6	0.41	0.127 - 0.822
7	0.16	0.018 - 0.406
8	0.92	0.424 - 1.551
9	0.70	0.277 - 1.253
10	0.06	0 - 0.202
11	0.81	0.345 - 1.387
12	0.41	0.161 - 0.759
13	0.14	0.020 - 0.338
14	0.28	0.097 - 0.540
15	0.03	0 - 0.13
16	0.59	0.23 - 1.05
17	0.29	0.089 - 0.576
18	0.10	0.049 - 0.373
19	0.03	0 - 0.17

c)

Node	Mean age	Confidence Interval
1	52.66	39.52 - 65.99
2	39.11	19.21 - 57.48
3	42.46	29.08 - 54.83
4	29.69	17.84 - 42.62
5	16.74	7.28 - 27.95
6	8.75	2.5 - 17.15
7	22.33	12.04 - 33.62
8	20.93	11.53 - 31.99
9	13.93	6.73 - 22.84
10	8.38	3.04 - 14.67
11	6.5	1.75 - 12.16
12	1.76	0.532 - 3.668
13	0.1	0.001 - 0.376
16	0.36	0.062 - 0.917
17	0.14	0.011 - 0.42
18	0.759	0.247 - 1.580
19	0.311	0.078 - 0.728
20	0.063	0.005 - 0.205
21	0.182	0.033 - 0.445
22	0.056	0.005 - 0.168
23	0.008	0 - 0.045
24	0.014	0 - 0.067
25	0.38	0.094 - 0.891
26	0.17	0.04 - 0.42
27	0.01	0 - 0.08

SUPPLEMENTARY FIGURES

Figure S1. Bayesian Majority-Rule consensus trees obtained by MrBayes from the single-gene analyses of the Platycodoneae dataset. Numbers next to branches indicate Bayesian posterior probability values (PP).

Figure S2. Bayesian Majority-Rule consensus trees inferred from the single-gene analyses of the *Canarina* dataset. Numbers above branches indicate Bayesian credibility values (PP); numbers below branches indicate maximum likelihood bootstrap support values.

Figure S3. Saturation plots for the single-gene nuclear (ITS) and chloroplast markers, showing the uncorrected pairwise distances (*p* distance) against corrected maximum likelihood distances (ML distance) derived in PAUP using the appropriate model of substitution.

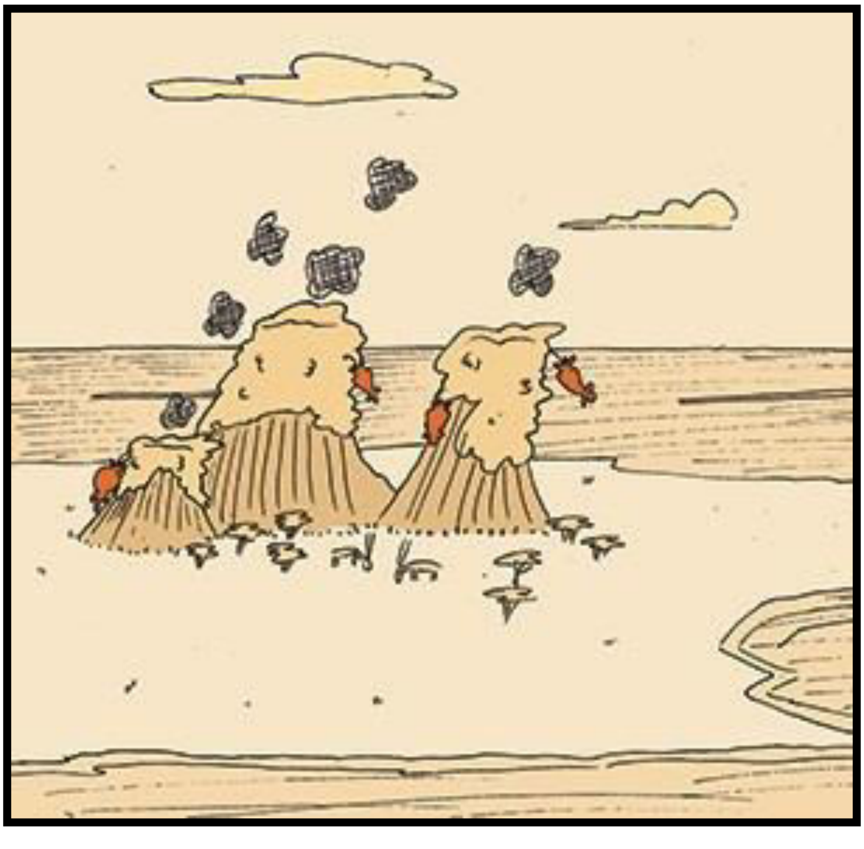
Figure S4. Maximum clade credibility (MCC) trees obtained from different BEAST analyses, showing nodes with mean ages and 95% HPD confidence intervals (values specified in Table S4): a) Platycodoneae dataset; b) standard dating of the *Canarina* dataset; c) nested dating analysis of all three linked datasets: Platycodoneae (left) and population-level *C. eminii* and *C. canariensis* (right).

Figure S5. Bayesian phylogeographic analysis of the *C. eminii* population-level dataset (nested dating approach) using an alternative coding of the geographic areas:(Elgon and Ruwenzori considered as different OTUs). The tree represents the MCC tree from BEAST with colored branch lengths representing the ancestral range with the highest posterior probability for each lineage (see legend). The nodal pie charts represents alternative ancestral ranges according to their posterior probability value.

Figure S6. Bayesian Majority-Rule consensus tree obtained by MrBayes from the *Canarina* concatenated chloroplast and nuclear dataset rooted with *Ostrowskia* (ITS, *psb*J-*pet*A, *trn*L-*trn*F, *pet*B-*pet*D, *trn*S-*trn*G, *trn*V-*ndh*C). Numbers above branches indicate Bayesian credibility values (PP); numbers below branches indicate maximum likelihood bootstrap support values.

SUPPLEMENTARY DATA
CHAPTER 2

Geographic barriers and Pleistocene climate change shaped patterns of genetic variation in the Eastern Afromontane Biodiversity Hotspot



SUPPLEMENTARY INFORMATION

Article title

Geographic barriers and Pleistocene climate change shaped patterns of genetic variation in the Eastern Afromontane biodiversity hotspot.

Authors

Mario Mairal*[1], Isabel Sanmartín[1], Alberto Herrero[1], Lisa Pokorny[2], Pablo Vargas[1], Juan J. Aldasoro†[3], and Marisa Alarcón†[3]

[1] *Real Jardín Botánico (RJB-CSIC), 28014 Madrid, Spain*

[2] *Royal Botanic Gardens, Kew (RBGK), Richmond, Surrey, TW9 3DS, UK*

[3] *Instituto Botánico de Barcelona (IBB-CSIC-ICUB), 08038 Barcelona, Spain*

SUPPORTING INFORMATION

Appendix S1. Expanded material and Methods

Appendix S2. Supplemental Tables and Figures

Appendix S3. Supplemental references

Appendix S1 Expanded Material and Methods

S1.1 DNA amplification, sequencing and alignment DNA

DNA of 72 individuals was extracted using the DNeasy Plant Mini Kit (QIAGEN Inc., California, USA) according to the manufacturer's instructions, from 18-23 mg of silica-gel dried leaves obtained from the fresh plant tissue collected from the field expeditions. PCR amplifications were performed using an Eppendorf Mastercycler Epgradient S (Westbury, NY). Table S3.3 shows the primer sequences for each region. The three intergenic spacers were amplified using the following PCR settings: 1) for *rpl*32-*trn*L$_{\mathrm{UAG}}$: 1 min pretreatment at 95°C, 30 cycles of 1 min at 96°C, 1 min at 52°C, and 2 min at 72°C; 2) for *trn*S$_{\mathrm{GCU}}$-*trn*$_{\mathrm{GUCC}}$: 1 min pre-treatment at 96°C, 30 cycles of 45 s at 96°C, 1 min at 54°C, and 1 min 30s at 72°C; and 3) for *pet*B1365–*pet*D738: 1 min pre-treatment at 96°C, 30 cycles of 1 min at 95°C, 2 min at 49°C, and 2 min at 72°C. All these reactions ended with a final extension step for 10 min at 72°C. Amplified products were treated with ExoSAP-IT (USB Corporation, Ohio) and submitted to Macrogen Inc. (Seoul, South Korea) for sequencing. We had difficulties to amplify some of the herbarium specimens, for this we developed internal primers and performed nested-PCRs. The resulting sequence data were assembled and edited using SeqEd v. 1.0.3 software (Applied Biosystems, California, USA). Sequences for each region were aligned using MAFFT software (Katoh *et al*., 2005). Sequences were manually adjusted where necessary by following alignment rules described in Kelchner (2000), and checked through the use of Bioedit software (Hall, 1999).

S1.2 Divergence time analyses

Model-fit of nucleotide substitution models was assessed via the Bayesian Information Criterion (BIC) as implemented in jModelTest 0.1.1 (Posada, 2008). The higher-level data set was calibrated with fossil-derived secondary age estimates (see Mairal *et al.* 2015), while the tree prior was unlinked to apply a coalescent constant size model to the population-level data set and a stochastic birth–death (Yule) prior to the species-level one (Mairal *et al.* 2015). The clock model was set to an uncorrelated log-normal prior to accommodate the change in mutation rate from species to populations, with a uniform distribution for the ucld.mean (10^{-4}– 10^{-1}) and a default exponential distribution for the ucld.stdev; the substitution model was set to GTR+G. Two MCMC chains were run for 50 million generations, sampling parameters every 1000 generation. We used Tracer v1.6 (Rambaut *et al.* 2007) to verify the following: whether a stationary distribution was attained, whether there was convergence among chains and whether effective sample sizes (ESS values) were >200 for all parameters. A 10% burn-in of the sampled populations was discarded (5 million). Post-burn-in trees were summarized into a maximum clade credibility tree using TREEANNOTATOR v.1.6.1, with mean values and 95% credible intervals for nodal ages, and were visualized in FIGTREE 1.3.1 (Rambaut & Drummond 2009).

S1.3 Bayesian discrete phylogeographic analyses

This method implements a continuous-time Markov Chain (CTMC) process to model biogeographic evolution, in which the discrete states correspond to geographic locations of the sequences and transition rates between states to migration rates or dispersal between areas (Ronquist & Sanmartín, 2011). Migration rates between areas and the geodispersal rate scaler were modelled using default gamma prior distributions

(Lemey *et al.*, 2009). Two replicate searches of 20 million generations each, sampling every 1000th generation, were combined in TreeAnnotator, after removing the 10% burn-in, to produce a maximum clade credibility (MCC) tree. Bayesian stochastic search variable selection (BSSVS, Lemey *et al.*, 2009) was used to identify the rates (colonization routes) that are best supported by the data.

Appendix S2 Supplemental Tables and Figures.

Table S2.1. List of the visited localities for this study and description of the sampling.

Country	**Hedberg collections**	***C. eminii* collections (X: explored, but not found)**	**Commentaries (not explored not explored)**
Ethiopia	Damot near source of Abay river	-	deforested
	Burye, road to Lake Tana	Dembecha (north of Debre Marcos), A9979	deforested, a few scattered trees contain plants of *C. eminii*
		Gifta (north of Debre Marcos), A9982	
	Shoa: Kachise (near Kachissy)	x	deforested
	Shoa, Entotto	x	deforested
	Shoa: Addis Ababa	x	deforested
	Shoa: Little Akaki river	x	deforested
	Shoa: Mannaghescia (Mannagascià)	x	deforested
	Shoa: Holeta (Oletta), valle di Metcha Coriccia	x	deforested
	Borana people, N of Agere Mariam	Agere Maryan, Yirga Chefé Road, A10060	*C. eminii* abundant
	Bale	Harenna Forest, Negele to Goba road, A10317 and A10322	*C. eminii* abundant
	Gallanot exploredSidamo: Welega (Wallega), Kidami	not explored	not explored
	Gallanot explored Sidamo: Shisha, Lake Uombo (Wombo)	not explored	not explored
	Gallanot explored Sidamo: Cencia	not explored	not explored
Uganda	Ruwenzori, Kangasaba	x	sparsely forested
	Ruwenzori, Mahoma Valley near Nyabitabu (Nyakalengija)	Ruwenzori Mts, Nyabitaba track, Sánchez Meseguer & Aldasoro 18	*C. eminii* abundant
	Mt. Elgon, W. slope between Bulambuli y Butandiga	Mt. Elgon, Sánchez Meseguer & Aldasoro 69	*C. eminii* abundant
	Mt. Elgon, Buginyanya	Mt. Elgon, Sánchez Meseguer & Aldasoro 71	*C. eminii* present
	N of Mt. Elgon at Kyessweri R. Sipi and Kabururon R. Sundet (Sebei)	Mt. Elgon, Sánchez Meseguer & Aldasoro 75	*C. eminii* present
	Mbale dist: Mt Elgon, Bumoni	not explored	not explored
	Bugisu distr., Bulago	not explored	not explored
	Karamoja Distr: Rom Mts	not explored	not explored, entry forbidden
	Karamoja Distr: Langia Mts	not explored	not explored, entry forbidden

	Karamoja Distr: Monte Kadam	not explored	not explored, entry forbidden
	Karamoja Distr: Monte Moroto	not explored	not explored, entry forbidden
South Sudan	Imatong Mts.	not explored	not explored, entry forbidden
	Lomwaga Mts.	not explored	not explored, entry forbidden
Kenya	Kapolet (Kaboret) forest	not explored	not explored
	Cherangani Hills	x	sparsely forested
	Elgeyo Escarpment	x	sparsely forested
	Kapsabet, Yala river	x	sparsely forested
	Nyanza Province	x	sparsely forested
	Aberdare Range	Aberdare Mts, Sánchez Meseguer & Aldasoro 103	*C. eminii* abundant
	Aberdare (east), Tuso	x	sparsely forested
	Dimbilil river, SW Mau	not explored	not explored
	Mt. Kenya, Embu distr., Thiba camp	x	sparsely forested
	Mt. Kenya, Kirinyaga distr.	x	sparsely forested

Despite intensive sampling effort (two field expeditions in 2009, 2015 to Ethiopia and one in 2010 to Kenya-Uganda), we failed to find *C. eminii* in many of the historically recorded localities (Hedberg, 1961). During the present field work, the number of localities where *C. eminii* grow in the northern Ethiopian Plateaus, and the Elgon Mt. was much smaller than that reported by Hedberg (1961). The highlands of northern Ethiopia as well as many parts of Kenya, now given over to cultivation or pasture, are largely devoid of forest vegetation. We noted that much of the forests in these localities were highly degraded, having either completely disappeared or being threatened by the advancement of extensive agriculture, cattle rising, forestry and other human activities. Although that species appeared there in previous studies, currently the absence of forested patches and adequate phorophytes likely impeded this species to survive there. All these difficulties hampered a deeper sampling of *C. eminii*. Moreover, some of the localities could not be visited due to security concerns, such as: South Sudan, eastern Congo, Burundi, South Sudan, and eastern Congo borders were not allowed to be visited during our sampling period. Besides, there are some places in north of Uganda where the authorities prohibit any entry.

Table S2.2. Plastid sequences of *C. eminii* newly generated for this study: localities and geographical coordinates, sample collector, and GenBank accessions number. Voucher specimens of all sampled species and populations were deposited in the herbarium of the *Instituto Botánico de Barcelona*- CSIC (BC).

Voucher specimen *Canarina eminii* Asch. ex Schweinf.	**trnS–trnG accession number**		***pet*B–*pet*D accession number**		***rpl*32-*trn*L^{UAG} accession number**
Ethiopia, Gifta, Debre Markos to Bahir Da Road, N of Bure, 09/08/2009, 2478m, 10º 50' 29'' N, 37º 02' 56'' E, *Aldasoro & Alarcón 9982* (BC)	KF028817	KM189335	KF028856	KM189234	KM189267
	KF028818	KM189336	KF028857	KM189235	KM189268
	KF028819	KM189337	KF028858	KM189236	KM189269
	KM189334	KM189338	KM189233	KM189237	KM189270
Ethiopia, between Dembecha and Bure, Debre Markos to Bahir Dar Road, 08/08/2009, 2493m, 10°40'50'' N, 37°14'53'' E, *Aldasoro & Alarcón 9979* (BC)	KF028814	KM189331	KF028853	KM189230	KM189260
	KF028815	KM189332	KF028854	KM189231	KM189261
	KF028816	KM189333	KF028855	KM189232	KM189262
	KM189330		KM189229		KM189263
Ethiopia, Agere Maryan, Yirga Chefé Road, 14/08/2009, 2434m, 6º 04' 54'' N, 38º 14' 33'' E, *Aldasoro & Alarcón 10060* (BC)	KF028820	KF028825	KF028859	KF028863	KM189275
	KF028821	KF028826	KF028860	KF028864	KM189276
	KF028822	KM189339	KF028861	KF028865	KM189277
	KF028823	KM189340	KF028862	KM189238	KM189278
	KF028824			KM189239	KM189279

Ethiopia, Harenna Forest, Negele to Goba Road, 19/08/2009, 2182m, 6º 42' 03" N, 39º 43' 35" E, *Aldasoro & Alarcón 10322* (BC)	KF028827	KF028831	KF028866	KF028870	KM189284
	KF028828	KF028832	KF028867	KF028871	KM189285
	KF028829	KM189341	KF028868	KM189240	KM189286
	KF028830		KF028869		KM189287
Ethiopia, Harenna Forest, Negele to Goba Road, 19/08/2009, 1939m, 6º 39' 3"N, 39º 43' 57"E, *Aldasoro & Alarcón 10317* (BC)	KF028833		KF028872		KM189290
Uganda, Mt. Elgon 31/08/2010, 2230 m, 1º 19' 44" N, 34º, 24' 58" E, *Sánchez-Meseguer & Aldasoro 69* (BC)	KF028834	KM189345	KF028873	KM189244	KM189292
	KF028835	KM189346	KF028874	KM189245	KM189293
	KF028836	KM189347	KF028875	KM189246	KM189294
	KF028837	KM189348	KF028876	KM189247	KM189295
	KF028838	KM189349	KF028877	KM189248	KM189296
	KF028839	KM189350	KF028878	KM189249	KM189297
	KM189342	KM189351	KM189241	KM189250	KM189298
	KM189343	KM189352	KM189242	KM189251	KM189299
	KM189344		KM189243		KM189300
Uganda, Ruwenzori Mts, 27/08/2010, 2599 m, 0º 21' 35" N, 29º 58' 20" E, *Sánchez-Meseguer & Aldasoro18* (BC)	KF028840	KF028846	KF028879	KF028885	KM189309
	KF028841	KM189353	KF028880	KM189252	KM189310
	KF028842	KM189354	KF028881	KM189253	KM189311
	KF028843	KM189355	KF028882	KM189254	KM189312
	KF028844	KM189356	KF028883	KM189255	KM189313
	KF028845	KM189357	KF028884	KM189256	KM189314
Kenya, Aberdare Mts, 03/09/2010, 2467 m, 0º 45' 52" N, 36º 44' 35" E, *Sánchez-Meseguer & Aldasoro103* (BC)	KF028847	KF028852	KF028886	KF028891	KM189321
	KF028848	KM189358	KF028887	KM189257	KM189322
	KF028849	KM189359	KF028888	KM189258	KM189323
	KF028850	KM189360	KF028889	KM189259	KM189324
	KF028851		KF028890		KM189325
Kenya, K2 N. Cherangani Hills Kaibwibich. IS377	XXX		XXX		XXX
Tanzania, Mbeya: Rungwe District. Livingstone Mountains; foot trail above Bumbigi. IS379	XXX		XXX		XXX
Tanzania, Mbeya: Rungwe District. Road between Igoma and Kitulo on south slope of Mporoto Mountains. IS378	XXX		XXX		XXX
Tanzania, Tukuyu district. IS382	XXX		XXX		XXX

Tanzania, Rungwe District. Mwakeleli, vincinity of Mwatesi River. IS381	XXX	XXX	XXX
Uganda, Mount Elgon, Sasa trail. IS376	XXX	XXX	XXX
Rwanda, Gikongoro Prefecture: Route Butare-Cyangugu, vers km 60. IS384	XXX	XXX	XXX
Burundi, Teza. IS386	XXX	XXX	XXX
Malawi, Misuku District, Sllindi Forest, Misuku Hills. IS 383	XXX	XXX	XXX

Table S2.3. Specific primer pairs used in the amplification and sequencing of plastid markers in *C. eminii*.

Primer name's	**Primer Sequence (5'–3')**	**Author primer's**
***trn*S^{GCU} -F**	GCC GCT TTA GTC CAC TCA GC	Hamilton, 1999
***trn*G^{UCC} -R**	GAA CGA ATC ACA CTT TTA CCA C	Hamilton, 1999
***pet*B^{1365}-F**	TTG ACY CGT TTT TAT AGT TTA	Löhne and Borsch, 2005
***pet*D^{738}-R**	AAT TTA GCY CTT AAT ACA GG	Löhne and Borsch, 2005
***rpl*32-F**	CAG TTC CAA AAA AAC GTA CTT C	Shaw *et al.*, 2007
***trnL*^{UAG}-R**	CTG CTT CCT AAG AGC AGC GT	Shaw *et al.*, 2007

Table S2.4. Hierarchical analysis of molecular variance (AMOVA) for *C. eminii* based on AFLP data at different levels.

AMOVA for all populations						
Groupings	**Source of variation**	**d.f.**	**Sum of squares**	**Variance components**	**Percentage of variation**	**Φ Statistics**
1 group	Among populations	5	2966.483	51.70776	40.42	-
	Within populations	55	4192.042	76.21894	59.58	Φ_{ST}= 0.4042
	Total	**60**	**7158.525**	**127.92670**		
2 groups: Eastern Rift-Western Rift:	Among groups	1	712.698	5.17812	3.98	Φ_{CT}= 0.0398
1) **Western Rift**: Gifta, Dembecha, Mt. Elgon and Ruwenzori Mts.	Among populations	4	2253.785	48.69257	37.43	Φ_{SC}= 0.3898
2) Eastern Rift: Agere Maryan, Harenna Forest and Aberdare Mts.	Within populations	55	4192.042	76.21894	58.59	Φ_{ST}= 0.4141
	Total	**60**	**7158.525**	**130.08963**		
3 groups:	Among groups	2	1443.203	3.30884	2.57	Φ_{CT}=

						0.0257
1. Abyssinian Plateau (Gifta and Dembecha)	Among populations	3	1523.280	49.05.202	38.15	Φ_{SC}= 0.3916
2. Agere Maryan, Harenna Forest and Aberdare Mts	Within populations	55	4192.042	76.21894	59.28	Φ_{ST}= 0.4072
3. Elgon and Rwenzori Mts.	**Total**	**60**	**7158.525**	**128.57979**		
4 groups:						
1. Abyssinian Plateau (Gifta and Dembecha)	Among groups	3	2187.769	14.13032	10.89	Φ_{CT}= 0.1089
2. Agere Maryan, Harenna Forest and Aberdare Mts.	Among populations	2	778.714	39.34721	30.34	Φ_{SC}= 0.3405
3. Elgon Mts.	Within populations	55	4192.042	76.21894	58.77	Φ_{ST}= 0.4123
4. Rwenzori Mts	**Total**	**60**	**7158.525**	**129.69647**		

AMOVA for Eastern Rift populations

Groupings	**Source of variation**	**d.f.**	**Sum of squares**	**Variance components**	**Percentage of variation**	**Φ Statistics**
1 group in the Eastern Rift:	Among populations	2	778.714	38.93272	32.87	
Harar massif (Agere Maryan, Harenna Forest) and Aberdare Mts.	Within populations	21	1669.869	79.51757	67.13	Φ_{ST}= 0.3287
	Total	**23**	**2448.583**	**118.45029**		
2 groups in the Eastern Rift	Among groups	1	395.356	4.91574	4.09	Φ_{CT}= 0.0409
1. Harar massif: Agere Maryan and Harenna Forest	Among populations	1	383.358	35.87003	29.82	Φ_{SC}= 0.3109
2. Aberdare Mts.	Within populations	21	1669.869	79.51757	66.10	Φ_{ST}= 0.3390
	Total	**23**	**2448.583**	**120.30335**		

AMOVA for Western Rift populations

Groupings	**Source of variation**	**d.f.**	**Sum of squares**	**Variance components**	**Percentage of variation**	**Φ Statistics**
1 group Western:	Among populations	2	1475.071	55.03159	42.59	
Debre Markos, Elgon and Rwenzori Mts.	Within populations	34	2522.173	74.18155	57.41	Φ_{ST}= 0.42590
	Total	**36**	**3997.243**	**129.21314**		
2 groups in the Western Rift	Among groups	1	730.505	-11.89280	-9.42	Φ_{CT}= -0.0942
1. Abyssinian plateau: Debre Markos	Among populations	1	744.565	63.99118	50.67	Φ_{SC}= 0.4631
2. Elgon and Rwenzori Mts.	Within populations	34	2522.173	74.18155	58.74	Φ_{ST}= 0.4126
	Total	**36**	**3997.243**	**126.27993**		

AMOVA for Ethiopian plateaus populations

Groupings	**Source of variation**	**d.f.**	**Sum of squares**	**Variance components**	**Percentage of variation**	**Φ Statistics**
2 groups Eastern Plateau - Western Plateau:	Among groups	1	726.589	12.54985	9.28	Φ_{CT}= 0.0928
1. Abyssinian Plateau: Gifta and Dembecha	Among populations	1	383.358	34.89585	25.81	Φ_{SC}= 0.2845
2. Harar Plateau: Agere Maryan and Harenna Forest	Within populations	30	2633.083	87.76944	64.91	Φ_{ST}= 0.3509
	Total	**32**	**3743.030**	**135.21515**		

Table S2.5. Migration routes supported by Bayes Factor comparisons with a cut off value >3 using Bayesian stochastic search variable selection (BSSVS, Lemey *et al.*, 2009).

Migration routes	Bayes Factor
Harenna Forest - Aberdare Mts.	7.6760
Rwenzori - Mt. Elgon	5.3979
Mt. Elgon - Rwenzori	4.6847
Aberdare Mts - Harenna Forest	4.0291
Rwenzori - Southern Mountain sky-islands	3.8762
Rwenzori - Abyssinian massif	3.6395
Mt.Elgon - Southern Mountain sky-islands	3.3288
Southern Mountain sky-islands - Rwenzori	3.1858
Southern Mountain sky-islands - Mt. Elgon	3.0084

Table S2.6. Pairwise F_{ST} between populations of *C. eminii* with nuclear data (below) and geographic distances in km (above).

	Gifta	Dembecha	Agere Maryan	Harenna Forest	Mt. Elgon	Rwenzori Mts.	Aberdare Mts.
Gifta	-	28	546	545.9	1091.84	1398.95	1115.13
Dembecha	0.0801	-	520.37	517.78	1080.55	1396.78	1098
Agere Maryan	0.2157	0.1593	-	177.87	675.86	1112.88	611.1
Harenna Forest	0.3028	0.2352	0.1706	-	836.87	1287.61	735.28
Mt. Elgon	0.3555	0.2877	0.2985	0.3698	-	503.09	266.4
Rwenzori Mts.	0.3306	0.2627	0.2887	0.2174	0.3779	-	752
Aberdare Mts.	0.3041	0.2390	0.2058	0.2247	0.3302	0.3142	-

Figure S2.1.

Nested analysis of the Platycodoneae dataset. Numbers above branches indicate mean ages and numbers below branches indicate Bayesian PP. For more details see Mairal *et al.* 2015a.

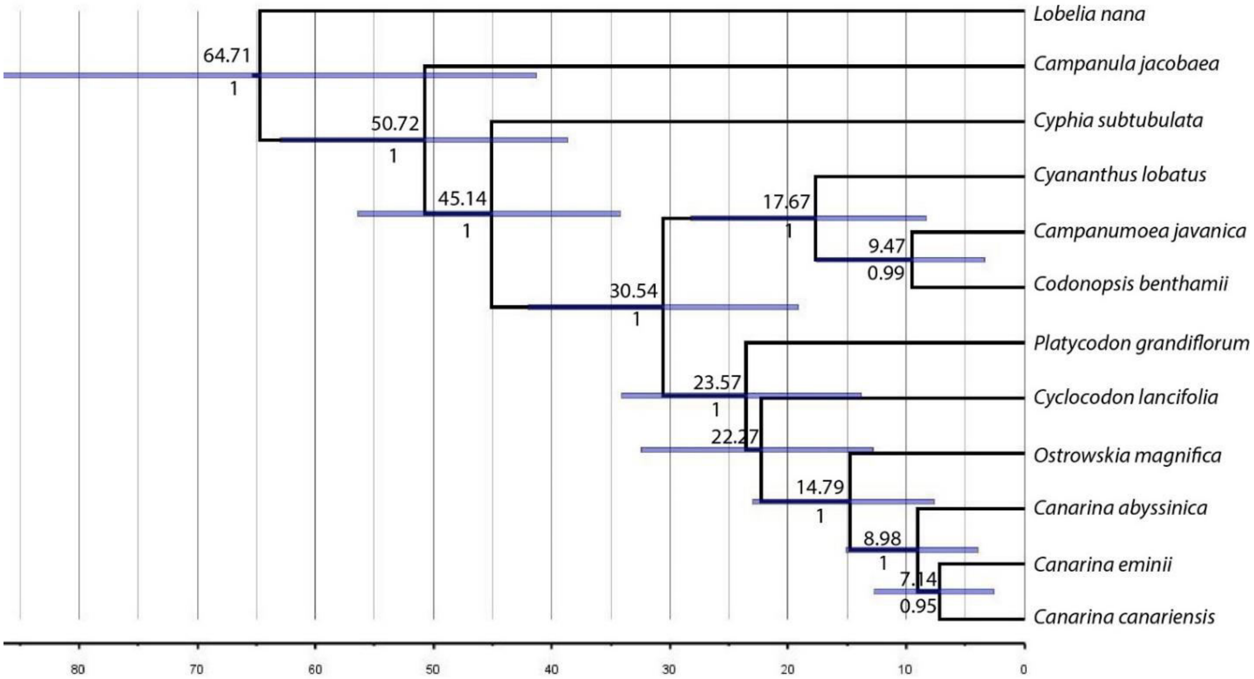
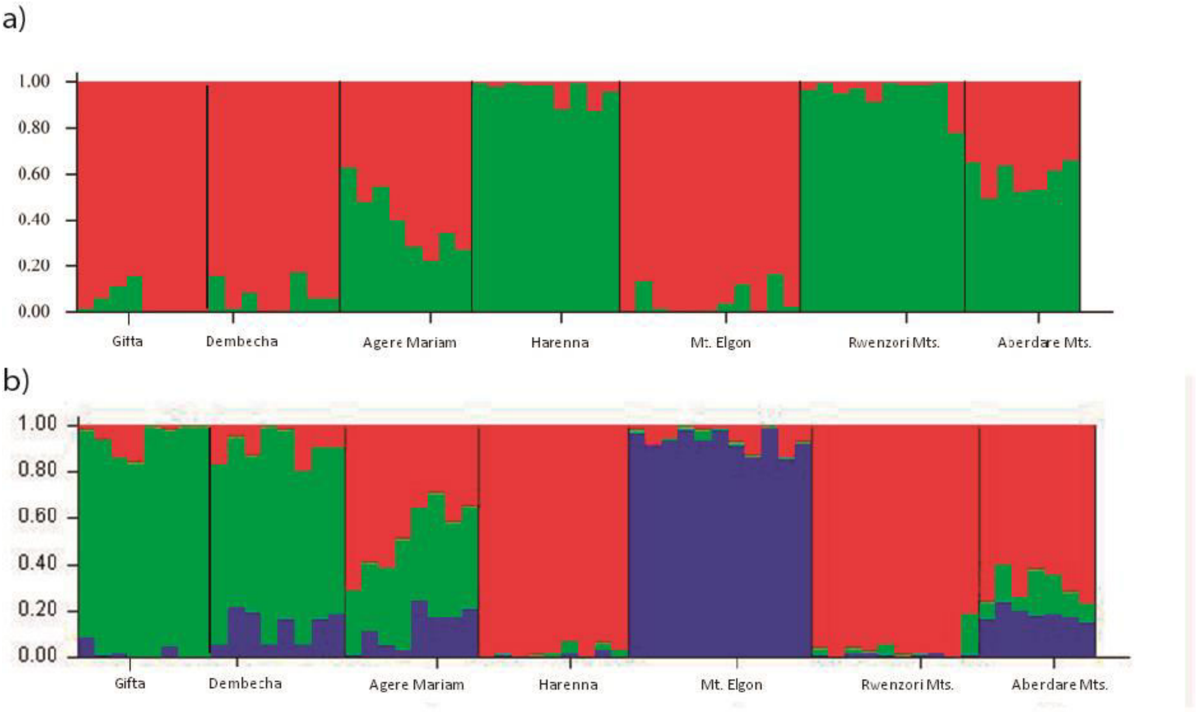
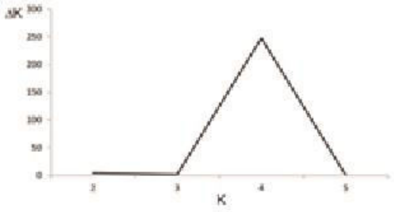
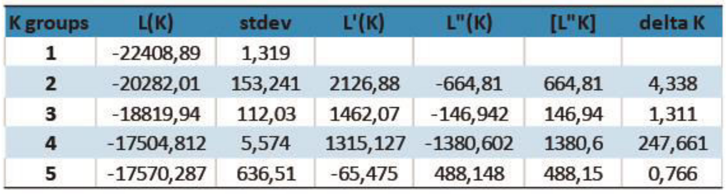


Figure S2.2. Histograms showing the Bayesian clustering of individuals within populations (STRUCTURE) for a) K=2 and b) K=3. Colours represent the proportion of individual membership to each inferred Bayesian group. c) The estimated probability of the likelihood function according to the Evanno method for STRUCTURE analyses of *C. eminii*. Maximum Δ*K* values correspond to the presumed true number of K cluster**s.**



c)





K groups	L(K)	stdev	L'(K)	L"(K)	[L"K]	delta K
1	-22408,89	1,319				
2	-20282,01	153,241	2126,88	-664,81	664,81	4,338
3	-18819,94	112,03	1462,07	-146,942	146,94	1,311
4	-17504,812	5,574	1315,127	-1380,602	1380,6	247,661
5	-17570,287	636,51	-65,475	488,148	488,15	0,766

Figure S2.3. Neighbor tree (A) and neighbor-net (B) for the AFLP data for individuals of *C. eminii*. Numbers are bootstrap values from neighbor-joining analysis using Nei-Li distances over 1000 bootstrap replicates. To root the neighbor tree, we used a population of *C. abyssinica* collected in Ethiopia (ab).

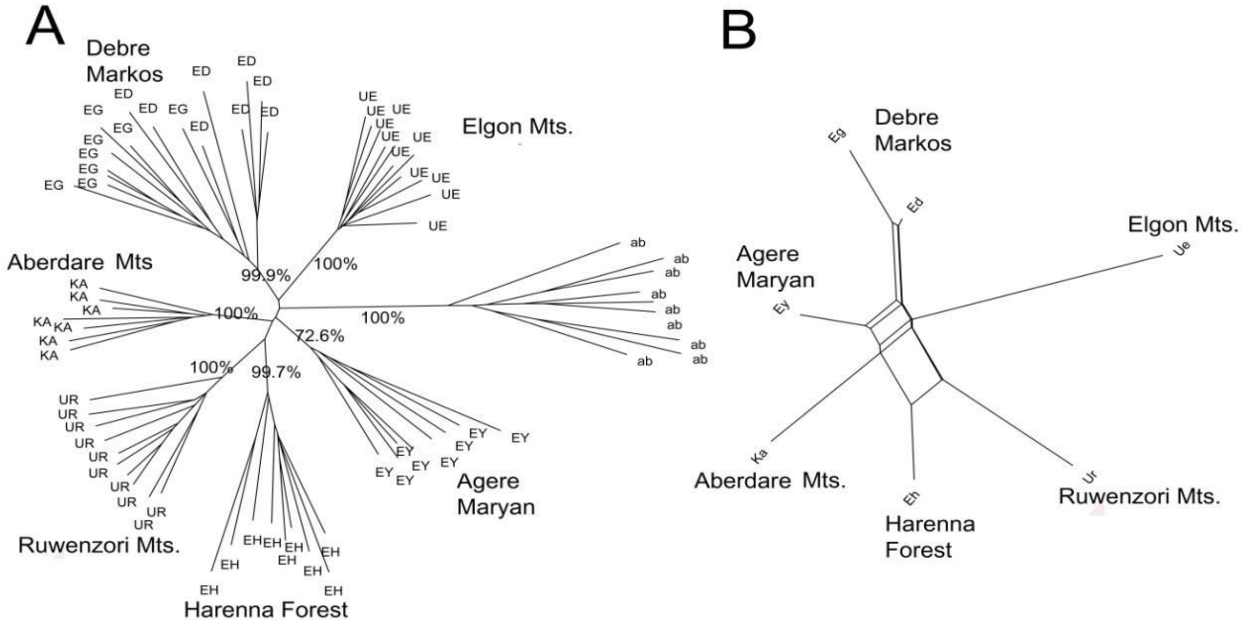


Figure S2.4. Principal coordinates analysis of AFLP data for individuals of *C. eminii.*

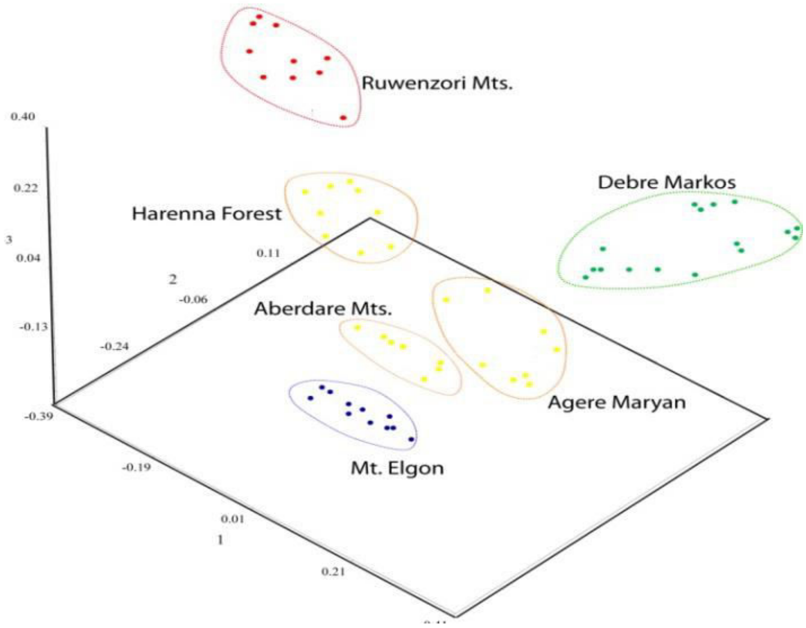


Figure S2.5. Regression plot among pairwise F_{ST} and logarithms of geographical distances. Acronyms are E: Elgon Mt., R: Rwenzori, D: Debre Markos (Gifta and Dembecha), Y: Agere Maryan, A: Abredare Mts., H: Harenna Forest.

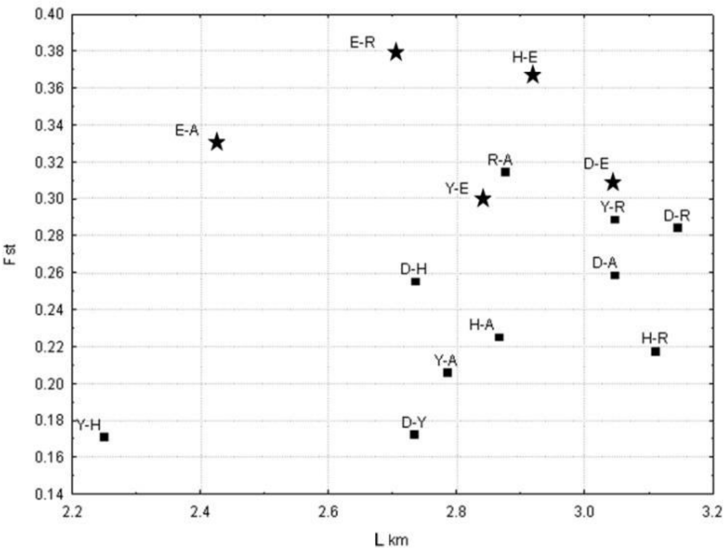
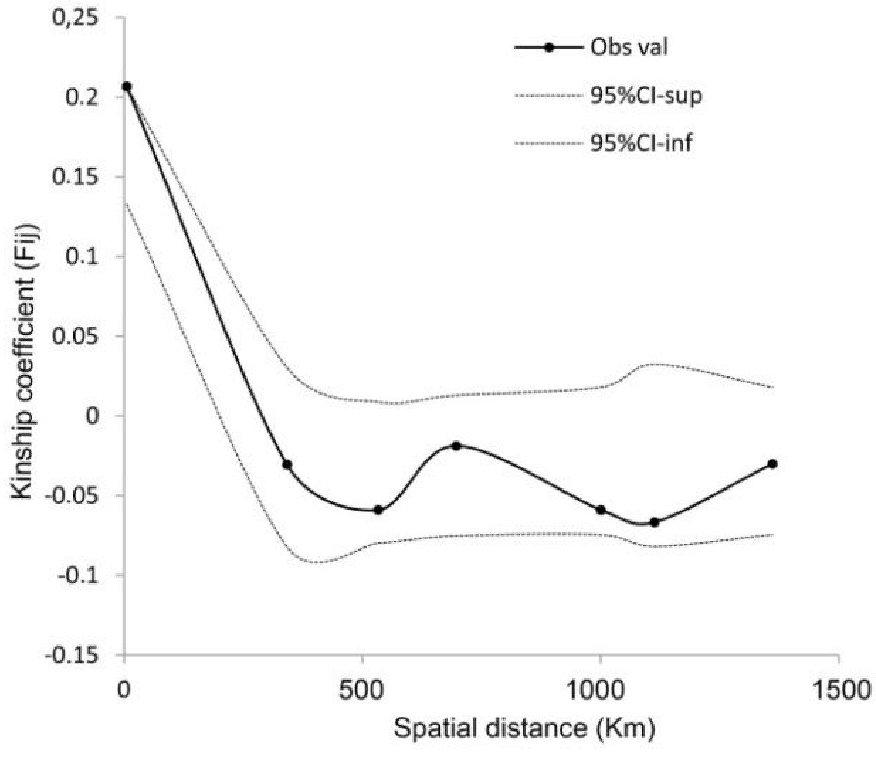


Figure S2.6. Distances intervals for *C. eminii* populations in SPAGeDi analysis (dotted lines show the 95% confidence level).



Appendix S3 Supplemental references

Arctander, P., Johansen, C. & Coutellec-Vreto, M.A. Phylogeography of three closely related African bovids (tribe Alcelaphini). *Molecular Biology and Evolution.* **16**, 1724-1739 (1999).

Belay, G. & Mori, A. Intraspecific phylogeographic mitochondrial DNA (D-loop) variation of Gelada baboon, *Theropithecus gelada*, in Ethiopia. *Biochemical Systematics and Ecology*. **34**, 554-561 (2006).

Braginets, O.P., Minakawa, N., Mbogo, C.M. & Ya, G. Population genetic structure of the African malaria mosquito *Anopheles funestus* in Kenya. *The American journal of tropical medicine and hygiene*. **69**, 303-308 (2003).

Colangelo, P. *et al*. A mitochondrial phylogeographic scenario for the most widespread African rodent, *Mastomys natalensis*. *Biological Journal of the Linnean Society*. **108**, 901-916 (2013).

Derero, A., Gailing, O. & Finkeldey. R. Maintenance of genetic diversity in *Cordia africana* Lam., a declining forest tree species in Ethiopia. *Tree Genetics & Genomes*. **7**, 1-9 (2010).

Dubach, J. *et al*. Molecular genetic variation across the southern and eastern geographic ranges of the African lion, *Panthera leo*. *Conservation Genetics*. **6**, 15-24 (2005).

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

Evans, B.J., Bliss, S.M., Mendel, S.A. & Tinsley, R.C. The Rift Valley is a major barrier to dispersal of African clawed frogs (*Xenopus*) in Ethiopia. *Molecular Ecology*. **20**, 4216-4230 (2011).

Field, L.M. *et al*. Analysis of genetic variability in *Anopheles arabiensis* and *Anopheles gambiae* using microsatellite loci. *Insect Molecular Biology*. **8**, 287-297 (1999).

Girman, D.J. *et al*. Patterns of population subdivision, gene flow and genetic variability in the African wild dog (*Lycaon pictus*). *Molecular Ecology*. **10**, 1703-1723 (2001).

Hall, T.A. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic acids symposium series*. **41**, 95-98 (1999).

Hamilton, M. Four primers pairs for the amplification of chloroplast intergenic regions with intraspecific variation. *Molecular Ecology*. **8**, 521-523 (1999).

Kadu, C.A.C., *et al*. Divergent pattern of nuclear genetic diversity across the range of the Afromontane *Prunus africana* mirrors variable climate of African highlands. *Annals of Botany*. **111**, 47-60 (2013).

Katoh, K., Kuma, K., Toh, H. & Miyata, T. MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Research*. **33**, 511-518 (2005).

Kelchner, S.A. The Evolution of Non-Coding Chloroplast DNA and Its Application in Plant Systematics. *Annals of the Missouri Botanical Garden*. **87**, 482 (2000).

Krafsur, E.S. Population structure of the tsetse fly *Glossina pallidipes* estimated by allozyme, microsatellite and mitochondrial gene diversities. *Insect Molecular Biology*. **11**, 37-45 (2002).

Lehmann, T. *et al*. The Rift Valley complex as a barrier to gene flow for *Anopheles gambia*e in Kenya. *Journal of Heredity*. **90**, 613-621 (1999).

Lehmann, T. *et al*. Brief communication. The Rift Valley complex as a barrier to gene flow for *Anopheles gambiae* in Kenya: the mtDNA perspective. *Journal of Heredity*. **91**, 165-168 (2000).

Lehmann, T. *et al*. Population Structure of *Anopheles gambiae* in Africa. *Journal of Heredity*. **94**, 133-147 (2003).

Löhne, C. & Borsch, T. Molecular evolution and phylogenetic utility of the *petD* group II intron: a case study in basal angiosperms. *Molecular Biology and Evolution*, **22**. 317-332 (2005).

Matthee, C.A. & Robinson, T.J. Molecular phylogeny of the springhare, *Pedetes capensis*, based on mitochondrial DNA sequences. *Molecular Biology and Evolution*- **14**, 20-29 (1997).

Muchugi, A. *et al*. Genetic structuring of important medicinal species of genus *Warburgia* as revealed by AFLP analysis. *Tree Genetics & Genomes*. **4**, 787-795 (2008).

Nicolas, V. *et al*. Phylogeographic structure and regional history of *Lemniscomys striatus* (Rodentia: Muridae) in tropical Africa. *Journal of Biogeography*. **35**, 2074-2089 (2008).

Omondi, S.F. *et al*. Genetic Diversity and Population Structure of *Acacia senegal* (L) Willd. in Kenya. *Tropical Plant Biology*. **3**, 59-70 (2010).

Pitra, C., Hansen, A.J., Lieckfeldt, D. & Arctander, P. An exceptional case of historical outbreeding in African sable antelope populations. *Molecular Ecology*. **11**, 1197-1208 (2002).

Pokorny, L., Oliván, G., Shaw, A. Phylogeographic patterns in two southern hemisphere species of *Calyptrochaeta* (Daltoniaceae, Bryophyta). *Systematic Botany*. **36**, 542-553 (2011).

Posada, D. jModelTest: phylogenetic model averaging. *Molecular Biology and Evolution*. **25**, 1253-1256 (2008).

Rambaut A., Drummond A.J. & Suchard M. Tracer v1. 6. http://tree.bio.ed.ac.uk/software/tracer/ (2007)

Rambaut, A. & Drummond, A. FigTree v1.3.1. http://tree. bio. ed. ac. uk/software/figtree (2009).

Ronquist, F. & Sanmartín, I. Phylogenetic methods in biogeography. *Annual Review of Ecology, Evolution, and Systematics*. **42**, 441 (2011).

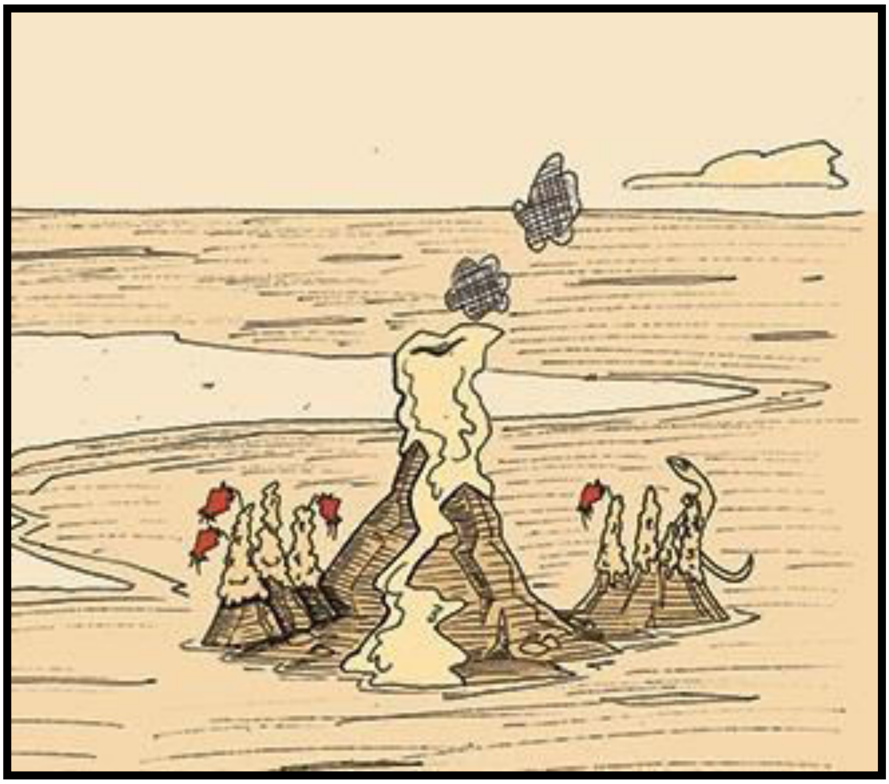
Sertse, D., Gailing, O., Eliades, N.G. & Finkeldey, R. Anthropogenic and natural causes influencing population genetic structure of *Juniperus procera* Hochst. ex Endl. in the Ethiopian highlands. *Genetic Resources and Crop Evolution*. **58**, 849-859 (2011).

Shaw, J., Lickey, E.B., Schilling, E.E. & Small, R.L. Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the tortoise and the hare III. *American Journal of Botany*. **94**, 275-288 (2007).

Silvestrini, M., *et al*. Genetic diversity and structure of Ethiopian, Yemen and Brazilian *Coffea arabica* L. accessions using microsatellites markers. *Genetic Resources and Crop Evolution*. **54**, 1367-1379 (2007).

SUPPLEMENTARY DATA
CHAPTER 3

Paleo-islands as refugia and sources of genetic diversity within volcanic archipelagos: The case of the widespread endemic *Canarina canariensis* (Campanulaceae)



Supplementary Tables

Table S1. Plant species which are restricted and shared between two or more of the ancient static areas considered: the three paleo-islands of Tenerife and La Gomera.

Species	La Gomera	Adeje	Teno	Anaga
Aeonium cuneatum			X	X
Aeonium pseudourbicum		X	X	
Argyranthemum coronopifolium			X	X
Asparagus fallax	X			X
Bystropogon odoratissimum		X	X	X
Ceropegia dichotoma subsp. *krainzii*	X	X		
Cistus chinamadensis	X			X
Convolvulus volubilis	X	X	X	X
Dicheranthus plocamoides	X		X	
Greenovia dodrentalis			X	X
Lavandula buchii			X	X
Lotus dumetorum			X	X
Luzula canariensis	X			X
Navaea phoenicea			X	X
Polycarpaea carnosa	X		X	X
Salvia broussonetii			X	X
Silene lagunensis			X	X
Solanum vespertilio sbsp. *vespertilio*			X	X
Sonchus fauces-orci		X	X	
Teline pallida	X		X	X
Tolpis crassiuscula		X	X	

Table S2. Sampling information and GenBank accession numbers for all of the taxa included in this study. Abbreviations indicating the main geographical areas discussed in the text are: Gran Canaria (GC); Tenerife (TF); Gomera (GO); La Palma (LP); El Hierro (EH). The three paleo-islands of Tenerife (Anaga, Teno and Roque del Conde) are indicated in parentheses.

Population Code	**Island_Locality (Paleoisland)**	**Voucher**	**Geographic coordinates**	**Nº samples**	**Genbank accesion number**		
					***Rpl*32-*trn*L**	***pet*B-*pet*D**	***trn*S-*trn*G**
GC-Cc	Gran Canaria - Cueva Corcho	Aldasoro & Alarcón 10355bis	28º 01' 31" N, 15º 35' 46" W	8	KP798033, KP798034, P798035, KP798036, KP798037,KP798038, KP798058, KP798062	KP798290, KP798291, KP798292, KP798349, KP798362, KP798371, KP798384, KP798397, KP798431	KP798146, KP798147, P798148, KP798149, KP798150,KP798151, KP798152, KP798153, KP798154
GC-Tm	Gran Canaria - Tilos de Moya	Mairal M. MM246 & Caujapé J. 7202	28° 5' 15.334" N, 15° 35' 34.342" W	11	KP798017, KP798018, KP798019, KP798020, KP798021, KP798022, KP798052, KP798053, KP798057, KP798063, KP798064	KP798314, KP798315, KP798316, KP798317, KP798318, KP798319, KP798372,KP798379, KP798380, KP798388, KP798390, XX000	KP798164, KP798165, KP798166, KP798167, KP798168, KP798169, KP798170, KP798213, KP798218, KP798219, KP798284
GC-Es	Gran Canaria - Camino de El Sao	Caujapé J. 7173 & Mairal M. MM48	28° 3' 27.324" N, 15° 39' 35.157" W	10	KP798023, KP798024, KP798025, KP798026, KP798027, KP798028, KP798029, KP798030, KP798031, KP798032	KP798304, KP798305, KP798306, KP798307, KP798308, KP798309, KP798310, KP798311, KP798312, KP798313	KP798155, KP798156, KP798157, KP798158, KP798159, KP798160, KP798161, KP798162, KP798163, XX000
GC-Ag KP798040	Gran Canaria - Andenes de Guayedra	Caujapé J. 7404	28° 3' 50.501" N, 15° 41' 47.252" W	11	KP798011, KP798012, KP798013, KP798014, KP798015, KP798016, KP798048, KP798049, KP798050, KP798051, KP798056	KP798320, KP798321, KP798322, KP798323, KP798324, KP798325, KP798350, KP798360, KP798363, KP798387, KP798392	KP798171, KP798172, KP798173, KP798174, KP798175, KP798176, KP798177, KP798178, KP798179, KP798180, KP798181
An-E	Tenerife - Monte Chamuscadas (Anaga Center)	Aldasoro & Alarcón 10321bis	28°32'13.68"N, 16°15'8.45"W	10	KP798039, KP798040, KP798041, KP798042, KP798054, KP798055, KP798059, KP798060, KP798068, KP798073	KP798285, KP798286, KP798287, KP798288, KP798289, KP798347, KP798348, KP798356, KP798364, KP798365	KP798136, KP798137, KP798138, KP798139, KP798140, KP798141, KP798142, KP798143, KP798144, KP798145
An-C	Tenerife - El Bailadero (Anaga East)	Mairal M. MM238	28° 33' 1.991" N, 16° 12' 12.495" W	11	KP797991, KP797992, KP797993, KP798043, KP798044, KP798045, KP798046, KP798047, KP798048, KP798061, KP798066	KP798338, KP798339, KP798340, KP798341, KP798342, KP798343, KP798344, KP798345, KP798346, KP798354, KP798361	KP798206, KP798207, KP798208, KP798209, KP798210, KP798211, KP798214, KP798215, KP798216, KP798217, KP798220
An-W	Tenerife - Tope del Carnero (Anaga West)	González A. 1000 & Mairal M. MM259	28° 33' 26.470" N, 16° 18' 56.631" W	6	KP797994, KP798067, KP798069, KP798070, KP798071, KP798127	KP798337, KP798353,KP798370, KP798383, KP798386,KP798391, KP798403	KP798200, KP798201, KP798202, KP798203, KP798204, KP798274
TF-Bj	Tenerife - Barranco de Badajoz	Caujapé J. 9959	28° 18' 12.603" N, 16° 25' 42.459" W	4	KP797995, KP797996, KP797997, KP797998	KP798335, KP798336, KP798382, KP798430	KP798195, KP798196, KP798197, KP798198, KP798199
TF-Br	Tenerife - Barranco Ruiz	Caujapé J. 4672 & Mairal M. MM209 &	28° 23' 7.825" N, 16° 37' 35.886" W	6	KP798005, KP798006, KP798007, KP798008, KP798009, KP798010	KP798326, KP798327, KP798328, KP798329, KP798330, KP798331	KP798182, KP798183, KP798184, KP798185, KP798186, KP798187
Conde	Tenerife - Barranco del Infierno (Roque del Conde)	Caujapé J. 9198 & Mairal M .MM 270	28° 7' 2.384" N, 16° 42' 20.783" W	11	KP797999, KP798000, KP798001, KP798002, KP798003, KP798004, KP798072, KP798109, KP798110, KP798128, KP798135	KP798332, KP798333, KP798334, KP798351, KP798355, KP798381, KP798396, KP798420, KP798421, KP798422, KP798429	KP798188, KP798189, KP798190, KP798191, KP798192, KP798193, KP798194, KP798257, KP798258, KP798259, KP798260

Ten-Ep	Tenerife - El Palmar, Monte del Agua (Teno)	Aldasoro & Alarcón 10339	28° 20' 30.681" N, 16° 49' 28.315" W	8	KP798123, KP798124, KP798125, KP798129, KP798130, KP798131, KP798132, KP798133	KP798401, KP798402, KP798405, KP798406, KP798407, KP798408, KP798409, KP798427	KP798268, KP798278, KP798401, KP798402, KP798405, KP798406, KP798407, KP798408, KP798409
Ten-A	Tenerife - Teno Alto (Teno)	González A. 1001 & Mairal M. MM 283	28° 20' 35.7" N, 16° 52' 30.302" W	7	KP798103, KP798104, KP798105, KP798106, KP798107, KP798108, KP798126	KP798400, KP798404, KP798423, KP798424, KP798425, KP798426, KP798428	KP798205, KP798262, KP798263, KP798264, KP798265, KP798266, KP798267
GO	La Gomera - Tamargada	Mairal M. MM025	28° 10' 1.959" N, 17° 13' 7.197" W	12	KP798113, KP798114, KP798115, KP798116, KP798117, KP798118, KP798119, KP798120, KP798121, KP798122, KP798111, KP798112	KP798398, KP798399,KP798410, KP798411, KP798412, KP798413, KP798414, KP798415, KP798416, KP798417,KP798418, KP798419	KP798227,KP798228,KP798229, KP798230,P798231,KP798232,KP798269, KP798270, KP798271, P798279, KP798255, KP798256
,LP-Lt	La Palma - Los Tilos	Mairal M. MM014	28° 47' 23.568" N, 17° 48' 6.864" W	10	KP798094, KP798095, P798096, KP798097, KP798098, P798099, KP798100, KP798101, KP798102, KP798134	KP798293, KP798294,KP798295, KP798296, KP798357,KP798366, KP798367, KP798368,KP798369, KP798373	KP798233, KP798234,KP798235, KP798236,KP798237,KP798238, KP798239, KP798240,KP798241, KP798272
LP-Ba	La Palma - Barranco del agua	Mairal M. MM013	28° 47' 37.309" N, 17° 47' 33.784" W	10	KP798074, KP798075, KP798076, KP798077, KP798078, KP798079, KP798080, KP798081, KP798082, KP798083	KP798298, KP798299, KP798300, KP798301, KP798302, KP798303, KP798358, KP798359, KP798378, KP798385	KP798248, KP798249, KP798250, KP798251, KP798252, KP798253, KP798254, KP798273, KP798276, KP798277
EH	El Hierro - Hoya del Pino	Mairal M. MM017	27° 43' 50.170" N, 18° 1' 35.723" W	9	KP798084, KP798085, KP798086, KP798087, KP798088, KP798089, KP798090, KP798091, KP798092, KP798093	KP798297, KP798374, KP798375, KP798376, KP798377, KP798393, KP798394, KP798395, KP798432	KP798242, KP798243, KP798244, KP798245, KP798246, KP798247, KP798280, KP798281, KP798282, KP798283

Table S3: Results from the sensitivity analyses to assess confidence in our age estimates. In BEAST we explored simpler models of substitution and molecular clock; results reported in the paper are shown in bold (the age of the crown-node, the first divergence event within *C. canariensis*, was constrained using a normal prior based on an age estimate from a previous analysis (Mairal *et al*. 2015)). Additionally, we estimated divergence times using the semiparametric Penalized Likelihood approach of Sanderson 2002, 2004) and based on a maximum-likelihood tree obtained in RAXML; the crown-age of *C. canariensis* was either fixed with the mean age from Mairal *et al*. 2015 or assigned a minimum age (i.e., the lower bound of the 95%HPD credibility interval from this study); *C. eminii* was included as outgroup in the analysis to provide another calibration point.

BEAST								
	Substitution Model	**Clock Model**	**Pathlikelihood.delta**	**Age estimate crown-node (Ma)**	**Age estimate Eastern-clade (Ma)**	**Age estimate Western-clade (Ma)**	**Coefficient of variation**	**Co-variance**
	JC	STRICT	-14547.3	0.82 (0.439 - 1.245)	0.241 (0.044 - 0.859)	0.514 (0.171 - 0.83)	-	-
	HKY	STRICT	-15847.6	0.827 (0.448 - 1.258)	0.24 (0.044 - 0564)	0.50 (0.189 - 0.898)	-	-
	GTR	STRICT	-13680.1	0.831 (0.450 - 1.257)	0.242 (0.042 - 0.561)	0.504 (0.19 - 0.901)	-	-
	GTR+G	**UCLD**	-12425.1	**0.877** (0.4523 - 1.365)	**0.254** (0.042 - 0.633)	**0.522** (0.16 - 0.967)	0.571	4.77 e-2
Penalized Likelihood	**ML tree**			**Settings for crown-age**	**Eastern-clade (Ma)**	**Western-clade (Ma)**	**Smoothing value**	
	Including outgroup *C. eminii*			Minimum age = 0.5	0.0928	0.1806	1	
	Including outgroup *C. eminii*			Fix age = 0.76	0.1028	0.1964	1	
	Only *C. canariensis*			Fix age = 0.76	0.1397	0.4175	6.30E+02	

Additional references:

Sanderson, MJ (2002) Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. *Molecular Biology and Evolution*, **19**, 101–109.

Sanderson, M.J. (2004) r8S, program and documentation. Version 1.7.1. Available at: http://loco.biosci.arizona.edu/r8s/.

Table S4. Haplotypes of *Canarina canariensis*, as determined by using the variable sites and indels found in the sixteen *C. canariensis* populations with markers *Pet*B-*Pet*D (937 bp), *Rpl*32-*trn*L (654 bp) and *trn*S-*trn*G (658 bp).

Haplotypes	**Nº indiv.**		***Pet*B-*Pet*D**		***Rpl*32-*trn*L**					***trn*S-*trn*G**	
		Sites	807	849	976	1123	1314	1346	1505	1759	2106
H1	75		T	T	G	G	A	A	A	A	T
H2	4		T	T	A	G	A	T	A	A	T
H3	25		T	T	G	T	A	A	T	G	G
H4	1		T	T	G	T	A	A	A	G	G
H5	3		T	T	G	G	A	T	A	A	T
H6	3		T	G	G	T	A	A	T	G	G
H7	1		T	G	G	G	A	A	A	A	T
H8	3		G	G	G	T	G	A	A	G	G
H9	28		G	T	G	T	G	A	A	G	G
H10	1		G	T	G	T	G	A	A	A	T

Table S5. G_{ST} and number of migrants (in parentheses) obtained from cloroplastic data in *Canarina canariensis*. Population codes are provided in Table S2.

Pop. Codes	An-W	An-C	An-E	GC-Cc	GC-Es	GC-Tm	GC-Ag	TF-Br	TF-Bj	Conde	Ten-A	Ten-Ep	GO	LP-Lt	EH
An-W	0 (0)														
An-C	0.20310 (1.96)	0 (0)													
An-E	0.25096 (1.49)	0.00011 (4620)	0 (0)												
GC-Cc	0.38292 (0.81)	0.04592 (10.39)	0.00069 (720)	0 (0)											
GC-Es	0.40520 (0.73)	0.05035 (9.43)	0.00000 (4.03 E+14)	1 (NA)	0 (0)										
GC-Tm	0.41402 (0.71)	0.05263 (9)	0.00011 (46209)	1 (NA)	1 (NA)	0 (0)									
GC-Ag	0.41402 (0.71)	0.05263 (9)	0.00011 (4620)	1 (NA)	1 (NA)	1 (NA)	0 (0)								
TF-Br	0.35135 (0.92)	0.04287 (11.16)	0.00415 (120)	1 (NA)	1 (NA)	1 (NA)	1 (NA)	0 (0)							
TF-Bj	0.30435 (1.14)	0.04597 (10.38)	0.01582 (31.11)	1 (NA)	1 (NA)	1 (NA)	1 (NA)	1 (NA)	0 (0)						
Conde	0.08998 (5.06)	0.07890 (5.84)	0.10647 (4.2)	0.16631 (2.51)	0.18314 (2.23)	0.19048 (2.13)	0.19048 (2.13)	0.14651 (2.91)	0.12591 (3.47)	0 (0)					
Ten-A	0.14356 (2.98)	0.40449 (0.74)	0.48166 (0.54)	0.60053 (0.33)	0.62401 (0.3)	0.63304 (0.29)	0.63304 (0.29)	0.56544 (0.38)	0.50790 (0.48)	0.12376 (3.54)	0 (0)				
Ten-Ep	0.38292 (0.81)	0.68288 (0.23)	0.80003 (0.12)	1 (0)	1 (0)	1 (0)	1 (0)	1 (0)	1 (0)	0.30899 (1.12)	0.04651 (10.25)	0 (0)			
GO	0.23108 (1.66)	0.52070 (0.46)	0.59168 (0.35)	0.69156 (0.22)	0.71798 (0.2)	0.72814 (0.19)	0.72814 (0.19)	0.65201 (0.27)	0.58692 (0.35)	0.20638 (1.92)	-0.36699 (-14.13)	0.04038 (11.88)	0 (0)		
LP-Lt	0.35718 (0.9)	0.57942 (0.36)	0.66667 (0.25)	0.80003 (0.12)	0.81818 (0.11)	0.82500 (0.11)	0.82500 (0.11)	0.77165 (0.15)	0.72125 (0.19)	0.41300 (0.71)	0.48166 (0.54)	0.80003 (0.12)	0.59168 (0.35)	0 (0)	
EH	0.21491 (1.83)	0.41765 (0.7)	0.48653 (0.53)	0.58540 (0.35)	0.61227 (0.32)	0.62271 (0.3)	0.62271 (0.3)	0.54603 (0.42)	0.48359 (0.53)	0.27891 (1.29)	0.32049 (1.06)	0.58540 (0.35)	0.43103 (0.66)	0.06805 (6.85)	0 (0)
LP-Ba	0.48498 (0.53)	0.70879 (0.21)	0.81818 (0.11)	1 (0)	1 (0)	1 (0)	1 (0)	1 (0)	1 (0)	0.51563 (0.47)	0.62401 (0.3)	1 (0)	0.71798 (0.2)	0.00000 (4 E+14)	0.14936 (2.85)

Table S6. Mean ages and 95% HPD confidence intervals from the BEAST analysis. Node numbers correspond to Figure S2.

Node	Mean age	Confidence Interval
1	0.878	0.4523 - 1.365
2	0.255	0.041 - 0.633
3	0.067	0.0009 - 0.2431
4	0.107	0.002 - 0.347
5	0.522	0.161 - 0.967
6	0.252	0.027 - 0.639
7	0.089	0.0005 - 0.3318
8	0.216	0.029 - 0.574
9	0.095	0.001 - 0.326

Table S7. Results from the DNA polymorphism, neutrality test and mismatch raggedness for *Canarina canariensis* haplotypes. The mismatch distribution analyses are shown in Figure S1.

	DNA polymorphism				Fu´s *Fs* test	Tajima´s *D* test	Mismatch
	π	θ	H (d)	n	Fs	D	Raggedness
GC + East TF	0.0002	0.0006	0.198	4	-1.07*	-1.760	0.5
West TF+ GO	0.0002	0.0004	0.271	3	-0.66	-1.325	0.3
LPA + EH	0.0001	0.0002	0.254	3	-1.68	-1.008	0.3

* significative

Table S8. Pairwise F_{ST} between populations of *Canarina canariensis* with nuclear data. Population codes are provided in Table S2.

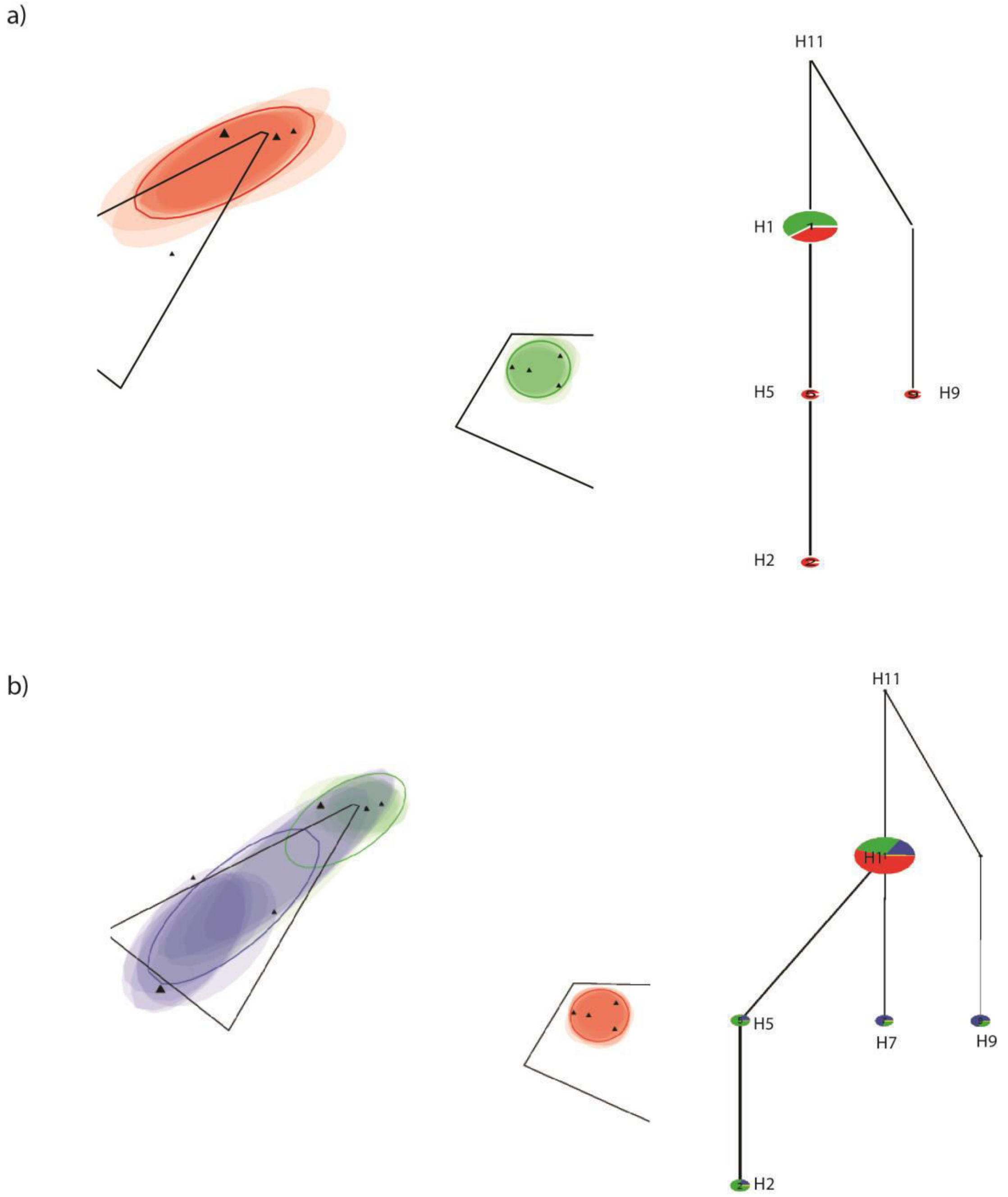
Pairwise F_{ST} between populations	GC-CC	An-W	An-E	An-C	Ten-Ep	Ten-A	GO	LP-Ba	LP-Lt	EH
GC-CC	0.0000	0.2667	0.0328	0.1445	0.2688	0.2530	0.2153	0.3185	0.2657	0.3249
An-W	0.2667	0.0000	0.2616	0.2537	0.2425	0.1235	0.2042	0.2877	0.2618	0.2746
An-E	0.0328	0.2616	0.0000	0.1389	0.2709	0.2547	0.2276	0.3237	0.2579	0.3268
An-C	0.1445	0.2537	0.1389	0.0000	0.2467	0.2084	0.2264	0.3466	0.2715	0.3484
Ten-Ep	0.2688	0.2425	0.2709	0.2467	0.0000	0.1790	0.1788	0.2798	0.2467	0.2877
Ten-A	0.2530	0.1235	0.2547	0.2084	0.1790	0.0000	0.1532	0.2143	0.2185	0.2114
GO	0.2153	0.2042	0.2276	0.2264	0.1788	0.1532	0.0000	0.2034	0.1419	0.2161
LP-Ba	0.3185	0.2877	0.3237	0.3466	0.2798	0.2143	0.2034	0.0000	0.0860	0.0704
LP-Lt	0.2657	0.2618	0.2579	0.2715	0.2467	0.2185	0.1419	0.0860	0.0000	0.1390
EH	0.3249	0.2746	0.3268	0.3484	0.2877	0.2114	0.2161	0.0704	0.1390	0.0000

Table S9. Hierarchical analysis of molecular variance (AMOVA) for *C. canariensis* based on allelic variation at different levels: (A) amongst groups, (B) amongst populations within groups and (C) within populations. All population codes are provided in Table S2, except Teno and Anaga: Teno= Ten-Ep + Ten-A; Anaga =An-C + An-E.

AMOVA groups	**Nº of groups (K)**	**Levels**			**F-statistics**		
		A	**B**	**C**	**Fsc**	**Fst**	**Fct**
All Populations	1	27.33	NA	72.67	NA	NA	NA
[TF + GC] / [GO + LP + EH]	2	10.66	20.14	69.21	0.22538	0.30795	0.10695
[TF + GC + GO] / [LP + EH]	2	9.95	21.17	68.89	0.23507	0.31114	0.09945
[GC + Anaga] / [An-W + Teno + GO + LP + EH]	2	13.46	19.08	67.46	0.22047	0.32540	0.13461
[GC + Anaga + An-W] / [Teno + GO + LP + EH]	2	10.97	20.18	68.85	0.22667	0.31154	0.10975
[GC + Anaga] / [An-W + Teno + GO] / [LP + EH]	3	14.87	15.29	69.84	0.17962	0.30164	0.14874
[GC + Anaga + An-W] / [Teno + GO] / [LP + EH]	3	12.36	17.29	70.34	0.19732	0.29657	0.12364
[GC + Anaga] / [An-W + Teno] / [GO + LP + EH]	3	15.22	15.23	69.55	0.17965	0.30451	0.15220
[GC + Anaga + An-W] / [Teno] / [GO + LP + EH]	3	12.84	17.31	69.85	0.19856	0.30150	0.12843
[GC]/ [TF] / [GO + LP + EH]	3	9.02	20.68	70.30	0.22731	0.29704	0.09024
[GC] / [TF + GO] / [LP + EH]	3	8.08	21.66	70.25	0.23570	0.29749	0.08085
[GC+TF] [GO] [LP+EH]	3	9.28	20.39	70.33	0.22481	0.29675	0.09280
[GC + Anaga] / [An-W] / [Teno + GO + LP + EH]	3	15.37	16.78	67.85	0.19824	0.32148	0.15371
[GC + Anaga] / [An-W + Ten-A] / [Ten-Ep + GO] / [LP + EH]	4	15.75	13.64	70.60	0.16192	0.29396	0.15755
[GC + Anaga] / [An-W]/ [Teno + GO] / [LP + EH]	4	16.90	13.02	70.08	0.15669	0.29918	0.16897
[GC + Anaga] / [An-W] / [Teno] / [GO + LP + EH]	4	17.40	13.14	69.46	0.15907	0.30541	0.17402
[GC + Anaga + An-W] / [Ten-A] / [Ten-Ep + GO] / [LP + EH]	4	11.58	17.56	70.86	0.19863	0.29142	0.11578
[GC+Anaga] / [An-W + Ten-A] / [Ten-Ep] / [GO+LP+EH]	4	17.44	13.06	69.50	0.15820	0.30500	0.17439
[GC + Anaga + An-W] / [Teno] / [GO] / [LP + EH]	4	12.38	16.77	70.85	0.19140	0.29152	0.12382
[GC] / [TF] / [GO] / [LP] / [EH]	5	3.62	24.27	72.10	0.25186	0.27897	0.03624
[GC] / [Anaga + An-W] / [Teno] / [GO]/ [LP + EH]	5	10.53	17.95	71.53	0.20059	0.28473	0.10525
[GC] / [Anaga] / [An-W + Teno] / [GO]/ [LP + EH]	5	13.13	15.56	71.31	0.17913	0.28692	0.13132
[GC + Anaga] / [An-W] / [Teno] / [GO]/ [LP + EH]	5	17.99	11.35	70.65	0.13846	0.29348	0.17993
[GC + Anaga] / [An-W + Ten-A] / [EP] / [GO]/ [LP + EH]	5	18.10	11.21	70.69	0.13683	0.29310	0.18105
[GC + Anaga] / [An-W] / [Ten-A] / [EP + GO]/ [LP + EH]	5	16.85	12.52	70.62	0.15061	0.29377	0.16854
[GC+An-E] [An-C] [Ten-Ep] [An-W + Ten-A] [GO+LP+EH]	5	18.92	11.21	69.88	0.13819	0.30123	0.18918

Supplementary Figures.

Figure S1. Clusters and networks obtained with Bayesian Phylogeographic and Environmental Clustering (BPEC) with plastid DNA sequences. The left handside shows the phylogeographical clusters and the ``most likely ancestral areas´´ (the greater the triangle, the greater the likelihood). The right handside shows the networks with the highest probability. The small black circles indicate inferred intermediate unobserved haplotypes. a) Eastern Clade; b) Eastern Clade + Conde; c) Western Clade; d) Western Clade + Conde.



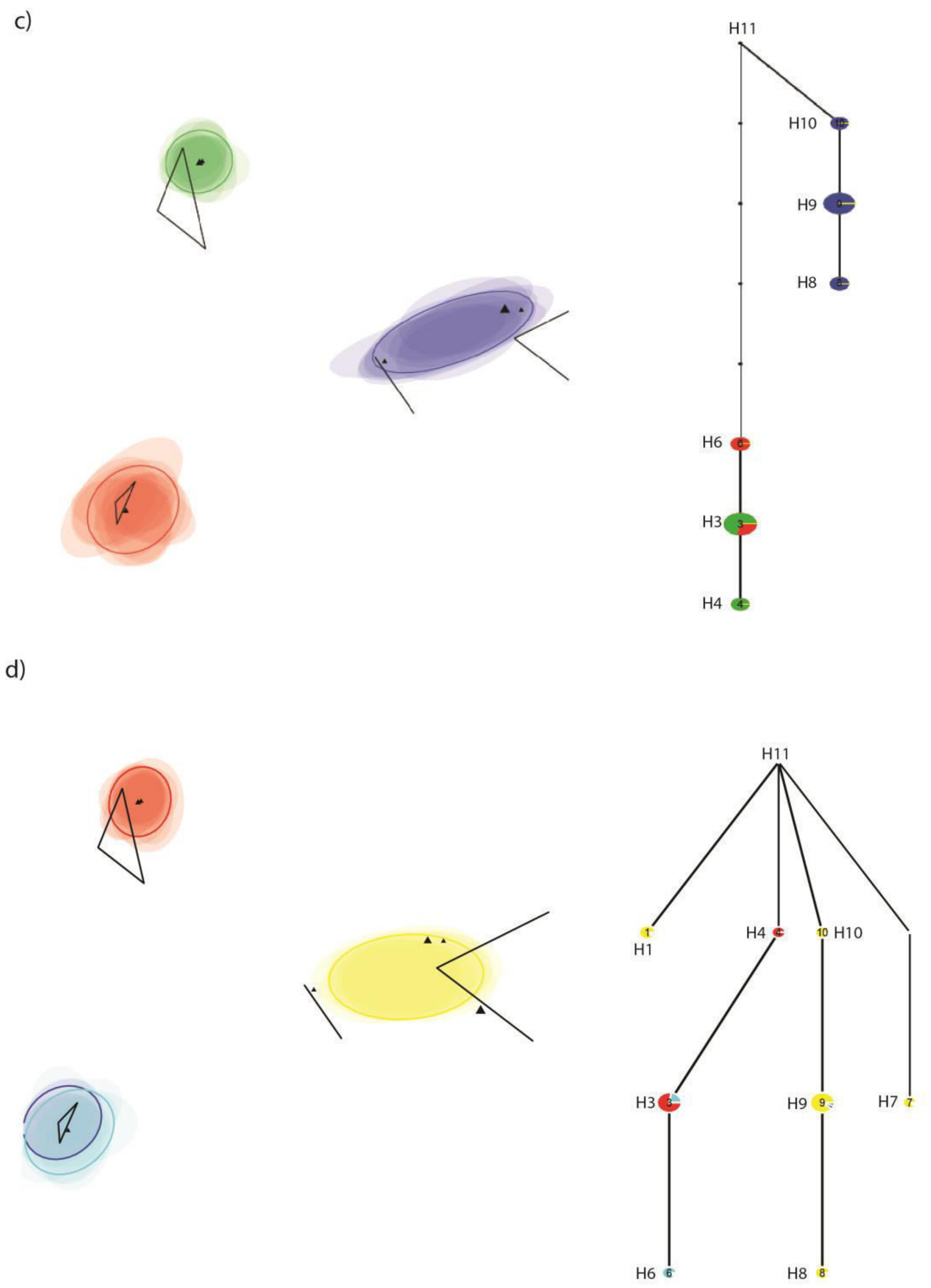
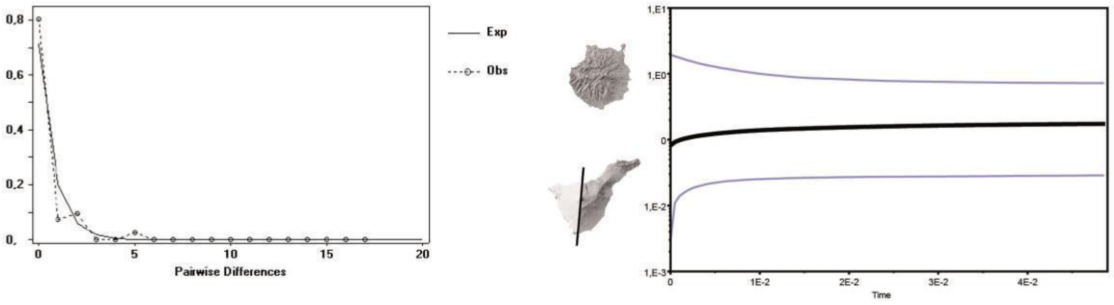
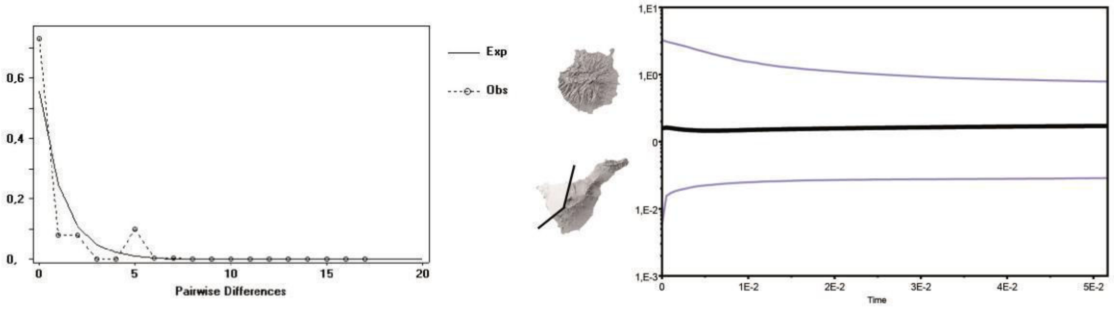

c)
H11
H10
H9
H8
H6
H3
H4
d)
H11
H1
H4
H10
H3
H9
H7
H6
H8

Figure S2. Mismatch distribution (left handside) and Extended Bayesian Skyline Plot analysis (right handside) for the different groups.

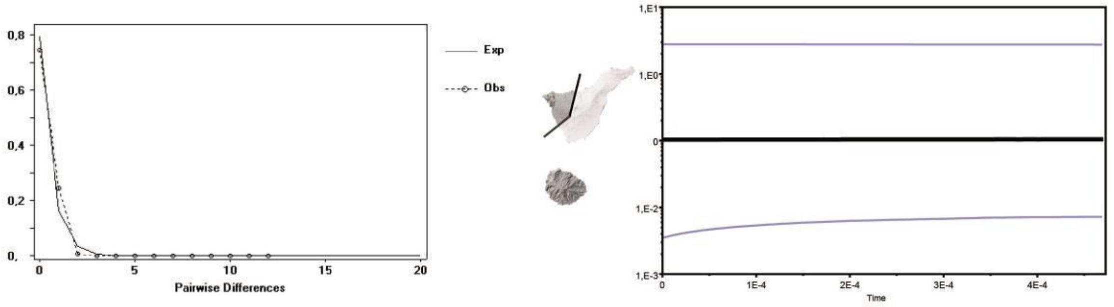
ET-GC



ET-GC + Conde



TENO-GO



LP-EH

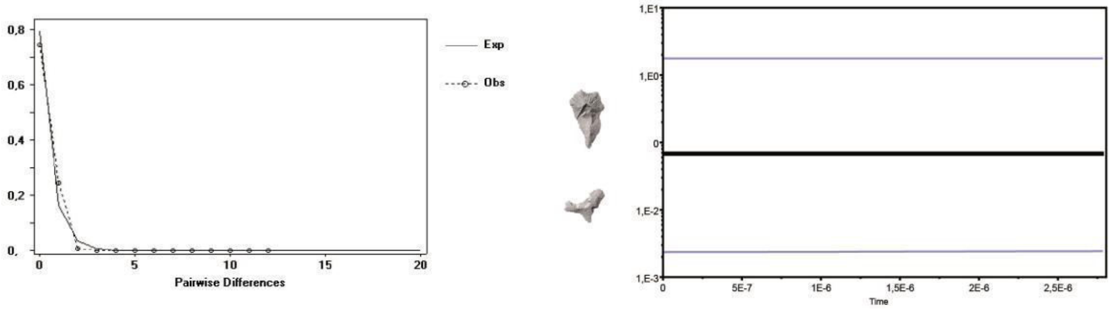


Figure S3. The MCC tree, from the BEAST analysis, showing nodes with mean ages and the 95% HPD confidence intervals (values specified in Table S5).

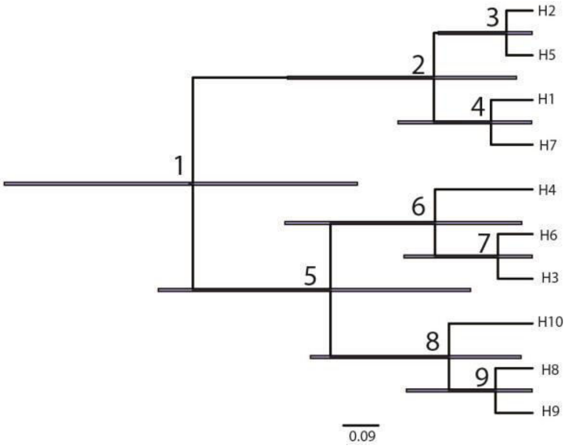


Figure S4. Principal coordinates analysis of AFLP data for individuals of *C. canariensis*. The right handside shows a geographical legend for the letter codes.

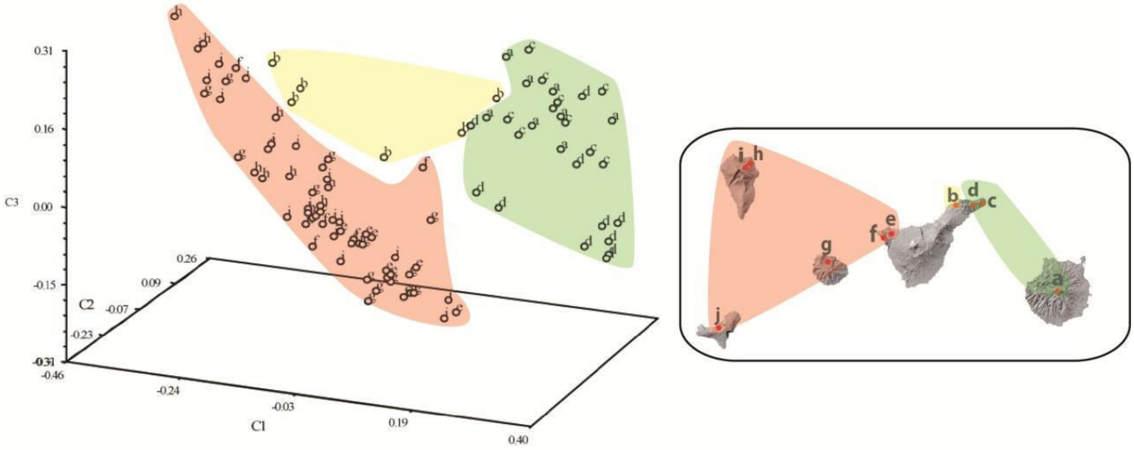
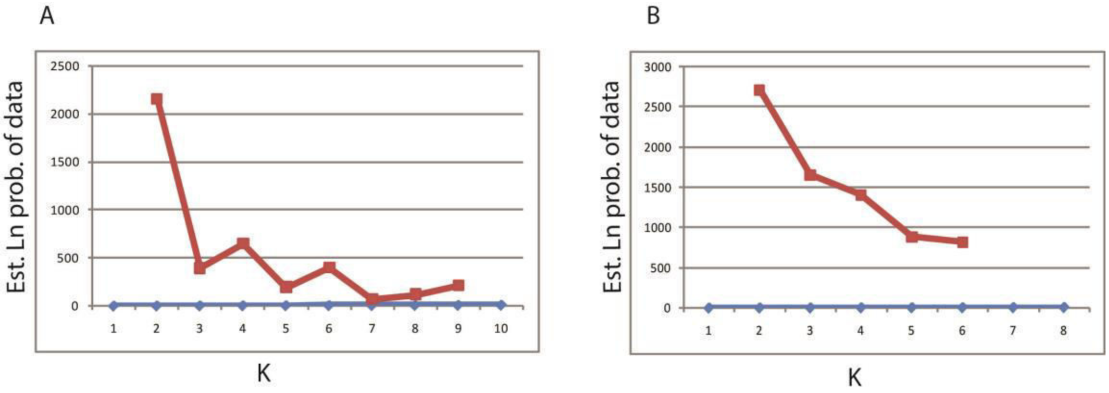


Figure S5. The estimated probability of the likelihood function according to the Evanno method for: a) STRUCTURE analysis for *C.canariensis*; b) STRUCTURE analysis using only Teneriffean populations.



A tale of two forests: ongoing aridification drives population decline and genetic diversity loss at continental scale in Afro-Macaronesian evergreen forests archipelago endemics.

Authors: Mario Mairal[1, 2], Juli Caujapé- Castells [3], Loïc Pellissier[4], Ruth Jaén-Molina[3], Nadir Álvarez[5], Myriam Heuertz[6], Isabel Sanmartín[1]

Supporting information

Table S1. Primers used for PCR amplification for each loci.

C. canariensis

nº primer	T ª annealing	Motifs	Motif length	Bases in all Motifs
P22	53.5ºC	AAC	3	36
P25	63ºC	ATT	3	30
P14	59ºC	TTC	3	42
P23	58ºC	TTC	3	36
P6	56ºC	ATGG	4	32
P20	56ºC	AAC	3	36
P9	56ºC	ATCT	4	28
P13	68ºC	AAC	3	45
P2	63ºC	AAAG	4	40

C. Eminii

nº primer	T ª annealing	Motifs	Motif length	Bases in all Motifs
P22	53.5ºC	AAC	3	36
P25	63ºC	ATT	3	30
P20	56ºC	AAC	3	36
P9	56ºC	ATCT	4	28
P13	68ºC	AAC	3	45
P26	56ºC	ACC	3	30

Table S2. Geographical coordinates used in the Ecological Niche Modelling of *Canarina canariensis*.

ID	Latitude	Longitude	ID	Latitude	Longitude
1	28.05685	-15.5878	55	28.56294	-16.1668
2	28.08217	-15.5977	56	28.55055	-16.2035
3	28.05101	-15.5899	57	28.55735	-16.3157
4	28.0107	-15.5273	58	28.3035	-16.4285
5	28.03159	-15.8456	59	28.38551	-16.6266
6	28.04485	-15.5835	60	28.11733	-16.7058
7	28.0717	-15.5383	61	28.34186	-16.8245
8	28.09024	-15.6008	62	28.34325	-16.8751
9	28.55015	-16.2041	63	28.16721	-17.2187
10	28.47	-16.12	64	28.18465	-17.2397
11	28.33082	-16.8189	65	28.78988	-17.8019
12	28.38	-16.33	66	28.7937	-17.7927
13	28.5167	-16.4	67	28.76382	-17.7749
14	28.54806	-16.2184	68	28.80294	-17.7936
15	28.04432	-15.4839	69	27.7306	-18.0266
16	28.55015	-16.2041	70	28.30796	-16.7649
17	28.45153	-16.4157	71	28.56457	-16.167
18	28.55	-16.22	72	28.56027	-16.2841
19	28.53296	-16.2638	73	28.16746	-17.2171
20	28.53296	-16.2638	74	28.32852	-16.7032
21	28.35517	-16.8124	75	28.35625	-16.7983
22	28.47	-16.12	76	28.3754	-16.6282
23	28.54638	-16.1874	77	28.12818	-16.7024
24	28.18365	-17.2717	78	28.35276	-16.807
25	28.18263	-17.2417	79	28.11534	-15.5949
26	28.15817	-17.1934	80	28.06012	-15.6998
27	28.15607	-17.2063	81	28.12556	-15.6124
28	28.1627	-17.2237	82	28.10558	-15.581
29	28.15607	-17.2063	83	28.08851	-15.57
30	28.15364	-17.171	84	28.07797	-15.5811
31	28.15063	-17.1919	85	28.07388	-15.5818
32	28.14946	-17.192	86	28.06939	-15.59
33	28.15364	-1.7E+08	87	28.06138	-15.5815
34	28.15364	-17.171	88	28.06447	-15.5885
35	28.15364	-17.171	89	28.0531	-15.5806
36	28.15364	-17.171	90	28.05293	-15.5706
37	28.55634	-16.2649	91	28.03443	-15.5917
38	28.15364	-17.171	92	28.0914	-15.601
39	28.14302	-17.2162	93	28.07927	-15.6009
40	28.82516	-17.8867	94	28.08047	-15.6122
41	28.1	-17.1	95	28.06942	-15.6205
42	28.01944	-15.5531	96	28.02449	-15.5598

43	28.61919	-17.7863	97	28.01641	-15.5619
44	27.77991	-17.9872	98	28.00772	-15.5611
45	27.70822	-17.9649	99	27.97659	-15.5084
46	28.67504	-17.7905	100	27.99774	-15.5696
47	28.51268	-16.352	101	27.96333	-15.4976
48	28.55	-16.2667	102	27.96386	-15.5081
49	28.02527	-15.5961	103	28.06138	-15.4763
50	28.06104	-15.5884	104	27.98292	-15.5382
51	28.09886	-15.5754	105	28.04472	-15.4689
52	28.08759	-15.5929	106	28.07527	-15.4985
53	28.05759	-15.6598	107	27.97473	-15.4785
54	28.06403	-15.6965	108	28.00748	-15.5289

Table S3. Geographical coordinates used in the Ecological Niche Modelling of *Canarina eminii*.

ID	Latitude	Longitude	ID	Latitude	Longitude
1	-2.25	28.68333	41	7.5	36.52
2	0.21	29.5	42	7.15	38.71667
3	7.56	39.09	43	7.55	39.08
4	-9.13	33.71	44	7.15	38.71667
5	9.02	36.29	45	1.084139	37.04889
6	-8.98	33.6	46	1.068056	37.24806
7	-9.17	33.88	47	6.700833	39.72639
8	-2.13333	28.83333	48	6.081667	38.2425
9	1.1	34.24	49	1.328889	34.41611
10	-9.18	33.65	50	0.764444	36.74306
11	9.02	36.29	51	0.359722	29.97222
12	8.58333	38.01667	52	9.018375	38.7722
13	-9	33.7	53	9.080695	38.77219
14	8.58333	38.01667	54	1.912099	30.61751
15	9.02	36.29	55	0.232985	29.83998
16	-2.23333	28.78333	56	0.332078	29.82401
17	8.58333	38.01667	57	9.083333	38.58333
18	0.21	29.5	58	-2.72901	28.75119
19	7.15	38.71667	59	3.778691	32.9173
20	7.56	39.09	60	-1.6016	27.97826
21	9.03333	36.63333	61	1.753087	34.68635
22	9.03333	36.63333	62	1.246711	34.38147
23	7.55	39.08	63	1.245984	34.39745
24	8.58333	38.01667	64	1.246712	34.38147
25	7.5	36.52	65	1.196913	34.67004
26	-3.14	29.34	66	2.515355	34.76573
27	7.55	39.08	67	0.203156	34.90793
28	8.58333	38.01667	68	-9.11165	33.52943

29	9.02	36.29
30	8.58333	38.01667
31	0.21	29.5
32	0.21	29.5
33	9.03333	36.63333
34	7.56	39.09
35	9.05	38.56667
36	7.15	38.71667
37	7.55	39.08
38	-3.18	29.6
39	8.58333	38.01667
40	9.02	36.29
69	0.615923	35.50846
70	-1.21223	36.8254
71	-0.14126	37.3079
72	6.858021	36.19624
73	-0.44012	36.67394
74	6.046251	37.60266
75	-0.63046	35.82881
76	6.046251	37.60266
77	1.211568	35.55595
78	1.102873	37.05502
79	1.098176	37.71058

Table S4. The estimated probability of the likelihood function according to the Evanno method for STRUCTURE analyses of *C. canariensis*. Maximum ΔK values correspond to the presumed true number of K clusters.

K	Reps	Mean LnP(K)	Stdev LnP(K)	Ln'(K)	|Ln''(K)|	Delta K
1	10	-2544.63	0.286938	—	—	—
2	10	-2235.31	0.128668	309.32	167.43	1301.252
3	10	-2093.42	11.37949	141.89	14.83	1.303221
4	10	-1966.36	0.636309	127.06	38.35	60.26948
5	10	-1877.65	17.13562	88.71	43.07	2.513478
6	10	-1832.01	0.617252	45.64	31.24	50.61142
7	10	-1817.61	1.04078	14.4	0.19	0.182555
8	10	-1803.02	5.173823	14.59	16.18	3.127281
9	10	-1804.61	18.22449	-1.59	0.44	0.024143
10	10	-1806.64	24.73828	-2.03	—	—

Table S5. The estimated probability of the likelihood function according to the Evanno method for STRUCTURE analyses of *C. eminii*. Maximum ΔK values correspond to the presumed true number of K clusters.

K	Reps	Mean LnP(K)	Stdev LnP(K)	Ln'(K)	|Ln''(K)|	Delta K
1	10	-558.2	1.140175	—	—	—
2	10	-491.4	0.982061	66.8	106.23	108.1704
3	10	-530.83	13.24991	-39.43	32.74	2.47096
4	10	-537.52	8.321966	-6.69	1.36	0.163423
5	10	-545.57	19.12375	-8.05	5.61	0.293352

6	10	-548.01	5.572642	-2.44	20.15	3.615879
7	10	-570.6	16.83482	-22.59	11.5	0.683108
8	10	-581.69	11.79844	-11.09	12.21	1.034882
9	10	-580.57	14.28185	1.12	130.96	9.169681
10	10	-710.41	401.1391	-129.84	—	—

Table S6. Hierarchical analysis of molecular variance (AMOVA) for *C. canariensis* and *C. eminii* based on allelic variation at different levels: (A) among groups, (B) among population within groups and (C) within populations. Abbreviations indicating the main geographical areas discussed in the text are: Gifta (Gif); Dembecha (Dem); Yirga (Yir); Egon (Elg); Aberdare (Ab); Ruwenzori (Ruw); Harenna Forest (HF); Gran Canaria (GC); Tenerife (TF); Gomera (GO); La Palma (LP); El Hierro (EH); Tilos de Moya (GC-Tm); Andenes de Guayedra (GC-Ag); El Sao (GC-Es); Barranco de Badajoz (Bb); Barranco Ruiz (Br).

		Levels			**F-statistics**		
AMOVA Groups		**A**	**B**	**C**	**Fsc**	**Fst**	**Fct**
C. eminii	**Nº of groups (K)**						
[Gif+Dem], [Yir], [Elg], [Ab], [Ruw], [HF]	6	23.22	-5.12	81.9	-0.07	0.181	0.231
[Gif+Dem+HF+Yir], [Ruw+Elg+Ab]	2	20.17	4.12	75.71	0.051	0.242	0.201
[Gif+Dem+HF+Yir], [Ruw+Ab],[Elg]	3	20.06	1.97	77.97	0.024	0.22	0.2
[Gif+Dem+Ruw], [Yir+Ab+HF], [Elg]	3	-4.56	21.03	83.54	0.201	0.164	-0.045
[Gif+Dem+Ruw+Elg], [Yir+Ab+HF]	2	-6.23	21.36	84.88	0.201	0.151	-0.062
[Gif+Dem+Ruw], [Yir+Ab+HF+Elg]	2	0.06	17.36	82.58	0.173	0.174	0.0006
C. canariensis	**Nº of groups (K)**						
Individual populations	17	22.21	12.68	65.11	0.163	0.348	0.222
[GC],[TF],[GO],[LP],[EH]	5	18.25	19.36	62.39	0.236	0.376	0.182
[GC-Tm+GC-Ar],[GC-Ag+GC-Es+Anaga+Bb],[Teno+Tf-Br+Tf-Bb+GO],[LP+EH]	4	18	19.62	62.37	0.239	0.376	0.18
[GC+Anaga+Br], [Teno+Bb+Br+LG], [EH+LP]	3	17.03	21.43	61.54	0.258	0.384	0.17
[GC+TF], [GO+LP+EH]	2	12.36	26.8	60.85	0.305	0.391	0.123

Table S7. Probabilities for each scenario performed with DIYABC. Confidence intervals are shown in square brackets.

C. canariensis **Oceanic islands**	**Population**	**HPD scenario uniform**	**HPD scenario expansion**	**HPD scenario bottleneck**	**HPD scenario decline**
Gran Canaria	8. Cueva Corcho	0.0324 [0.0260,0.0627]	0.0263 [0.0128,0.0399]	0.0183 [0.01034,0.2135]	0.9293 [0.9066,0.9519]

Gran Canaria	11.Andenes de Guayedra	0.1407 [0.1082,0.1732]	0.1453 [0.1316,0.2010]	0.021 [0.0092,0.0549]	0.6930 [0.6497,0.7363]
Gran Canaria	10. El Sao	0,2369 [0.1908,0.2831]	0,1961 [0.1559,0.2364]	0,1767 [0.1383,0.2152]	0,3902 [0.3325,0.4478]
Gran Canaria	9. Tilos de Moya	0.0205 [0.0192,0.0657]	0.0295 [0.0283,0.0640]	0.0013 [0.0001,0.0027]	0.9400 [0.7882,1.0000]
Tenerife (Anaga)	12. Camino Chamuscadas	0.2367 [0.2082,0.2651]	0.1651 [0.1414,0.1908]	0.0028 [0.0013,0.042]	0.5972 [0.5646,0.6298]
Tenerife (Anaga)	13. El Bailadero	0.02513 [0.0000,0.1720]	0.0212 [0.0000,0.0963]	0.02717 [0.01328,0.05896]	0.9265 [0.7775,1.0000]
Tenerife (Teno)	16. El Palmar	0.0211 [0.0202,0.0459]	0.0159 [0.0065,0.0252]	0.0119 [0.0054,0.0403]	0.9511 [0.9354,0.9668]
Tenerife (Anaga)	14. Tope del Carnero	0.2508 [0.2179,0.2837]	0.0958 [0.0726,0.1190]	0.0534 [0.01932, 0.7503]	0.6000 [0.1706,1.0000]
Tenerife (Teno)	15. Teno Alto	0.0271 [0.0147,0.0395]	0.0185 [0.0098,0.0313]	0.0021 [0.0008,0.0034]	0.9523 [0.9360,0.9686]
Tenerife (Adeje)	17. Barranco del Infierno	0.0033 [0.0000,0.0077]	0.0015 [0.0000,0.0067]	0.0012 [0,0.0024]	0.9940 [0.9880,1.0000]
Tenerife	18. Barranco de Badajoz	0.0616 [0.0353,0.0880]	0.0448 [0.0208,0.0688]	0.0436 [0.0210,0.0663]	0.8499 [0.8096,0.8902]
Tenerife	19. Barranco Ruiz	0,0527 [0.0281,0.0774]	0,0483 [0,0258,0,0708]	0,0215 [0.0074,0.0356]	0,8775 [0.8423,0.9127]
La Palma	22. Barranco del Agua	0.0500 [0.0315,0.0684]	0.0872 [0.0639,0.1105]	0.1375 [0.05067,0.2623]	0.7253 [0.6768,0.7738]
La Palma	21. Los Tilos	0.0918 [0.0854,0.1582]	0.0757 [0.0463,0.1051]	0.0301 [0.01301,0.0529]	0.8024 [0.7579,0.8469]
El Hierro	24. El Hierro	0.1003 [0.0624,0.1382]	0.1189 [0.1035,0.1963]	0.0031 [0.0008,0.0054]	0.7498 [0.6942,0.8054]
La Gomera	20. Tamargada	0.3262 [0.2930,0.3594]	0.2733 [0.2502,0.3145]	0.0569 [0.0252,0.0803]	0.3436 [0.3115,0.3758]

***C. eminii* Continental islands**	**Population**	**HPD scenario uniform**	**HPD scenario bottleneck**	**HPD scenario expansion**	**HPD scenario decline**
Abyssinian massif	1. Gifta	0.2531 [0.2246,0.2817]	0.2812 [0.2507,0.3117]	0.3197 [0.2881,0.3513]	0.1460 [0.1228,0.1691]
Harar massif	4. Harenna	0.1633 [0.1324,0.1942]	0.2744 [0.2376,0.3113]	0.2274 [0.1925,0.2624]	0.3348 [0.2951,0.3746]
Harar massif	3. Yirga	0.2121 [0.1821,0.2421]	0.2671 [0.2334,0.3008]	0.2796 [0.2453,0.3138]	0.2413 [0.2103,0.2722]
Rwenzori Mts.	7. Rwenzori	0.1222 [0.0970,0.1474]	0.0862 [0.0643,0.1081]	0.0929 [0.0703,0.1155]	0.6986 [0.6631,0.7342]
Elgon Mt.	6. Elgon	0.1059 [0.0815,0.1303]	0.0771 [0.0557,0.0984]	0.1259 [0.0994,0.1523]	0.6912 [0.6544,0.7280]
Aberdare Mts.	5. Aberdare	0.1082 [0.0475,0.1690]	0.0668 [0.0293,0.1043]	0.0739 [0.0317,0.1161]	0.7511 [0.6710,0.8312]

Figure S1. Graphical representation of the four scenarios simulating changes in population size with DIYABC. 1: constant population size; 2: bottleneck event; 3: expansion event; and 4: decline event. "0" corresponds to the present time.

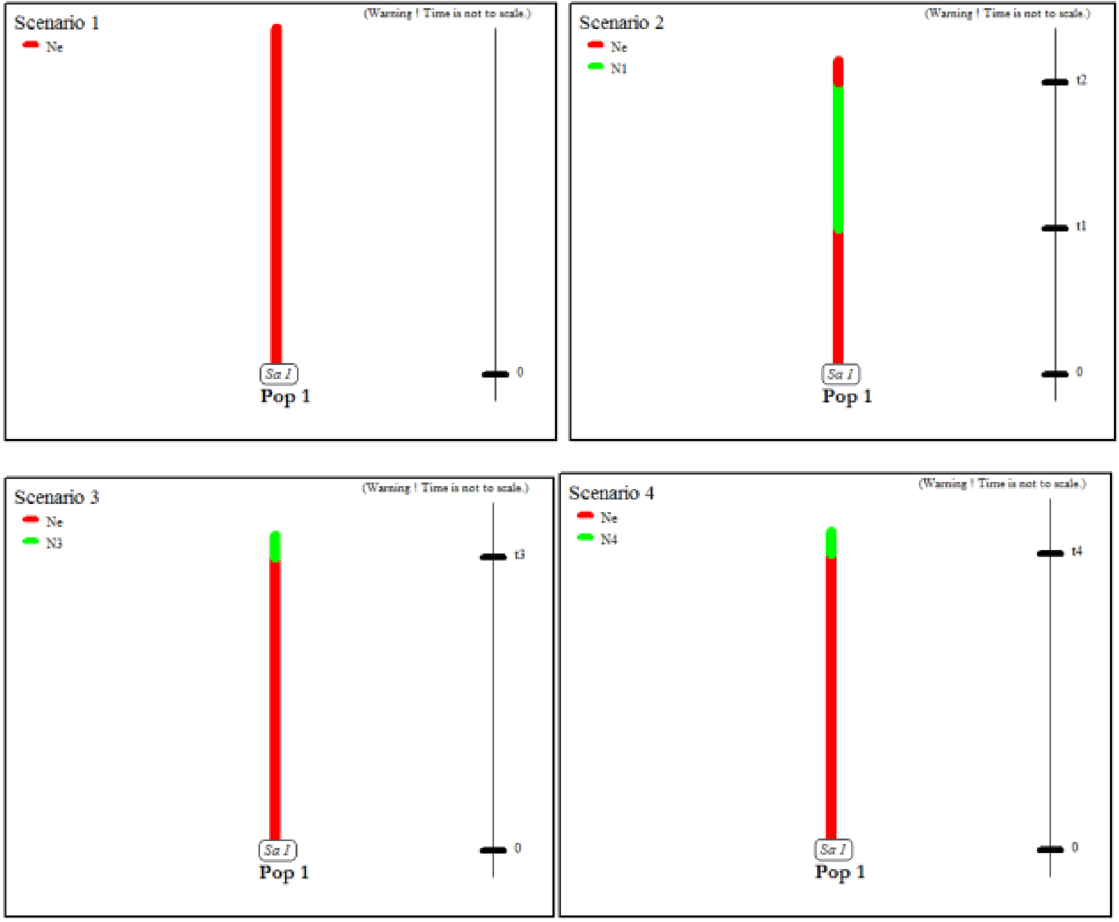
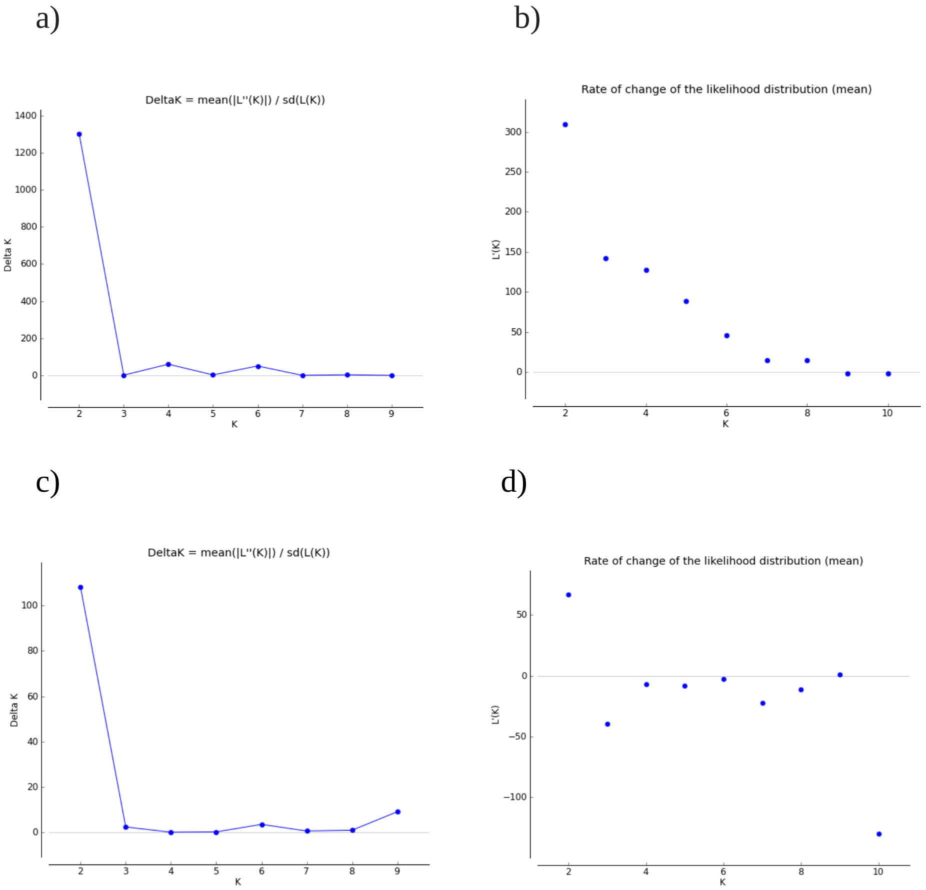


Figure S2. Estimated probability of the likelihood function according to the Evanno method (Delta K, on the left) and the absolute values of the rate of change of the likelihood distribution (on the right) for *C. canariensis* (a,b) and *C. eminii* (c.d).



Supplementary references.

Harrell JFE (2012) rms: Regression modeling strategies. R package version 3.4-0.

Hastie T. (2016) gam: Generalized additive models. R package version, 1. https://cran.r-project.org/web/packages/gam/ gam.pdf.

Hijmans RJ, van Etten J (2016) raster: raster: Geographic data analysis and modeling. R package version, 2–15. https://cran.r-project.org/web/packages/raster/ras- ter.pdf

Hijmans RJ, van Etten J (2016) raster: raster: Geographic data analysis and modeling. R package version, 2–15. https://cran.r-project.org/web/packages/raster/ras- ter.pdf

Liaw A, Wiener M (2002) Classification and Regression by randomForest. R news, 2, 18-22.

Ridgeway G (2015) Generalized boosted regression models. Documentation on the R Package 'gbm', version 1? 5, 7. https://cran.r-project.org/web/packages/gbm/gbm.pdf.

Team RC (2012) R: A language and environment for statistical computing.

Thuiller W, Georges D, Engler R (2013) biomod2: Ensemble platform for species distribution modeling. R package version, 2, r560.

Urbanek S (2010) rJava: Low-level R to Java interface.

VanDerWal J, Falconi L, Januchowski S, Shoo L, Storlie C (2011) SDMTools: Species distribution modelling tools: Tools for processing data associated with species distribution modelling exercises. R package version, 1.

SUPPLEMENTARY DATA
CHAPTER 5

Living on the edge: timing of Rand Flora disjunctions congruent with ongoing aridification in Africa

STUDY GROUPS

1) *Adenocarpus* DC. (Fabaceae) (Fig. 3) contains approximately 20 species, and most of these inhabit NW Africa and the Iberian Peninsula. However, one well-supported monophyletic group (Percy and Cronk, 2002) shows a disjunct pattern where the three Macaronesian species and the widespread Mediterranean *A. complicatus* J. Gay form a group sister to the only subtropical species in Africa, *A. mannii* (Hook. f.) Hook. f. According to Percy and Cronk (2002) the distribution of *A. mannii* can be explained by either a W-E dispersal route of Mediterranean-Macaronesian species via the Saharan mountains or an eastern route of dispersal of Mediterranean species via the Red Sea Hills. Despite efforts to reconstruct its biogeographic history, Cubas et al. (2010) could not distinguish between these to routes in part due to a lack of support in the node leading to this species.

2) Aeonium alliance (Crassulaceae): although the Crassulaceae are a morphologically diverse, large family (ca. 1500 spp.), the Macaronesian representatives constitute a well-supported monophyletic lineage (Mort et al., 2002), which includes three genera endemic to Macaronesia —*Aichryson* Webb & Berthel. (10 spp.), *Greenovia* Webb & Berthel. (4 spp.), and *Monanthes* Haw. (10 spp.)— and genus *Aeonium* Webb & Berthel. (40 spp.), which comprises 37 Macaronesian endemic species, one species endemic to Morocco and two E African endemics (*A. leucoblepharum* Webb ex A. Richard and *A. stuessyi* H.Y. Liu). The molecular phylogeny reconstructed by Mort et al. (2002) revealed that *Aeonium* is paraphyletic, with *Greenovia* embedded within. The Moroccan species (*A. korneliuslemsii* H.Y. Liu) and the eastern African *A. leucoblepharum* appear nested within the Canarian radiation clade, and are related, although weakly supported, to the Cape Verde endemism *A. gorgoneum* J. A. Schmidt (Mort et al., 2007; Kim et al., 2008).

3) *Campanula* L. s.s., *Azorina* group (Campanulaceae): this clade of ca. 30 species (Mansion et al., 2012) is mainly distributed in North Africa but also includes

species from the Iberian Peninsula, South Arabia, Socotra, Eastern Africa, West Asia, and Macaronesian endemics. Alarcon et al. (2013) showed that *C. jacobaea* C. Sm. ex Webb (native to the Cape Verde archipelago) and *C. balfourii* Wagn. & Vierh. (endemic to Socotra) are sister species, with northwestern African endemic *C. hypocrateriformis* Dobignard as their closest known relative.

4) Camptoloma Benth. (Scrophulariaceae) is a small genus (3 spp.) with a widely disjunct African distribution: *Camptoloma canariense* (Webb & Berthel.) Hilliard is endemic to Gran Canaria, whereas *C. lyperiiflorum* Hilliard is found in the Horn of Africa, and *C. rotundifolium* Benth. is restricted to southern Africa (i.e., Angola, and South Africa). Broad-scale phylogenetic analyses within the family (Kornhall et al., 2001; Oxelman et al., 2005) have revealed that the genus is phylogenetically and geographically positioned between the Holarctic tribe Schrophularieae and the Southern Hemisphere tribes Manuleeae and Seleagineae (see also Hilliard, 1999). South African *Phygelius capensis* E. Mey. ex Benth., is reconstructed as sister to Camptoloma, though with low support, both of them are related to tribes Teedieae (4 genera, ca. 10 spp.) and Buddlejeae (4 genera, ca. 100 spp.). *Camptoloma* is monophyletic, with South African species *C. rotundifolium* sister to a clade comprising *C. canariense* and *C. lyperiiflorum* (Kornhall et al., 2001; Oxelman et al., 2005). This wide E-W disjunction and restricted species-level distribution make Camptoloma an interesting candidate to study the effect of contraction-extinction (aridification) events across the Sahara.

5) *Campylanthus* Roth. (Plantaginaceae): this is a small paelotropical genus (ca. 20 species) of shrubs and subshrubs with a markedly disjunct distribution. Two species are endemic to Macaronesia, *C. salsoloides* (L. f.) Roth from the Canary Islands, and *C. glaber* Benth. from Cape Verde, while the remaining species occur in the Horn of Africa, East Africa, Southern Arabia and Pakistan (Hjertson et al., 2008). A

new molecular phylogeny, including 14 species, confirmed the disjunction between the Macaronesian endemics and the East Africa-South Arabian clade (Thiv et al., 2010).

6) *Canarina* L. belongs to tribe Platycodoneae (ca. 90 spp.), an early diverging lineage within family Campanulaceae with nine genera, all except Canarina confined to East and Central Asia. The genus comprises only three species showing an extreme disjunction pattern. *Canarina eminii* Aschers. ex Schweinf., and *C. abyssinica* Engl. occur in East Africa, appearing mainly in “sky islands”, but they differ in habitat and life-forms. *Canarina eminii* is an epiphytic herb endemic to the highly fragmented Afromontane forest belts, extending from the Ethiopian Highlands to Malawi. *Canarina abyssinica* is a rock dweller of upland forests, restricted to few patches in East Africa, from Southern Sudan to Tanzania. On the other side of the distribution appears *C. canariensis* (L.) Vatke., which is restricted to the edges of the laurisilva forest in the Canary Islands. The Asian genus *Ostrowskia* is the closest relative of genus *Canarina* (Mansion et al., 2012).

7) Cicer L. is a Fabaceae genus that belongs in subfamily Papilionoideae, usually with herbaceous or semi-shrubby biotypes. The genus seems to have originated in the Early Miocene Neotethys, and consists of ca. 40 species distributed throughout the Northern Hemisphere (Javadi et al., 2007). The Canarian endemic *C. canariense* A.Santos & G.P.Lewis and the Eastern African *C. cuneatum* A.Rich. (Egypt, Ethiopia, Sudan, and Saudi Arabia) form an early-diverging strongly supported clade sister to the remaining species of *Cicer*, though probably Moroccan endemic *C. atlanticum* Coss. Ex Maire (see monograph by van der Maesen, 1972), which was not sampled for DNA, also falls in this clade. Javadi et al. (2007) suggest that the geographical disjunction present in these species could be explained in terms of their ancestor having had a broader distribution in more favorable climate. They are indeed the only ones with climbing habits found in relatively humid habitats.

8) Colchicum L. (Colchicaceae): former genus *Androcymbium* Willd. is paraphyletic in relation to *Colchicum* (Vinnersten & Reeves 2003, Manning et al. 2007) and has therefore been transferred to *Colchicum* (Manning et al. 2007) together with *Bulbocodium* L. and *Merendera* Ramond. While *Colchicum* s. str., together with *Bulbocodium* and *Merendera*, are genera centered in the Mediterranean Basin and eastwards, the ca. 60 species formerly belonging to *Androcymbium* are mainly found in Africa, with several disjunctions between south, east, and northwest, and a high concentration of species (ca. 50) in southern Africa, especially in the winter-rainfall region. The disjunct distribution between south and north Africa is according to Caujapé-Castells et al. (2001) best explained by an origin in southern Africa, followed by dispersal northwards via the African "Arid Corridor", prior to the desiccation of Africa in the Miocene.

9) *Euphorbia* subgen. *Athymalus* (sects. *Anthacanthae* and *Balsamis*) (Euphorbiaceae): This subgenus —with about 150 species and a great diversity of life forms including herbs, geophyte, shrubs, small trees and numerous succulent forms— has its main center of diversity in southern Africa and a secondary center in the Horn of Africa. Most species in this group have restricted geographic ranges. The exception is *E. balsamifera* Aiton (sect. *Balsamis*), which has a wide distribution exhibiting the W-E Rand Flora pattern. *Euphorbia balsamifera* subsp. *balsamifera* is distributed in the Canary Islands and the west coast of Africa (Morocco and Mauritania), whereas *E. balsamifera* subsp. *adenensis* (Deflers) Bally occurs in eastern Africa (i.e., Ethiopia, Somalia, NE Sudan), the southern Arabian Peninsula (Yemen, Oman) and in one locality in Socotra. A third morphologically divergent (long linear leaves) subspecies, *E. balsamifera* subsp. *sepium* (N.E.Br.) Maire, occurs in the western Sahel region. Section *Balsamis* also includes taxa from the southern Arabian Peninsula (Peirson et al., 2013). Section *Anthacanthae* is composed of mainly southern African succulent species, though subsect. *Platycephalae* (e.g., *E. grantii* Oliv., *E. platycephala* Pax, *E.*

omariana M.G. Gilbert), which is sister to subsect. *Florispinae*, has a wider distribution throughout eastern Africa.

10) *Euphorbia* L. subgen. *Esula* sect. *Aphyllis* (Euphorbiaceae): This section represents one of the few introductions of the mostly northern temperate subgenus *Esula* into non-Mediterranean Africa. Section *Aphyllis* consists of two subclades — subsect. *Africanae* and subsect. *Macaronesicae*— whose names reflect their disjunct distribution (Barres et al., 2011; Riina et al., 2013). Subsection *Macaronesicae* is found in Macaronesian islands (except for the Azores) and the Atlantic coast of both Morocco and Western Sahara, from sea level to 1600(–2200) m. Species in subsect. *Africanae* occur in continental Africa (mainly central-eastern and southern Africa), southern Arabian Peninsula, Socotra, and Madagascar, from 100 to 2300 m. Members of sect. *Aphyllis* are shrubs, succulent to semi-succulent or sub-woody, much-branched or dendroid, some with branches pencil-like, often green (chlorophyllous). They grow in arid to semi-arid habitats such as stony plains and slopes, sparsely woody grasslands, loamy to sandy areas, often associated with low and open vegetation in which succulents predominate. A key physiological feature present in these species, as well as in other subclades within Euphorbia, is the crassulacean acid metabolism (CAM), a common physiological adaptation in plants growing in water limited habitats (Horn et al., 2014).

11) *Euphorbia* subgen. *Esula* sect. *Esula* (African clade): Another lineage within subgen. *Esula* present in Africa, and interesting because of some of its members exhibit a S-E Rand Flora disjunction, is a small subclade (here informally named “AfrEsula” clade) of ca. 10 species within sect. *Esula* (Riina et al., 2013). Most of the species in the AfrEsula clade are ericoid-leaved perennial herbs and shrubs (e.g. *E. ericoides* Lam., *E. genistoides* P.J. Bergius) distributed in southern and central-eastern Africa, Madagascar and Réunion, with one species reaching the Arabian Peninsula. The southern African species occur in the mesic parts of the winter-rainfall Cape

floristic region, in the summer-rainfall grasslands and forest margins of Natal, and extending into Zimbabwe and Mozambique. One species, *E. cyparissioides,* shows the broadest distribution range, occurring from Nigeria to east and south tropical Africa.

12) *Geranium* subgen. *Robertium* (Picard) Rouy (Geraniaceae): genus *Geranium* L. is divided into two subgenera (Fiz et al., 2008): *Geranium* (>380 species) and *Robertium* (24 species). Within the subgenus *Robertium*, the species endemic to Macaronesia, Morocco and the Iberian Península form the sister group to a clade formed by species from E Africa. Together these two clades are sister to a clade of Asian species. Fiz et al. (2008) suggested that the Afro-Macaronesian clade originated from Asian lineages that migrated across the Arabian Plate, after its collision with Eurasia in the Miocene.

13) *Hypericum* L. (Hypericaceae): This large genus comprises ca. 500 species distributed in most temperate regions around the world, but it is also found in tropical and subtropical mountains. In Africa there are ca. 50 species distributed in Macaronesia, NW Africa, Eastern African Mountains and S Africa. A Rand Flora disjunction based on morphological similarities has been described between the Macaronesian section (i.e., *Webbia*) and the African section *Campylosporus* (Robson, 1985). A large-scale analysis, covering nearly 250 species worldwide and nuclear and plastid markers, shows that African species do not form a monophyletic group, but appear scattered along the phylogeny in at least five different clades, indicating independent biogeographic histories or, alternatively, an ancestral presence in Africa decimated by extinction (Meseguer et al., 2013). In some clades, a disjunction can be observed between Macaronesia-NW Africa and East Africa, such as *H. somaliense* N. Robson (Somalia)-*H. tomentosum* Durand ex Steud. (Mediterranean) or *H. canariense* L. (Canary Islands)-*H. socotranum* R.D. Good (Socotra). An interesting disjunction at the population level is present in the Afromontane, closely related species *H. bequaertii* De Wild., *H. revolutum* Vahl, *H. roeperianum* G.W. Schimp. ex A. Rich., *and H.*

quartinianum A. Rich., with isolated populations at both sides of Africa: Eastern African mountains and the Cameroon volcanic line.

14) *Kleinia* Mill. (Asteraceae)*:* This gynuroid genus nested in *Senecio* L. s. l. comprises ca. 50 species, distributed mainly in E Africa, S Africa, and Madagascar, but also present in N Africa, with one species in the Canary Islands, 2 in northwest Africa (Maghreb), and 5–6 species in S Arabia and Socotra; in addition, 2–5 species are found in India and Sri Lanka. The taxonomic delimitation of *Kleinia* with genus *Senecio* has hindered the systematics of the former. A recent molecular phylogeny of Senecioneae (Pelser et al., 2007), based on nuclear ITS, showed that *Kleinia* forms a well-supported monophyletic group. Two sister lineages were recognized: one formed by Arabian, E African and S African species, and another comprising sub-Saharan and NW African species. Within the latter group, there is a clade with a W-E geographic disjunction: Canarian and NW African species —*K. neriifolia* Haw. and *K. anteuphorbium* (L.) Haw.— form a group which is sister to a clade that includes species from Southern Arabia, East Africa and Madagascar.

15) *Plocama* W. Aiton is a genus in the Rubiaceae included in tribe Putorieae with ca. 30 species, which generally grow as shrubs characterized by its fleshy indehiscent fruits. Backlund et al. (2007) reviewed tribe Putorieae identified several well-supported clades when reconstructing the phylogeny, but failed to identify morphological synapomorphies, which lead to the synonimization of a number of genera within Putorieae (e.g. Gaillonia A. Rich. ex DC., Crocyllis E. Mey. ex Benth., etc.). One of these clades consists of a number of disjunct species native to W, E and S Africa. *Plocama tinctoria* (Balf. f.) M. Backlund & Thulin (from Socotra and Somalia) and *P. yemenensis* (Thulin) M. Backlund & Thulin (from Oman and continental Yemen) form a clade sister to *P. crocyllis* (Sond.) M. Backlund & Thulin (from Namibia and South Africa). Meanwhile *P. pendula* Aiton (Canary Islands) is sister to the aforementioned species. Though Socotran *P. puberula, P. putorioides*, and *P.*

thymoides, Somali *P. calcicola* and *P. somaliensis*, and Yemeni *P. jolana* were not sample for DNA, they probably belong in the aforementioned E African-S Arabian clade (Backlund et al., 2007).

16) *Sideroxylon* L. (Sapotaceae)*:* Tribe Sideroxylae comprises ca. 80 species distributed among three genera, namely *Sideroxylon* (ca. 75 spp.), *Nesoluma* Baill. (three spp.), and *Argania* Roem. & Schult. (one sp.). Smedmark et al. (2006) showed that both *Nesoluma* and *Argania* were nested within larger *Sideroxylon*. Two sister species within this tribe are of interest to us: Moroccan *Argania spinosa* (L.) Skeels (originally described as Sideroxylon spinosum) and E African/S Arabian *Sideroxylon mascatense* (A. DC.) T.D. Penn. (Smedmark et al., 2006; Smedmark and Anderberg, 2007).

REFERENCES

Alarcón, M., Roquet, C., García-Fernández, A., Vargas, P., and Aldasoro, J. J. (2013). Phylogenetic and phylogeographic evidence for a Pleistocene disjunction between *Campanula jacobaea* (Cape Verde Islands) and *C. balfourii* (Socotra). *Mol. Phylogenet. Evol.* 69(3), 828–836.

Backlund, M., Bremer, B., and Thulin, M. (2007). Paraphyly of Paederieae, recognition of Putorieae and expansion of Plocama (Rubiaceae-Rubioideae). *Taxon* 56(2), 315–328.

Barres, L., Vilatersana, R., Molero, J., Susanna, A., and Galbany-Casals, M. (2011). Molecular phylogeny of *Euphorbia* subg. *Esula* sect. *Aphyllis* (Euphorbiaceae) inferred from nrDNA and cpDNA markers with biogeographic insights. *Taxon* 60(3), 705–720.

Caujapé-Castells, J., Jansen, R. K., Membrives, N., Pedrola-Monfort, J., Montserrat, J. M., and Ardanuy, A. (2001). Historical biogeography of *Androcymbium* willd. (Colchicaceae) in Africa: evidence from cpDNA RFLPs. *Bot. J. Linn. Soc.* 136(4), 379–392.

Cubas, P., Pardo, C., Tahiri, H., and Castroviejo, S. (2010). Phylogeny and evolutionary diversification of *Adenocarpus* DC. (Leguminosae). *Taxon* 59(3), 720–732.

Fiz, O., Vargas, P., Alarcón, M., Aedo, C., García, J. L., & Aldasoro, J. J. (2008). Phylogeny and historical biogeography of Geraniaceae in relation to climate changes and pollination ecology. *Syst. Bot.* 33(2), 326–342.

Hilliard, O. M. (1999). *The tribe Selagineae (Scrophulariaceae)*. Kew: Royal Botanic Gardens.

Hjertson, M., Henrot, J., and Thulin, M. (2008). *Campylanthus hajarensis* sp. nov. and a new record of *Campylanthus* (Scrophulariaceae) from Oman. *Nord. J. Bot.* 26(1–2), 35–37.

Horn, J. W., Xi, Z., Riina, R., Peirson, J. A., Yang, Y., Dorsey, B. L., Berry, P. E., Davis C. C., and Wurdack, K. J. (2014). Evolutionary burst in *Euphorbia* (Euphorbiaceae) are linked with Photosynthetic pathway. *Evolution*, 68(12), 3485–3504.

Javadi, F., Wojciechowski, M. F., and Yamaguchi, H. (2007). Geographical diversification of the genus *Cicer* (Leguminosae: Papilionoideae) inferred from molecular phylogenetic analyses of chloroplast and nuclear DNA sequences. *Bot. J. Linn. Soc.* 154(2), 175–186.

Kim, S. C., McGowen, M. R., Lubinsky, P., Barber, J. C., Mort, M. E., & Santos-Guerra, A. (2008). Timing and tempo of early and successive adaptive radiations in Macaronesia. *PLoS ONE* 3(5), e2139.

Kornhall, P., Heidari, N., and Bremer, B. (2001). Selagineae and Manuleeae, two tribes or one? Phylogenetic studies in the Scrophulariaceae. *Plant Syst. Evol.* 228(3–4), 199–218.

Manning, J., Forest, F., and Vinnersten, A. (2007). The genus *Colchicum* L. redefined to include *Androcymbium* Willd. based on molecular evidence. *Taxon* 56(3), 872-882.

Mansion, G., Parolly, G., Crowl, A. A., Mavrodiev, E., Cellinese, N., Oganesian, M., Fraunhofer, K., Kamari, G., Phitos, D., Haberle, R., Akaydin, G., Ikinci, N., Raus, T., and Borsch, T. (2012). How to handle speciose clades? Mass taxon-sampling as a strategy towards illuminating the natural history of *Campanula* (Campanuloideae). *PloS ONE* 7(11), e50076.

Meseguer, A. S., Aldasoro, J. J., and Sanmartín, I. (2013). Bayesian inference of phylogeny, morphology and range evolution reveals a complex evolutionary history in St. John's wort (*Hypericum*). *Mol. Phylogenet. Evol.* 67(2), 379–403.

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

Mort, M. E., Soltis, D. E., Soltis, P. S., Santos-Guerra, A., and Francisco-Ortega, J. (2007). Physiological evolution and association between physiology and growth form in *Aeonium* (Crassulaceae). *Taxon*, 453–464.

Oxelman, B., Kornhall, P., Olmstead, R. G., and Bremer, B. (2005). Further disintegration of Scrophulariaceae. *Taxon* 54(2), 411–425.

Peirson, J. A., Bruyns, P. V., Riina, R., Morawetz, J. J., and Berry, P. E. (2013). A molecular phylogeny and classification of the largely succulent and mainly African *Euphorbia* subg. *Athymalus* (Euphorbiaceae). *Taxon* 62(6), 1178–1199.

Pelser, P. B., Nordenstam, B., Kadereit, J. W., and Watson, L. E. (2007). An ITS phylogeny of tribe Senecioneae (Asteraceae) and a new delimitation of Senecio L. *Taxon 56*(4), 1077–1077.

Percy, D. M., and Cronk, Q. C. (2002). Different fates of island brooms: contrasting evolution in *Adenocarpus*, *Genista*, and *Teline* (Genisteae, Fabaceae) in the Canary Islands and Madeira. *Am. J. Bot.* 89(5), 854–864.

Riina, R., Peirson, J. A., Geltman, D. V., Molero, J., Frajman, B., Pahlevani, A., Barres, L., Morawetz, J. J., Salmaki, Y., Zarre, S., Kryukov, A., Bruyns, P. V., and Berry,

P. E. (2013). A worldwide molecular phylogeny and classification of the leafy spurges, *Euphorbia* subgenus *Esula* (Euphorbiaceae). *Taxon* 62(2), 316–342.

Robson, N.K.B. 1985. Studies in the genus *Hypericum* L. (Guttiferae). 3. Sections 1. *Campylosporus* to 6a. *Umbraculoides*. *Bull. Br. Mus. (Nat. Hist.) Bot. Ser.* 12, 163–211.

Smedmark, J. E., Swenson, U., and Anderberg, A. A. (2006). Accounting for variation of substitution rates through time in Bayesian phylogeny reconstruction of Sapotoideae (Sapotaceae). *Mol. Phylogenet. Evol.* 39(3), 706–721.

Smedmark, J. E., and Anderberg, A. A. (2007). Boreotropical migration explains hybridization between geographically distant lineages in the pantropical clade Sideroxyleae (Sapotaceae). *Am. J. Bot.* 94(9), 1491–1505.

Thiv, M., Thulin, M., Hjertson, M., Kropf, M., and Linder, H. P. (2010). Evidence for a vicariant origin of Macaronesian–Eritreo/Arabian disjunctions in *Campylanthus* Roth (Plantaginaceae). *Mol. Phylogenet. Evol.* 54(2), 607–616.

van der Maesen, L. J. G. (1972). *Cicer* L., a monograph of the genus, with special reference to the chickpea (*Cicer arietinum* L.), its ecology and cultivation. Wageningen: Veenman H. and Zonen N.V., 341 p.

Vinnersten, A., and Reeves, G. (2003). Phylogenetic relationships within Colchicaceae. *Am. J. Bot.* 90(10), 1455–1462.

Table S1. Net diversification rates (bd.ms) for all RF disjunct clades and their encompassing lineages (bold = highest crown.p, red when n≤2) under three possible scenarios: no extinction (Ɛ = 0), turnover at equilibrium (Ɛ = 0.5), and high extinction (Ɛ = 0.9). Probability (crown.p) of obtaining a clade with the same size and age as the RF disjunction, given the background diversification rate of the encompassing clade/s and at increasing extinction fractions (bold = highest crown.p, italics p<0.05). Stem and Crown ages in Myr.

				bd.ms			**crown.p**		
Lineage	**Richness**	**Stem**	**Crown**	**0.0**	**0.5**	**0.9**	**0.0**	**0.5**	**0.9**
Ad. mannii	15	5.3	3.7	**0.54456840**	0.48213740	0.22226870	**0.42537290**	0.39819950	0.38212080
Adenocarpus	30	9.4	6.6	0.41031060	0.37111330	**0.19830660**	0.15449390	0.21772780	**0.32660400**
Genisteae	560	41.2	24.0	0.23478290	0.22286220	**0.16624820**	*0.00548584*	*0.03543695*	**0.24812990**
Fabaceae	18860	71.0	61.0	0.15002710	0.14531170	**0.12280980**	*0.00012871*	*0.00371182*	**0.14011130**
Angiosperms	262196	–	132.0	0.08927046	0.08709107	**0.07668942**	*0.00000071*	*0.00010133*	***0.04302373***
Ae. leucoblepharum	**7**	3.0	1.7	**0.73691940**	0.63558720	0.24384520	**0.45155500**	0.42202850	0.40046890
Aeonium	40	14.2	10.4	0.28805120	0.26249090	**0.14781760**	*0.03536212*	0.09026290	**0.23454930**
Aeonium alliance	60	18.8	15.3	0.22230050	0.20445530	**0.12286310**	*0.01366463*	*0.04853158*	**0.18531150**
Crassulaceae	1500	56.0	43.0	0.15395520	0.14727870	**0.11547230**	*0.00313954*	*0.01889902*	**0.17038110**
Angiosperms	262196	–	132.0	0.08927046	0.08709107	**0.07668942**	*0.00029312*	*0.00315710*	**0.09121793**
Ca. jacobea	**2.01**	1.5	1.0	**0.00498754**	0.00496327	0.00386251	**0.95778500**	0.95494310	0.91589870
Azorina clade	26	11.2	10.1	**0.25395540**	0.22877140	0.11885870	**0.99280450**	0.80424570	0.70180750
Campanula	500	42.5	40.0	**0.13803650**	0.13088890	0.09696299	**0.98825780**	0.83944730	0.68768560
Campanulaceae	2400	80.0	76.0	**0.09329048**	0.08951007	0.07148787	**0.98499850**	0.86787070	0.67857540
Angiosperms	262196	–	132.0	**0.08927046**	0.08709107	0.07668942	**0.98462000**	0.86987020	0.67932940
Cm. rotundifolium	**2.01**	5.5	4.1	**0.00121647**	0.00121055	0.00094207	**0.95778500**	0.95494310	0.91589870
Cm. canariense	**3**	10.2	5.5	**0.07372093**	0.06395225	0.02141510	**0.55555560**	0.52082350	0.47325100
Camptoloma + Phygelius	**4**	10.6	10.2	**0.06795561**	0.05811604	0.01985904	**0.52645660**	0.50135390	0.46083480
Buddlejoideae	150	25.5	13.3	**0.32462320**	0.30343650	0.20412470	**0.97186820**	0.87383000	0.91676350

				bd.ms			**crown.p**		
Lineage	**Richness**	**Stem**	**Crown**	**0.0**	**0.5**	**0.9**	**0.0**	**0.5**	**0.9**
Scrophulariaceae	4800	43.0	38.5	**0.20216170**	0.19469420	0.15907430	**0.89180050**	0.77090530	0.88183380
Angiosperms	262196	–	132.0	0.08927046	0.08709107	**0.07668942**	0.62542940	0.58489290	**0.74284480**
Cy. salsoloides	18	13.7	5.1	**0.43082830**	0.38375150	0.18441020	**0.42194630**	0.39510270	0.37969900
Campylanthus	20	20.0	7.5	0.30701130	0.27438570	**0.13494740**	0.10214550	0.16017020	**0.22844080**
Globularieae	40	18.8	13.7	0.21866660	0.19926320	**0.11221190**	*0.01084014*	*0.04471008*	**0.15760690**
Plantaginaceae	1600	61.0	42.0	0.15915740	0.15232110	**0.11974940**	*0.00067409*	*0.01078351*	**0.18079200**
Angiosperms	262196	–	132.0	0.08927046	0.08709107	**0.07668942**	*0.00000114*	*0.00018943*	**0.06041009**
Cn. canariensis	**2.01**	8.2	6.5	**0.00076731**	0.00076358	0.00059423	**0.95778500**	0.95494310	0.91589870
Canarina	**3**	13.8	8.2	**0.04944696**	0.04289480	0.01436378	**0.99432670**	0.79735980	0.68574300
Platycodoneae	90	41.9	29.1	**0.13081310**	0.12126410	0.07700898	**0.99869090**	0.84497210	0.88199630
Campanulaceae	2400	80.0	76.0	**0.09329048**	0.08951007	0.07148787	**0.99755980**	0.81627100	0.87301420
Angiosperms	262196	–	132.0	**0.08927046**	0.08709107	0.07668942	**0.99738120**	0.81420940	0.88150240
Ci. canariense	**3**	12.2	3.4	**0.11925440**	0.10345220	0.03464207	**0.55555560**	0.52082350	0.47325100
Cicer	30	17.5	12.2	**0.22197130**	0.20076620	0.10728060	**0.77895850**	0.66313190	0.71042020
Vicioid clade	860	33.0	32.1	0.18890300	0.17997310	**0.13749020**	0.72322220	0.63828420	**0.76513850**
Fabaceae	18860	71.0	61.0	0.15002710	0.14531170	**0.12280980**	0.63947140	0.59148560	**0.74062200**
Angiosperms	262196	–	132.0	0.08927046	0.08709107	**0.07668942**	0.45503790	0.48584890	**0.63385480**
Co. schimperianum	**5**	6.3	4.9	**0.18699810**	0.15998920	0.05691240	**0.47520000**	0.44385090	0.41688380
Colchicum	160	32.3	25.5	0.17184420	0.16077950	**0.10885740**	0.42269740	0.44584400	**0.62769770**
Colchiceae	185	48.0	43.3	0.10455450	0.09802111	**0.06729433**	0.18023150	0.26406810	**0.46976100**
Colchicaceae	280	86.5	67.3	0.07342708	0.06919952	**0.04921930**	0.08535675	0.16504810	**0.37322020**
Angiosperms	262196	–	132.0	0.08927046	0.08709107	**0.07668942**	0.13062990	0.22741510	**0.51230800**
Eu. schimperiana	20	8.2	7.1	**0.32430780**	0.28984410	0.14255010	**0.42026500**	0.39358480	0.37850870
Esula	96	11.0	8.5	**0.45543540**	0.42267230	0.27056700	**0.82890250**	0.71703890	0.73821350
Euphorbia	2000	54.7	47.8	0.14451370	0.13850460	**0.10986400**	*0.00252786*	*0.02451810*	**0.24556170**
Euphorbiaceae	6600	69.2	64.6	0.12541300	0.12096180	**0.09972609**	*0.00062042*	*0.01108944*	**0.20327080**

Lineage	Richness	Stem	Crown	bd.ms			crown.p		
				0.0	0.5	0.9	0.0	0.5	0.9
Angiosperms	262196	–	132.0	0.08927046	0.08709107	**0.07668942**	*0.00001293*	*0.00108248*	**0.11175160**
E. balsamifera	**3**	7.2	3.8	**0.10670130**	0.09256247	0.03099554	**0.55555560**	0.52082350	0.47325100
E. omariana	30	8.2	7.4	**0.36595270**	0.33099290	0.17686810	**0.41534300**	0.38914730	0.37501580
Anthacanthae (+*Balsamis*)	170	11.9	10.8	**0.41135660**	0.38520160	0.26234810	**0.95612060**	0.84510890	0.90099900
Euphorbia	2000	54.7	47.8	0.14451370	0.13850460	**0.10986400**	0.66656400	0.60508700	**0.74058510**
Euphorbiaceae	6600	69.2	64.6	0.12541300	0.12096180	**0.09972609**	0.61447090	0.57617230	**0.71903000**
Angiosperms	262196	N/A	132.0	0.08927046	0.08709107	**0.07668942**	0.49259990	0.50841650	**0.65921960**
Eu. usambarica	12	5.6	4.4	**0.40721810**	0.35776220	0.15644910	**0.43068160**	0.40300640	0.38586230
Eu. tuckeyana	23	9.3	5.6	**0.43613340**	0.39146610	0.19834370	**0.41831760**	0.39182810	0.37712820
Aphyllis + *Exiguae* II	26	12.8	10.4	0.24662970	0.22217220	**0.11543010**	*0.01439753*	*0.04916386*	**0.12601220**
Euphorbia	2000	54.7	47.8	0.14451370	0.13850460	**0.10986400**	*0.00004386*	*0.00222117*	**0.11074760**
Euphorbiaceae	6600	69.2	64.6	0.12541300	0.12096180	**0.09972609**	*0.00000658*	*0.00071004*	**0.08463664**
Angiosperms	262196	N/A	132.0	0.08927046	0.08709107	**0.07668942**	*0.00000004*	*0.00002784*	***0.03636881***
G. robertianum	20	11.0	8.1	**0.28426980**	0.25406090	0.12495130	**0.42026500**	0.39358480	0.37850870
Robertium	30	15.3	11.0	0.24618640	0.22266800	**0.11898400**	0.28105320	0.30437140	**0.34601080**
Geranium	430	26.4	15.3	0.35102210	0.33235420	**0.24382760**	0.75725780	0.67324010	**0.76906770**
Geraniaceae	800	56.0	39.0	0.15362730	0.14627930	**0.11133050**	*0.01250196*	0.05990260	**0.30877360**
Angiosperms	262196	N/A	132.0	0.08927046	0.08709107	**0.07668942**	*0.00002672*	*0.00193325*	**0.13973240**
H. quartinianum	**2.01**	4.0	1.3	**0.00383657**	0.00381790	0.00297116	**0.95778500**	0.95494310	0.91589870
Campylosporus	**5**	16.0	4.0	**0.22907270**	0.19598680	0.06971769	**0.99385050**	0.79998560	0.68437510
H. scopulorum	**10**	21.0	17.3	**0.09303109**	0.08120901	0.03397611	**0.43620760**	0.40802270	0.38974430
Androsaemum	15	27.1	21.0	**0.09594776**	0.08494803	0.03916164	**0.46646460**	0.43687500	0.45206240
Hypericum	500	49.9	34.9	0.15820800	0.15001590	**0.11113240**	0.49518800	0.50072870	**0.70362740**
Hypericaceae	585	61.0	53.9	0.10535180	0.10004270	**0.07482297**	0.16194090	0.25492070	**0.53140710**
Angiosperms	262196	N/A	132.0	0.08927046	0.08709107	**0.07668942**	0.08694648	0.18865990	**0.54252850**
K. neriifolia	**10**	11.2	6.8	0.23668200	0.20660530	0.08643922	**0.43620760**	0.40802270	0.38974430

				bd.ms			**crown.p**		
Lineage	**Richness**	**Stem**	**Crown**	**0.0**	**0.5**	**0.9**	**0.0**	**0.5**	**0.9**
Kleinia	60	14.8	11.8	0.28823710	0.26509890	**0.15930560**	0.63133870	0.57122420	**0.66096100**
Senecioneae	3000	27.0	25.5	0.28679300	0.27552290	**0.22178340**	0.62647160	0.59647250	**0.79029560**
Asteraceae	23000	44.0	40.0	0.23375260	0.22656150	**0.19224400**	0.42414780	0.46759840	**0.73754260**
Angiosperms	262196	N/A	132.0	0.08927046	0.08709107	**0.07668942**	*0.00985193*	0.05161934	**0.33947270**
P. crocyllis	**9**	6.2	3.7	0.40650740	0.35354310	0.14415430	**0.44002160**	0.41149290	0.39241680
P. pendula	**10**	9.5	6.2	0.25958680	0.22659930	0.09480430	**0.43620760**	0.40802270	0.38974430
Plocama	34	22.7	10.2	0.27776600	0.25207530	**0.13789870**	0.50303780	0.47700480	**0.55683810**
Putorieae	36	34.0	31.9	0.09060727	0.08234886	**0.04551950**	*0.00647617*	*0.03079895*	**0.13585340**
Rubiaceae	13600	65.0	57.0	0.15481890	0.14977300	**0.12569480**	0.08541263	0.18312110	**0.51502430**
Angiosperms	262196	N/A	132.0	0.08927046	0.08709107	**0.07668942**	*0.00596619*	*0.03764727*	**0.30256190**
S. spinosus	**2.01**	47.3	17.5	**0.00028500**	0.00028362	0.00022071	**0.95778500**	0.95494310	0.91589870
African clade	20	50.5	47.3	**0.04868045**	0.04350725	0.02139758	**0.99869610**	0.84122420	0.84539730
Sideroxylon	80	74.5	56.3	**0.06552184**	0.06060775	0.03791194	**0.99934930**	0.88020500	0.91303010
Sapotaceae	800	77.0	74.5	**0.08042234**	0.07657574	0.05828039	**0.99963740**	0.90931830	0.95039740
Angiosperms	262196	–	132.0	**0.08927046**	0.08709107	0.07668942	**0.99974150**	0.92471130	0.96791170

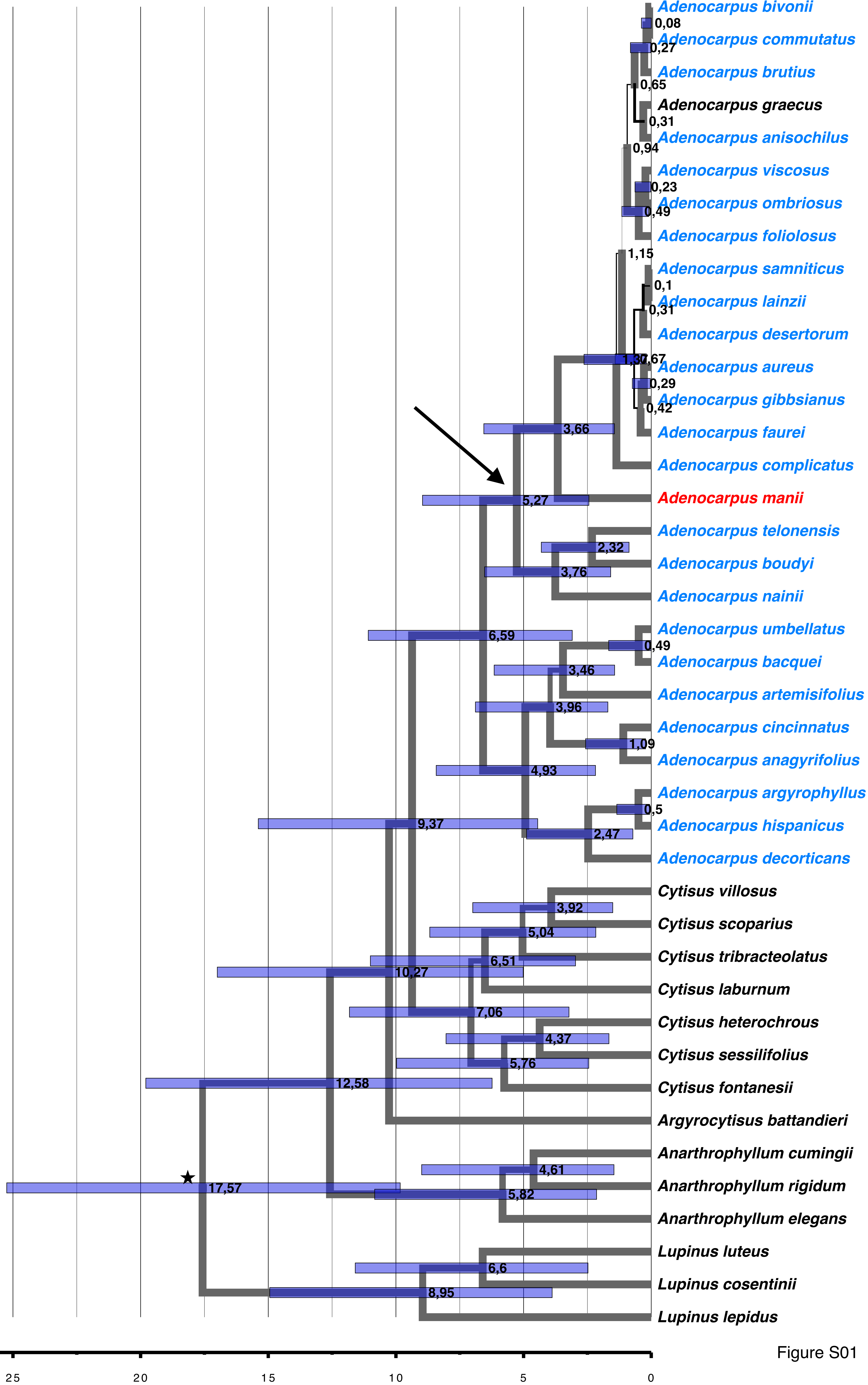

Adenocarpus bivonii
0,08
Adenocarpus commutatus
0,27
Adenocarpus brutius
0,65
Adenocarpus graecus
0,31
Adenocarpus anisochilus
0,94
Adenocarpus viscosus
0,23
Adenocarpus ombriosus
0,49
Adenocarpus foliolosus
1,15
Adenocarpus samniticus
0,1
Adenocarpus lainzii
0,31
Adenocarpus desertorum
1,3
0,67
Adenocarpus aureus
0,29
Adenocarpus gibbsianus
0,42
3,66
Adenocarpus faurei
Adenocarpus complicatus
5,27
Adenocarpus manii
Adenocarpus telonensis
2,32
Adenocarpus boudyi
3,76
Adenocarpus nainii
6,59
Adenocarpus umbellatus
0,49
Adenocarpus bacquei
3,46
Adenocarpus artemisifolius
3,96
Adenocarpus cincinnatus
1,09
Adenocarpus anagyrifolius
4,93
Adenocarpus argyrophyllus
0,5
9,37
Adenocarpus hispanicus
2,47
Adenocarpus decorticans
Cytisus villosus
3,92
Cytisus scoparius
5,04
6,51
Cytisus tribracteolatus
10,27
Cytisus laburnum
7,06
Cytisus heterochrous
4,37
Cytisus sessilifolius
5,76
12,58
Cytisus fontanesii
Argyrocytisus battandieri
Anarthrophyllum cumingii
4,61
★
17,57
Anarthrophyllum rigidum
5,82
Anarthrophyllum elegans
Lupinus luteus
6,6
Lupinus cosentinii
8,95
Lupinus lepidus
25
20
15
10
5
0

Figure S01

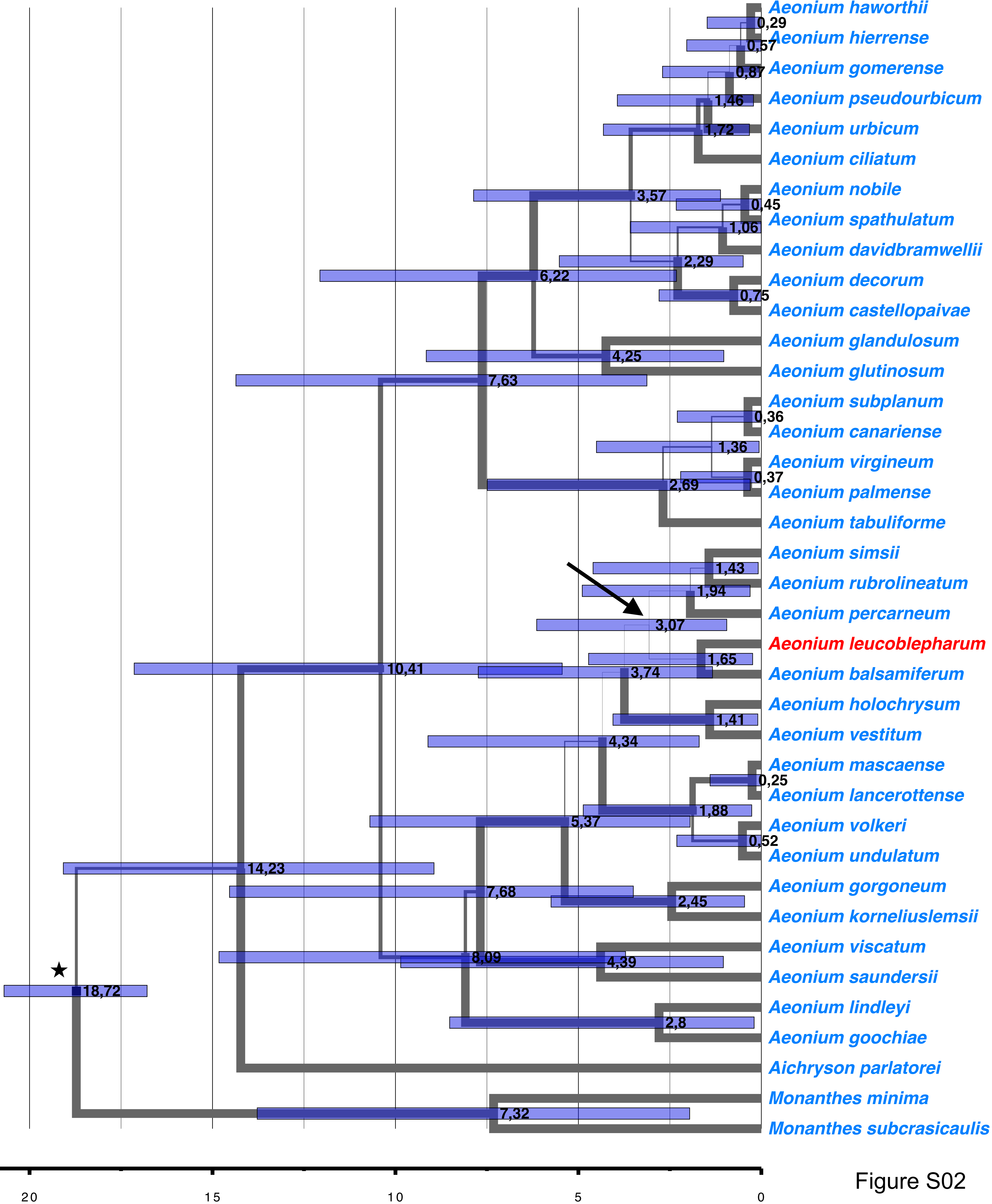

Aeonium haworthii
0,29
Aeonium hierrense
0,57
Aeonium gomerense
0,87
Aeonium pseudourbicum
1,46
Aeonium urbicum
1,72
Aeonium ciliatum
3,57
Aeonium nobile
0,45
Aeonium spathulatum
1,06
Aeonium davidbramwellii
2,29
6,22
Aeonium decorum
0,75
Aeonium castellopaivae
Aeonium glandulosum
4,25
Aeonium glutinosum
7,63
Aeonium subplanum
0,36
Aeonium canariense
1,36
Aeonium virgineum
0,37
2,69
Aeonium palmense
Aeonium tabuliforme
Aeonium simsii
1,43
Aeonium rubrolineatum
1,94
Aeonium percarneum
3,07
Aeonium leucoblepharum
1,65
10,41
Aeonium balsamiferum
3,74
Aeonium holochrysum
1,41
Aeonium vestitum
4,34
Aeonium mascaense
0,25
Aeonium lancerottense
1,88
5,37
Aeonium volkeri
0,52
Aeonium undulatum
14,23
Aeonium gorgoneum
7,68
2,45
Aeonium korneliuslemsii
Aeonium viscatum
8,09
4,39
★
18,72
Aeonium saundersii
Aeonium lindleyi
2,8
Aeonium goochiae
Aichryson parlatorei
Monanthes minima
7,32
Monanthes subcrasicaulis
20
15
10
5
0

Figure S02

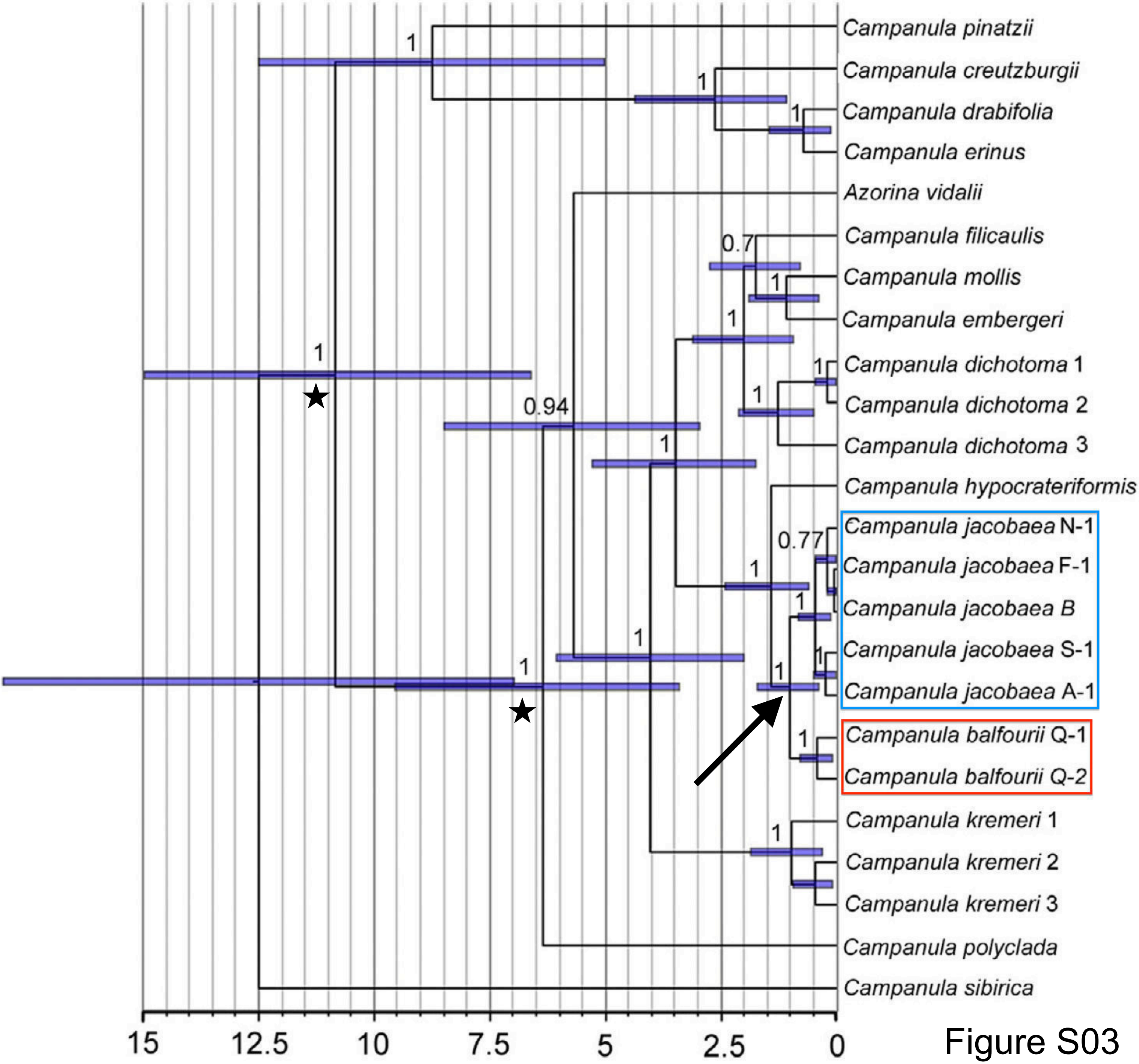

Campanula pinatzii
Campanula creutzburgii
Campanula drabifolia
Campanula erinus
Azorina vidalii
Campanula filicaulis
Campanula mollis
Campanula embergeri
Campanula dichotoma 1
Campanula dichotoma 2
Campanula dichotoma 3
Campanula hypocrateriformis
Campanula jacobaea N-1
Campanula jacobaea F-1
Campanula jacobaea B
Campanula jacobaea S-1
Campanula jacobaea A-1
Campanula balfourii Q-1
Campanula balfourii Q-2
Campanula kremeri 1
Campanula kremeri 2
Campanula kremeri 3
Campanula polyclada
Campanula sibirica
1
1
1
1
0.7
1
1
1
1
0.94
1
1
0.77
1
1
1
1
1
1
1
15
12.5
10
7.5
5
2.5
0

Figure S03

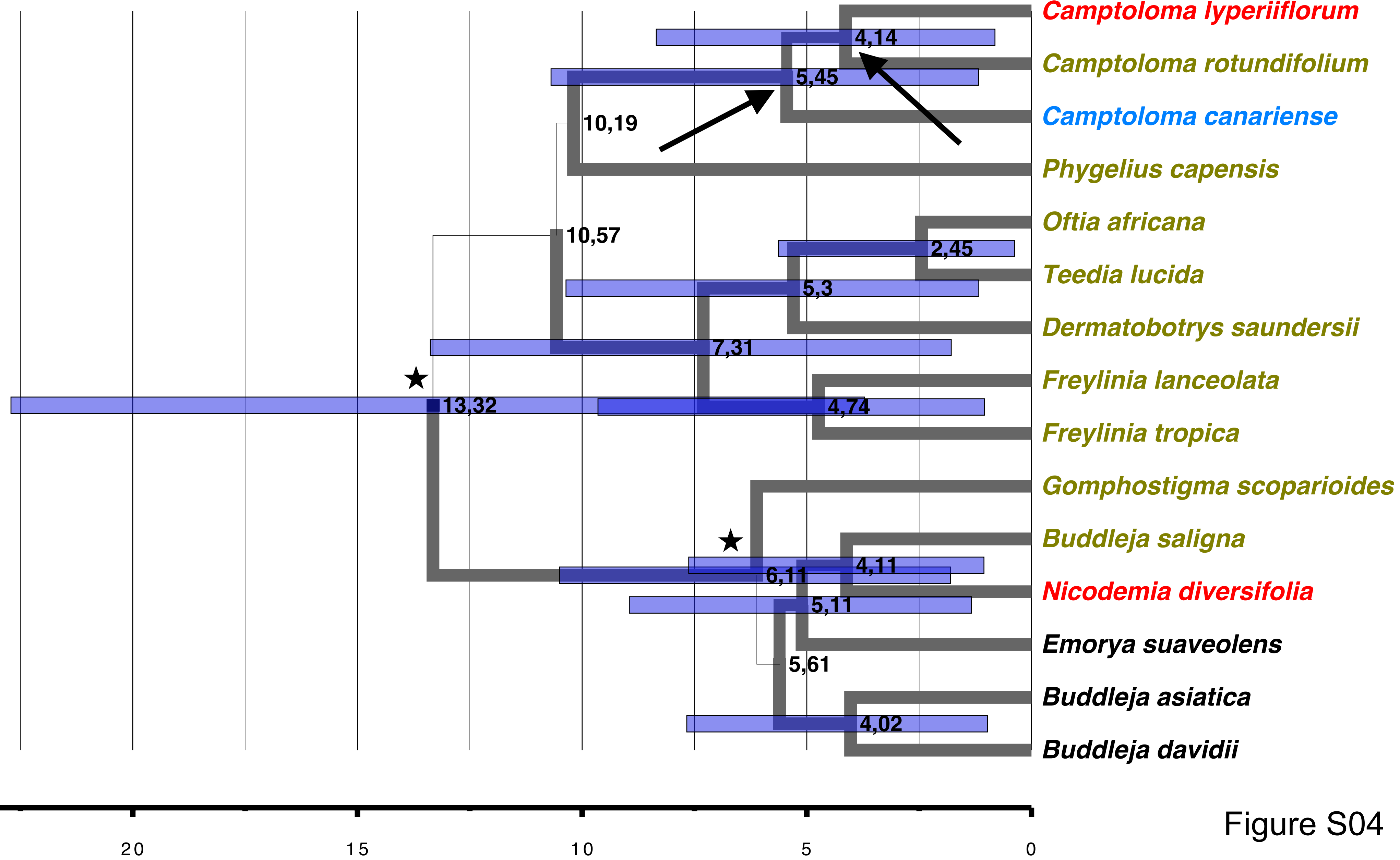

Camptoloma lyperiiflorum
Camptoloma rotundifolium
Camptoloma canariense
Phygelius capensis
Oftia africana
Teedia lucida
Dermatobotrys saundersii
Freylinia lanceolata
Freylinia tropica
Gomphostigma scoparioides
Buddleja saligna
Nicodemia diversifolia
Emorya suaveolens
Buddleja asiatica
Buddleja davidii
4,14
5,45
10,19
10,57
2,45
5,3
7,31
13,32
4,74
4,11
6,11
5,11
5,61
4,02
★
★
20
15
10
5
0

Figure S04

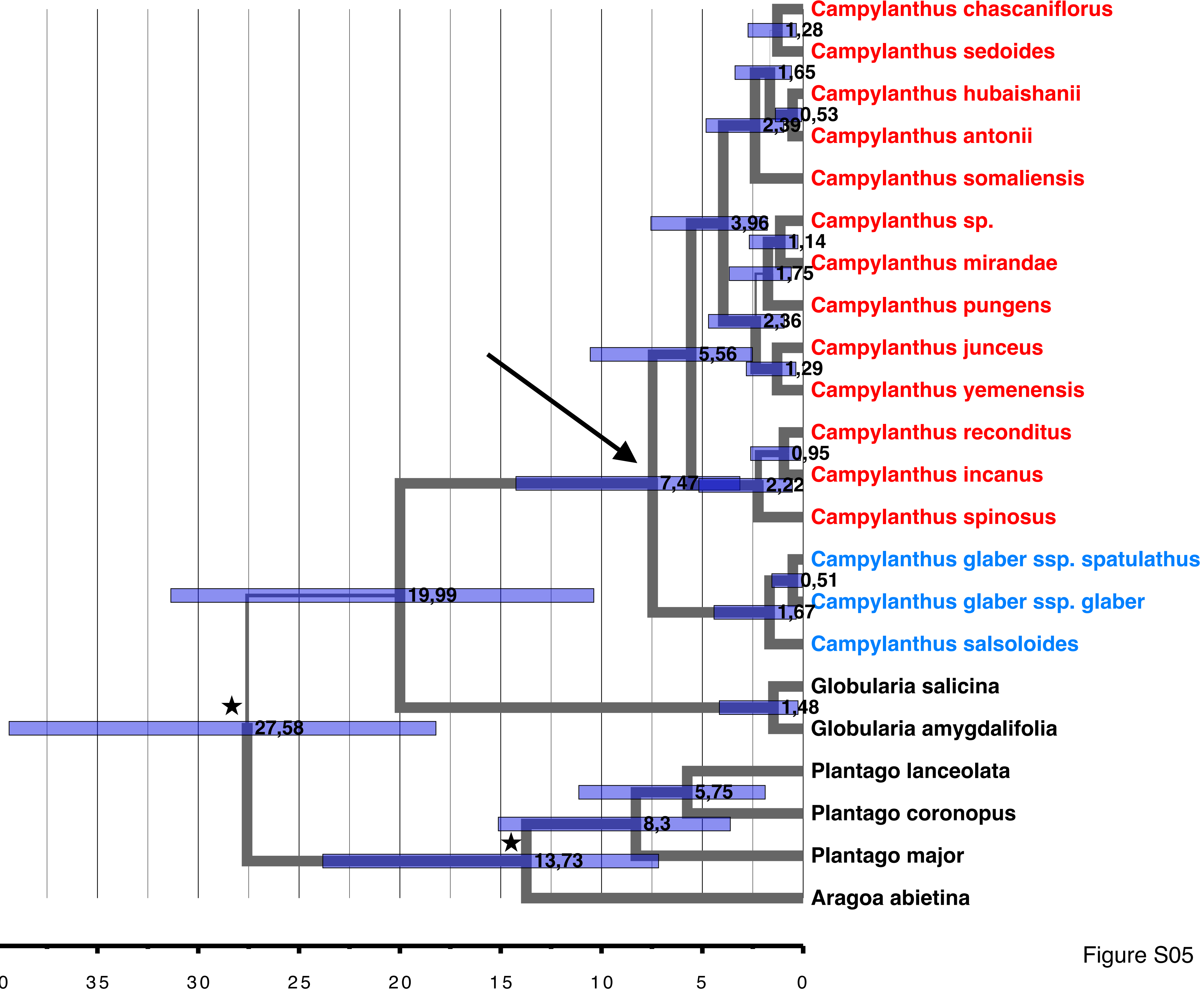

Campylanthus chascaniflorus
1,28
Campylanthus sedoides
1,65
Campylanthus hubaishanii
0,53
2,39
Campylanthus antonii
Campylanthus somaliensis
3,96
Campylanthus sp.
1,14
1,75
Campylanthus mirandae
Campylanthus pungens
2,36
5,56
Campylanthus junceus
1,29
Campylanthus yemenensis
Campylanthus reconditus
0,95
Campylanthus incanus
7,47
2,22
Campylanthus spinosus
Campylanthus glaber ssp. spatulathus
0,51
19,99
Campylanthus glaber ssp. glaber
1,67
Campylanthus salsoloides
Globularia salicina
1,48
★
27,58
Globularia amygdalifolia
Plantago lanceolata
5,75
Plantago coronopus
8,3
★
13,73
Plantago major
Aragoa abietina
0
35
30
25
20
15
10
5
0

Figure S05

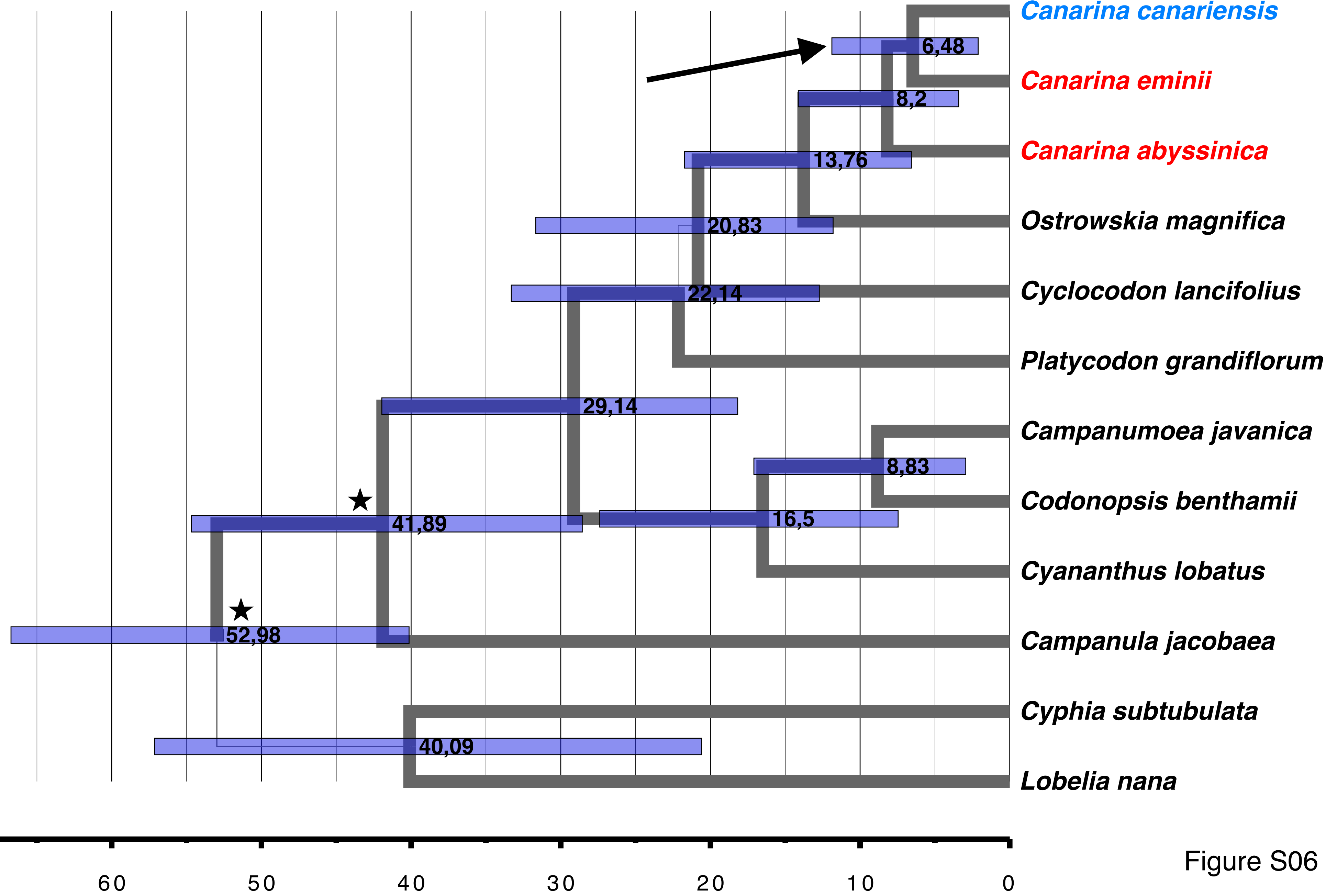

Canarina canariensis
6,48
Canarina eminii
8,2
Canarina abyssinica
13,76
Ostrowskia magnifica
20,83
Cyclocodon lancifolius
22,14
Platycodon grandiflorum
29,14
Campanumoea javanica
8,83
Codonopsis benthamii
★
41,89
16,5
Cyananthus lobatus
★
52,98
Campanula jacobaea
Cyphia subtubulata
40,09
Lobelia nana
60
50
40
30
20
10
0

Figure S06

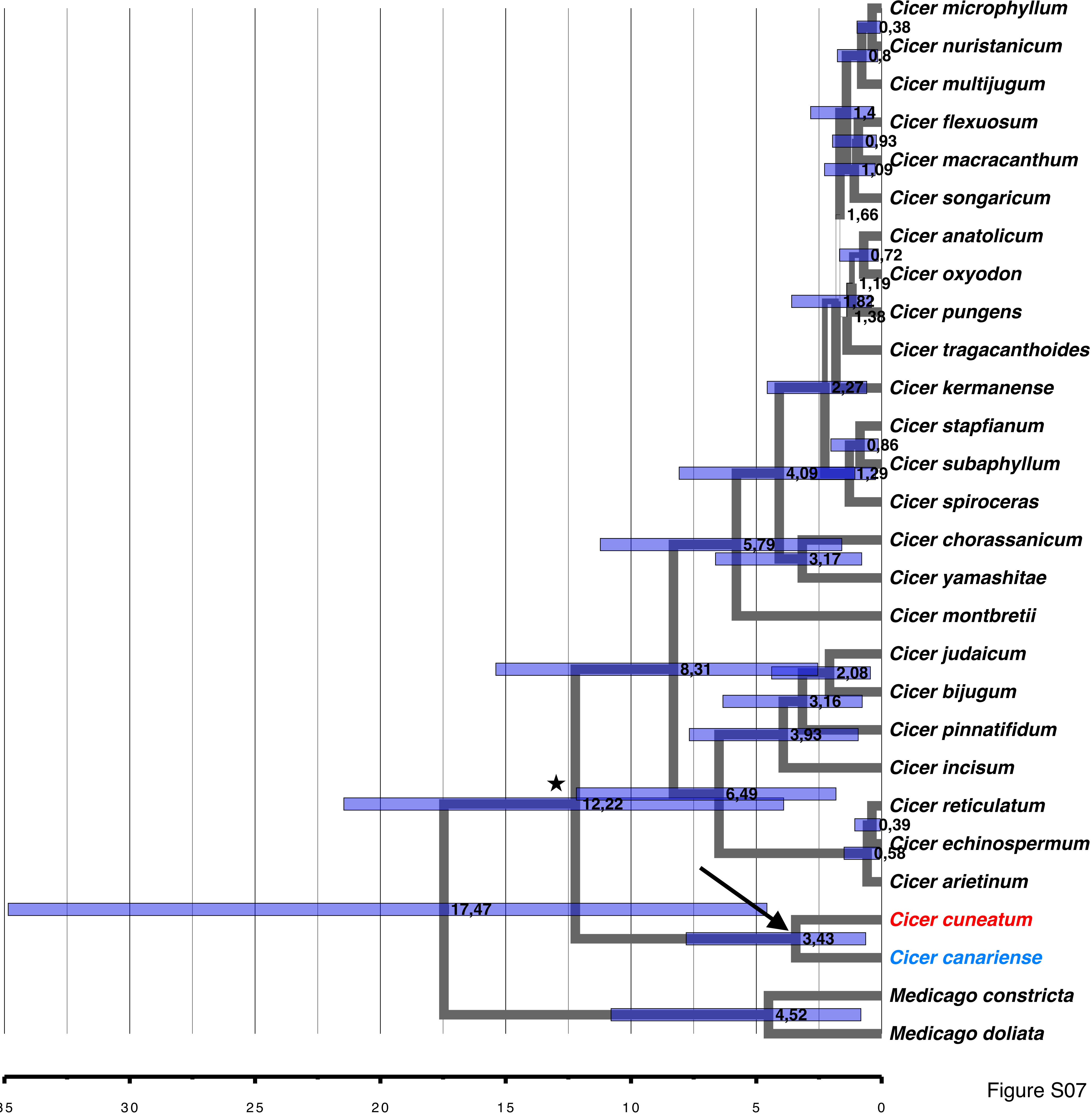

Cicer microphyllum
0,38
Cicer nuristanicum
0,8
Cicer multijugum
1,4
Cicer flexuosum
0,93
Cicer macracanthum
1,09
Cicer songaricum
1,66
Cicer anatolicum
0,72
Cicer oxyodon
1,19
1,82
Cicer pungens
1,38
Cicer tragacanthoides
2,27
Cicer kermanense
Cicer stapfianum
0,86
Cicer subaphyllum
4,09
1,29
Cicer spiroceras
Cicer chorassanicum
5,79
3,17
Cicer yamashitae
Cicer montbretii
Cicer judaicum
8,31
2,08
Cicer bijugum
3,16
Cicer pinnatifidum
3,93
Cicer incisum
★
6,49
12,22
Cicer reticulatum
0,39
Cicer echinospermum
0,58
Cicer arietinum
17,47
Cicer cuneatum
3,43
Cicer canariense
Medicago constricta
4,52
Medicago doliata
35
30
25
20
15
10
5
0

Figure S07


Colchicum hierrense
1,94
Colchicum psammophilum
3,26
Colchicum wyssianum
4,34
Colchicum gramineum
5,03
Colchicum rechingeri
6,33
3,03
Colchicum palaestinum
9,97
Colchicum roseum
Colchicum asteroides
1,23
Colchicum decipiens
12,93
Colchicum longipes
1,04
Colchicum albanense
4,86
Colchicum schimperianum
6,32
Colchicum austrocapense
8,43
Colchicum leistneri
Colchicum ciliolatum
4,96
Colchicum circinatum ssp. circinatum
6,56
Colchicum circinatum ssp. vestitum
1,26
Colchicum vanjaarsveldii
24,95
9,67
Colchicum capense
2,98
Colchicum villosum
12,71
Colchicum cedarbergense
6
Colchicum irroratum
7,64
Colchicum hantamense
15,22
10,83
Colchicum praeirroratum
Colchicum burchellii ssp. burchellii
20,19
4,42
Colchicum burchellii ssp. pulchrum
22,38
Colchicum henssenianum
Colchicum eghimocymbion
Colchicum androcymbioides
6,06
Colchicum lusitanum
31,03
9,86
Colchicum filifolium
5,45
Colchicum montanum
14,01
Colchicum autumnale
12,61
Colchicum luteum
5,25
Colchicum kesselringii
7,62
Colchicum robustum
17,9
Colchicum dregei
5,14
Colchicum exiguum ssp. exiguum
★
8,45
39,93
Colchicum vogelii
22,66
16,22
Colchicum cuspidatum
Colchicum bellum
5,8
Colchicum huntleyi
23,62
13,39
Colchicum walteri
9,27
Colchicum poeltianum
Colchicum melanthioides ssp. melanthioides
9,48
Colchicum melanthioides var. striatum
Gloriosa superba
50
40
30
20
10
0

Figure S08

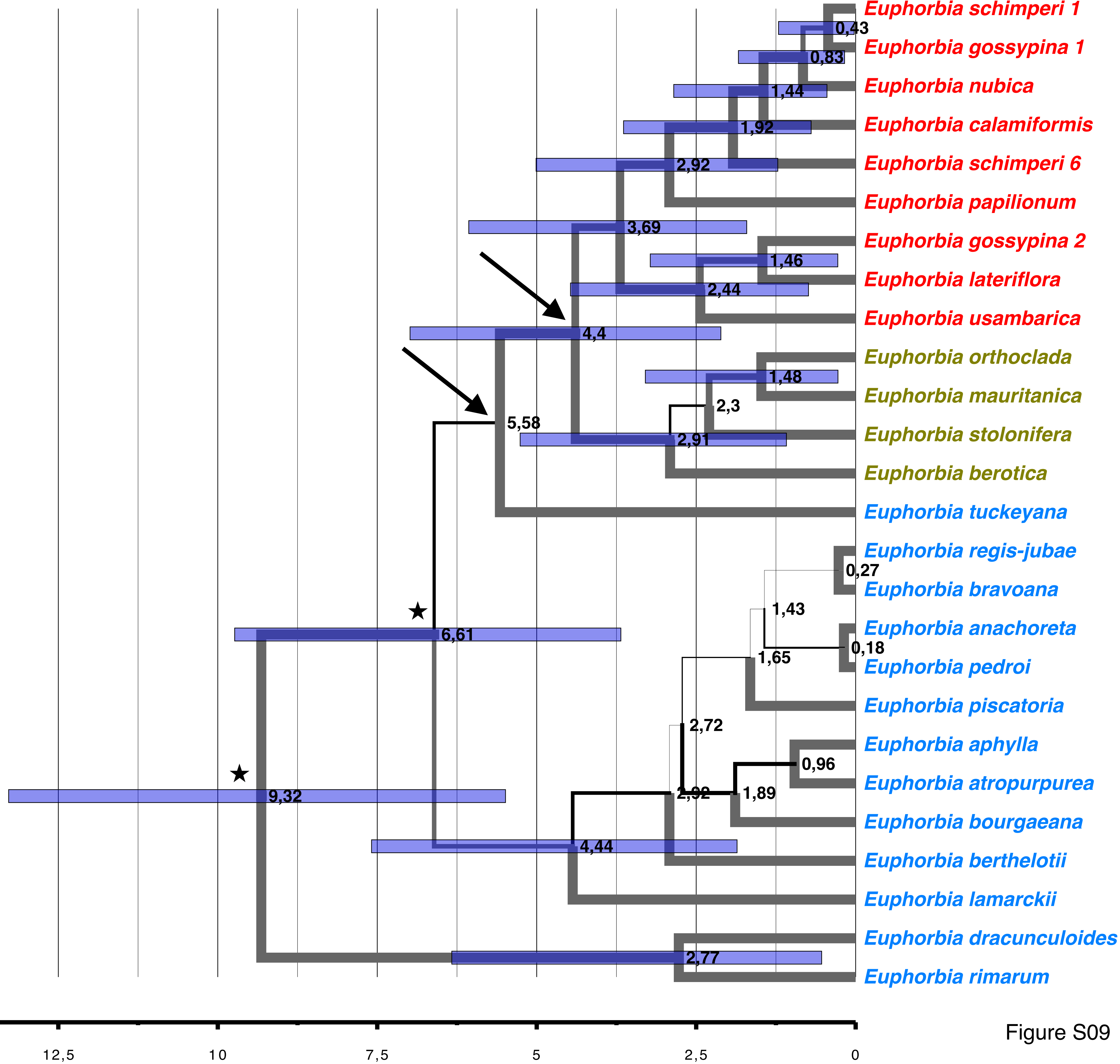

Euphorbia schimperi 1
0,43
Euphorbia gossypina 1
0,83
Euphorbia nubica
1,44
Euphorbia calamiformis
1,92
Euphorbia schimperi 6
2,92
Euphorbia papilionum
3,69
Euphorbia gossypina 2
1,46
Euphorbia lateriflora
2,44
Euphorbia usambarica
4,4
Euphorbia orthoclada
1,48
Euphorbia mauritanica
2,3
5,58
Euphorbia stolonifera
2,91
Euphorbia berotica
Euphorbia tuckeyana
Euphorbia regis-jubae
0,27
Euphorbia bravoana
1,43
★
6,61
Euphorbia anachoreta
0,18
1,65
Euphorbia pedroi
Euphorbia piscatoria
2,72
Euphorbia aphylla
0,96
★
9,32
Euphorbia atropurpurea
2,92
1,89
Euphorbia bourgaeana
4,44
Euphorbia berthelotii
Euphorbia lamarckii
Euphorbia dracunculoides
2,77
Euphorbia rimarum
12,5
10
7,5
5
2,5
0

Figure S09

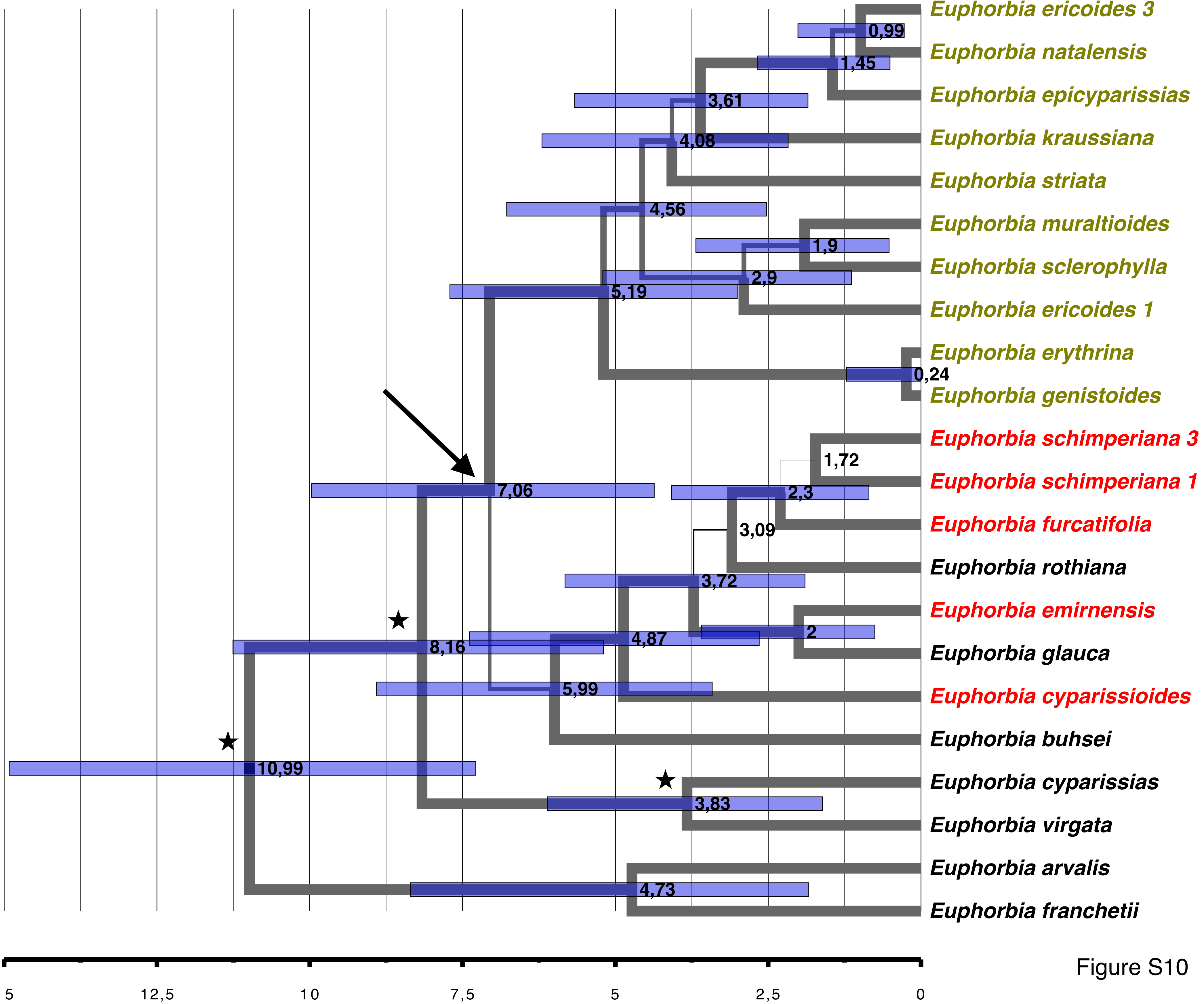

Euphorbia ericoides 3
0,99
Euphorbia natalensis
1,45
Euphorbia epicyparissias
3,61
Euphorbia kraussiana
4,08
Euphorbia striata
4,56
Euphorbia muraltioides
1,9
Euphorbia sclerophylla
2,9
5,19
Euphorbia ericoides 1
Euphorbia erythrina
0,24
Euphorbia genistoides
Euphorbia schimperiana 3
1,72
Euphorbia schimperiana 1
7,06
2,3
Euphorbia furcatifolia
3,09
Euphorbia rothiana
3,72
Euphorbia emirnensis
★
2
4,87
8,16
Euphorbia glauca
5,99
Euphorbia cyparissioides
Euphorbia buhsei
★
10,99
★
Euphorbia cyparissias
3,83
Euphorbia virgata
Euphorbia arvalis
4,73
Euphorbia franchetii
5
12,5
10
7,5
5
2,5
0

Figure S10


Euphorbia brevirama
Euphorbia crassipes 2
Euphorbia friedrichiae
Euphorbia melanohydrata
Euphorbia clavarioides
Euphorbia esculenta
Euphorbia multiceps 2
Euphorbia decepta 1
Euphorbia davyi
Euphorbia crassipes 1
Euphorbia decepta 2
Euphorbia braunsii
Euphorbia procumbens
Euphorbia flanaganii
Euphorbia hypogaea 2
Euphorbia albipollinifera
Euphorbia hypogaea 1
Euphorbia namibensis
Euphorbia namaquensis 2
Euphorbia filiflora
Euphorbia caput-medusae 2
Euphorbia fasciculata
Euphorbia schoenlandii
Euphorbia restituta
Euphorbia caput-medusae 1
Euphorbia hallii
Euphorbia namaquensis 1
Euphorbia multiceps 1
Euphorbia arida
Euphorbia celata 1
Euphorbia quadrata
Euphorbia dregeana
Euphorbia pedemontana
Euphorbia hamata 1
Euphorbia celata 2
Euphorbia hamata 2
Euphorbia gariepina
Euphorbia indurescens
Euphorbia monteiroi
Euphorbia lignosa
Euphorbia patula
Euphorbia polycephala
Euphorbia bruynsii
Euphorbia globosa
Euphorbia wilmaniae
Euphorbia pseudotuberosa
Euphorbia trichadenia
Euphorbia mammillaris 1
Euphorbia cumulata
Euphorbia meloformis 4
Euphorbia pulvinata
Euphorbia pentagona
Euphorbia obesa ssp. obesa
Euphorbia obesa ssp. symmetrica
Euphorbia ferox
Euphorbia jansenvillensis 1
Euphorbia meloformis 1
Euphorbia polygona
Euphorbia stellispina
Euphorbia jansenvillensis 2
Euphorbia pillansii
Euphorbia pseudoglobosa 2
Euphorbia clandestina
Euphorbia pubiglans
Euphorbia pseudoglobosa 1
Euphorbia mammillaris 2
Euphorbia heptagona
Euphorbia susannae
Euphorbia silenifolia 2
Euphorbia multifolia 2
Euphorbia clava
Euphorbia bubalina
Euphorbia tuberosa
Euphorbia silenifolia 1
Euphorbia loricata 2
Euphorbia loricata 1
Euphorbia multifolia 1
Euphorbia oxystegia
Euphorbia bupleurifolia
Euphorbia omariana
Euphorbia platycephala
Euphorbia grantii
Euphorbia masirahensis
Euphorbia larica
Euphorbia rubriseminalis
Euphorbia meuleniana
Euphorbia balsamifera ssp. balsamifera
Euphorbia balsamifera ssp. adenensis
Euphorbia matabelensis
Euphorbia bongensis
Euphorbia oatesii
Euphorbia cuneata
Euphorbia smithii
Euphorbia caperonioides
Euphorbia insarmentosa
Euphorbia crotonoides
Euphorbia benthamii
Euphorbia marie-cladieae
Euphorbia socotrana
Euphorbia hamaderoensis
Euphorbia scheffleri
Euphorbia sp.
Euphorbia acalyphoides
Euphorbia hadramautica
Euphorbia longituberculosa
Euphorbia antso
0,37
1,39
2,08
3,51
4,13
1,15
3,51
1,4
4,47
0,85
1,99
2,55
1,33
2,74
1,76
3,17
5,24
0,46
1,08
1,66
2,58
1,03
3,3
1,44
4,17
1,66
3,17
1,02
6,72
2
3,27
3,44
0,95
3,94
1,19
4,74
0,6
2,69
2,12
5,73
0,49
0,88
1,77
1,24
2,46
3,84
0,9
1,02
1,72
1,03
2,41
★
8,21
0,63
1,21
3,19
1,19
4,01
1,72
2,54
0,88
1,19
0,31
5,15
1,72
0,27
2,67
1,2
5,44
3,07
1,52
★
10,68
4,53
0,74
6,18
1,72
1,06
1,84
7,43
4,99
2,06
2,43
5,43
14,07
0,48
1,61
6,18
7,24
3,79
2,2
3,47
6,89
3,4
★
16,49
9,47
4,39
5,88
2,95
11,41
0,57
1,21
★
19,59
3,17
1,11
2,99
8,04
25
20
15
10
5
0

Figure S11

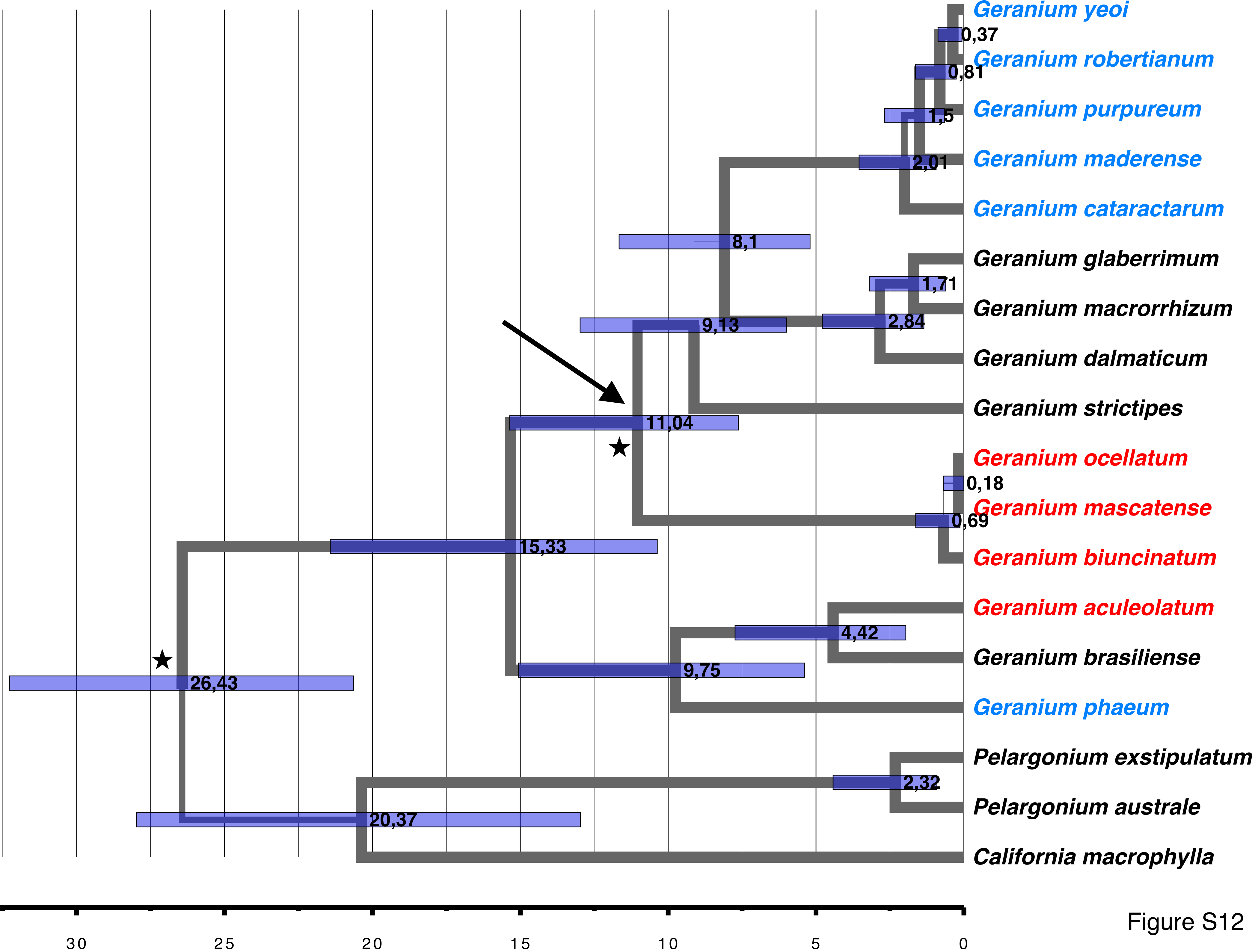

Geranium yeoi
0,37
Geranium robertianum
0,81
Geranium purpureum
1,5
Geranium maderense
2,01
Geranium cataractarum
8,1
Geranium glaberrimum
1,71
Geranium macrorrhizum
9,13
2,84
Geranium dalmaticum
Geranium strictipes
11,04
★
Geranium ocellatum
0,18
Geranium mascatense
0,69
15,33
Geranium biuncinatum
Geranium aculeolatum
4,42
Geranium brasiliense
★
26,43
9,75
Geranium phaeum
Pelargonium exstipulatum
2,32
Pelargonium australe
20,37
California macrophylla
30
25
20
15
10
5
0

Figure S12


Hypericum scioanum 1
0,13
Hypericum scioanum 2
2,98
Hypericum peplidifolium 1
0,56
3,65
Hypericum peplidifolium 2
5,23
Hypericum aethiopicum
Hypericum conjungens
3,32
Hypericum elegans
5,87
Hypericum kiboense 2
2,01
Hypericum kiboense 3
3,07
Hypericum kiboense 1
6,77
Hypericum tetrapterum
2,02
Hypericum undulatum
3,94
Hypericum maculatum 1
0,45
Hypericum maculatum 2
6,17
7,87
Hypericum perforatum 1
1,18
Hypericum perforatum 2
3,95
Hypericum perforatum 3
0,25
Hypericum perforatum 4
9,21
Hypericum aethiopicum ssp. sonderi
Hypericum olympicum 1
0,61
Hypericum olympicum 2
Hypericum linariifolium 1
1,06
Hypericum linariifolium 2
1,72
10,84
2,62
Hypericum humifusum 2
Hypericum humifusum 1
8,37
Hypericum papuanum
6,72
Hypericum perfoliatum
9,48
Hypericum scouleri
3,99
Hypericum formosum
9,86
6,93
Hypericum erectum
Hypericum richeri ssp. burseri
5,35
Hypericum montbretii
13,31
Hypericum glandulosum 1
0,19
Hypericum glandulosum 2
1,43
4,46
Hypericum reflexum
Hypericum coadunatum
1,87
6,52
Hypericum naudinianum
Hypericum caprifolium 1
1,22
Hypericum caprifolium 2
10,27
Hypericum tomentosum 1
1,85
Hypericum tomentosum 2
3,73
Hypericum somaliense
6,67
Hypericum pubescens
15,79
1,56
9,42
Hypericum psilophytum
11,4
Hypericum nummularioides
Hypericum aucheri
1,77
Hypericum triquetrifolium
5,67
Hypericum athoum
3,42
7,96
Hypericum annulatum
Hypericum montanum 1
0,39
Hypericum montanum 2
Hypericum cerastoides
18,86
Hypericum empetrifolium var. oliganthum
0,54
Hypericum xylosteifolium
0,94
6,37
Hypericum wightianum
Hypericum coris 1
0,94
9,74
Hypericum coris 2
Hypericum nummularium 1
10,88
0,83
Hypericum nummularium 2
Hypericum ericoides
Hypericum callithyrsum
5,68
14,84
Hypericum hyssopifolium
7,46
Hypericum helianthemoides
10,06
4,67
Hypericum scabrum
10,58
Hypericum confertum
12,22
Hypericum hirsutum
Hypericum linarioides
24,4
Hypericum choisianum
0,29
Hypericum patulum
1,92
Hypericum henryi
3,72
Hypericum acmosepalum
0,74
Hypericum lancasteri 2
2,16
4,4
Hypericum pseudohenryi
Hypericum curvisepalum
1,49
7,08
Hypericum lancasteri 1
2,79
Hypericum hookerianum
9,48
Hypericum synstylum
Hypericum socotranum
3,17
11,29
7,62
Hypericum balfourii
Hypericum longistylum
6,48
Hypericum ascyron
12,34
Hypericum elatoides
Hypericum balearicum 1
1,08
6,16
Hypericum balearicum 2
9,38
Hypericum calycinum
27,16
Hypericum dogonbadanicum
15,99
Hypericum bequaertii 1
0,12
Hypericum revolutum ssp. keniense
0,62
Hypericum revolutum ssp. revolutum
0,12
2,49
Hypericum bequaertii 2
Hypericum roeperianum 1
0,16
4,02
Hypericum roeperianum 2
Hypericum quartinianum 1
1,29
Hypericum quartinianum 2
Hypericum foliosum
0,9
Hypericum hircinum ssp. metroi
3,8
Hypericum grandifolium 1
0,94
5,21
Hypericum grandifolium 2
10,84
Hypericum androsaemum
17,3
Hypericum canariense 1
0,17
Hypericum canariense 2
21,02
Hypericum pamphylicum
Hypericum scopulorum
0,4
Hypericum tortuosum
33,75
Hypericum brachyphyllum
1,76
Hypericum fasciculatum 1
3
Hypericum galioides
4,83
Hypericum tenuifolium
2,87
Hypericum lissophloeus
4,06
6,89
Hypericum fasciculatum 2
Hypericum prolificum 1
1,06
11,62
Hypericum prolificum 2
3,49
Hypericum densiflorum
Hypericum ellipticum
13,6
Hypericum hypericoides 1
0,79
Hypericum hypericoides 2
7,41
Hypericum cistifolium 1
0,93
9,35
Hypericum cistifolium 2
Hypericum crux-andreae
5,88
Hypericum tetrapetalum
Hypericum laricifolium 1
0,45
Hypericum laricifolium 2
1,65
★
34,93
21,93
2,47
Hypericum mexicanum
Hypericum strictum
3,86
Hypericum brevistylum 1
0,24
7,27
Hypericum brevistylum 2
2,31
Hypericum quitense
8,77
Hypericum gentianoides
Hypericum drummondii
29,3
Hypericum mutilum ssp. mutilum
0,92
13,69
Hypericum mutilum ssp. boreale
3,29
Hypericum boreale
4,75
Hypericum lalandii
6,75
Hypericum canadense
2,42
Hypericum majus
49,9
Triadenum fraseri
Hypericum elodes 1
0,16
Hypericum elodes 2
★
53,88
27,72
Hypericum aegypticum ssp. webbii
1,33
Hypericum aegypticum ssp. aegypticum
Harungana madagascarensis
11,97
Vismia glaziovii
24,66
Psorospermum senegalense
Eliea articulata
60
50
40
30
20
10
0

Figure S13

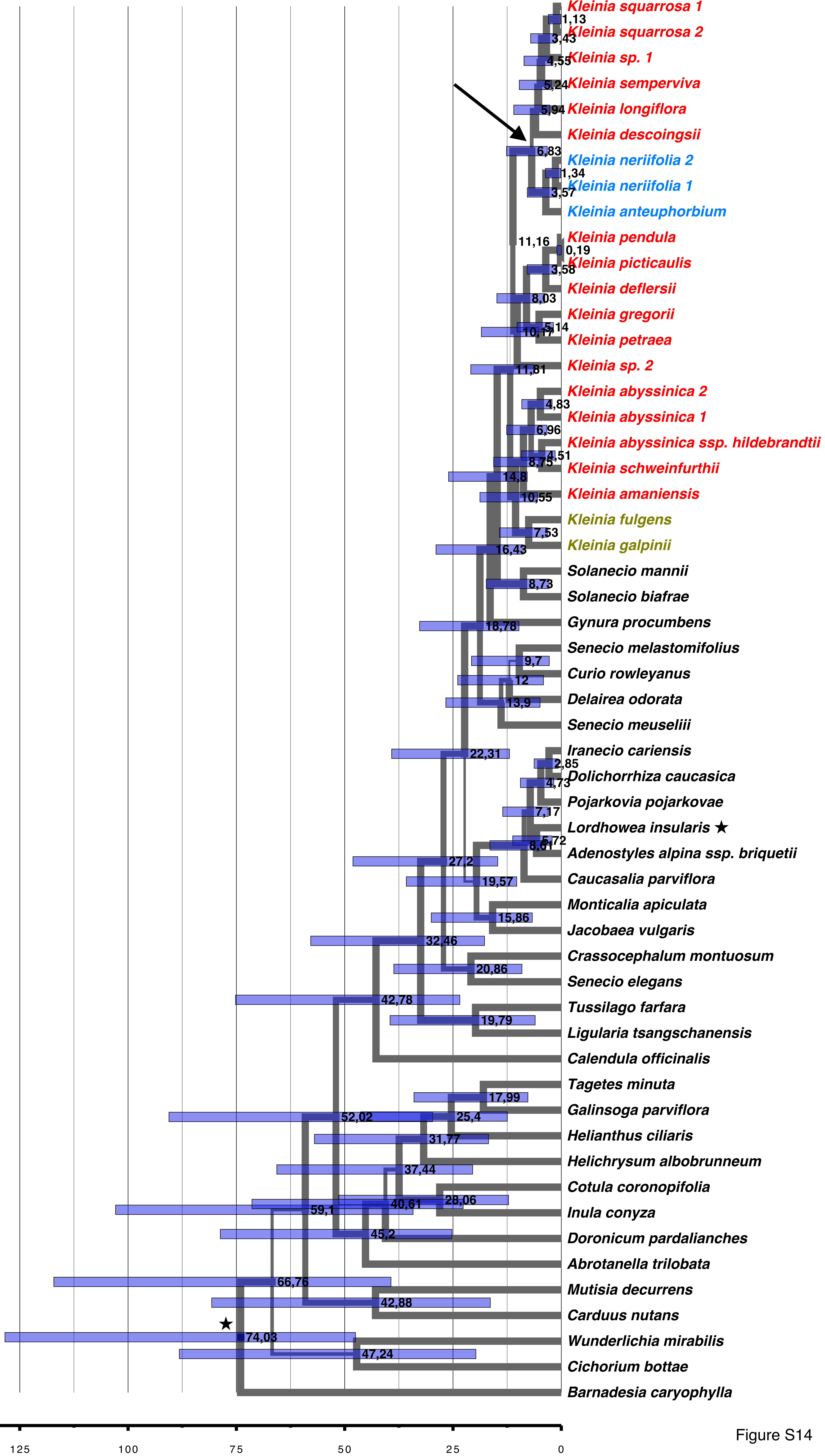

Kleinia squarrosa 1
1,13
Kleinia squarrosa 2
3,43
Kleinia sp. 1
4,55
Kleinia semperviva
5,24
Kleinia longiflora
5,94
Kleinia descoingsii
6,83
Kleinia neriifolia 2
1,34
Kleinia neriifolia 1
3,57
Kleinia anteuphorbium
11,16
Kleinia pendula
0,19
Kleinia picticaulis
3,58
Kleinia deflersii
8,03
Kleinia gregorii
5,14
10,17
Kleinia petraea
Kleinia sp. 2
11,81
Kleinia abyssinica 2
4,83
Kleinia abyssinica 1
6,96
Kleinia abyssinica ssp. hildebrandtii
4,51
8,75
Kleinia schweinfurthii
14,8
10,55
Kleinia amaniensis
Kleinia fulgens
7,53
Kleinia galpinii
16,43
Solanecio mannii
8,73
Solanecio biafrae
Gynura procumbens
18,78
Senecio melastomifolius
9,7
Curio rowleyanus
12
Delairea odorata
13,9
Senecio meuseliii
22,31
Iranecio cariensis
2,85
Dolichorrhiza caucasica
4,73
Pojarkovia pojarkovae
7,17
Lordhowea insularis ★
5,72
8,61
Adenostyles alpina ssp. briquetii
27,2
Caucasalia parviflora
19,57
Monticalia apiculata
15,86
Jacobaea vulgaris
32,46
Crassocephalum montuosum
20,86
Senecio elegans
42,78
Tussilago farfara
19,79
Ligularia tsangschanensis
Calendula officinalis
Tagetes minuta
17,99
Galinsoga parviflora
52,02
25,4
Helianthus ciliaris
31,77
Helichrysum albobrunneum
37,44
Cotula coronopifolia
28,06
40,61
59,1
Inula conyza
45,2
Doronicum pardalianches
Abrotanella trilobata
66,76
Mutisia decurrens
42,88
Carduus nutans
★
74,03
Wunderlichia mirabilis
47,24
Cichorium bottae
Barnadesia caryophylla
125
100
75
50
25
0

Figure S14

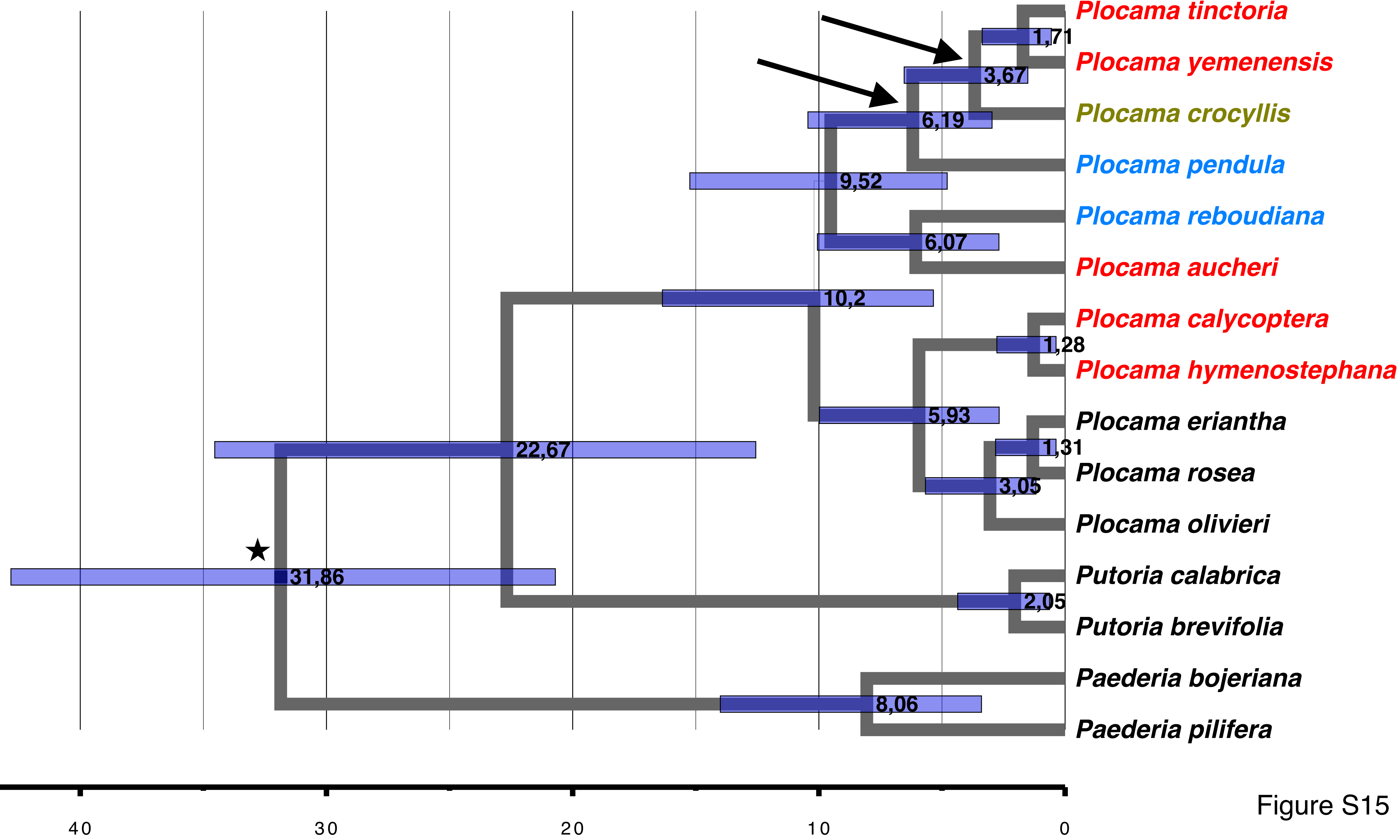

Plocama tinctoria
Plocama yemenensis
Plocama crocyllis
Plocama pendula
Plocama reboudiana
Plocama aucheri
Plocama calycoptera
Plocama hymenostephana
Plocama eriantha
Plocama rosea
Plocama olivieri
Putoria calabrica
Putoria brevifolia
Paederia bojeriana
Paederia pilifera
1,71
3,67
6,19
9,52
6,07
10,2
1,28
5,93
22,67
1,31
3,05
31,86
2,05
8,06
★
40
30
20
10
0

Figure S15

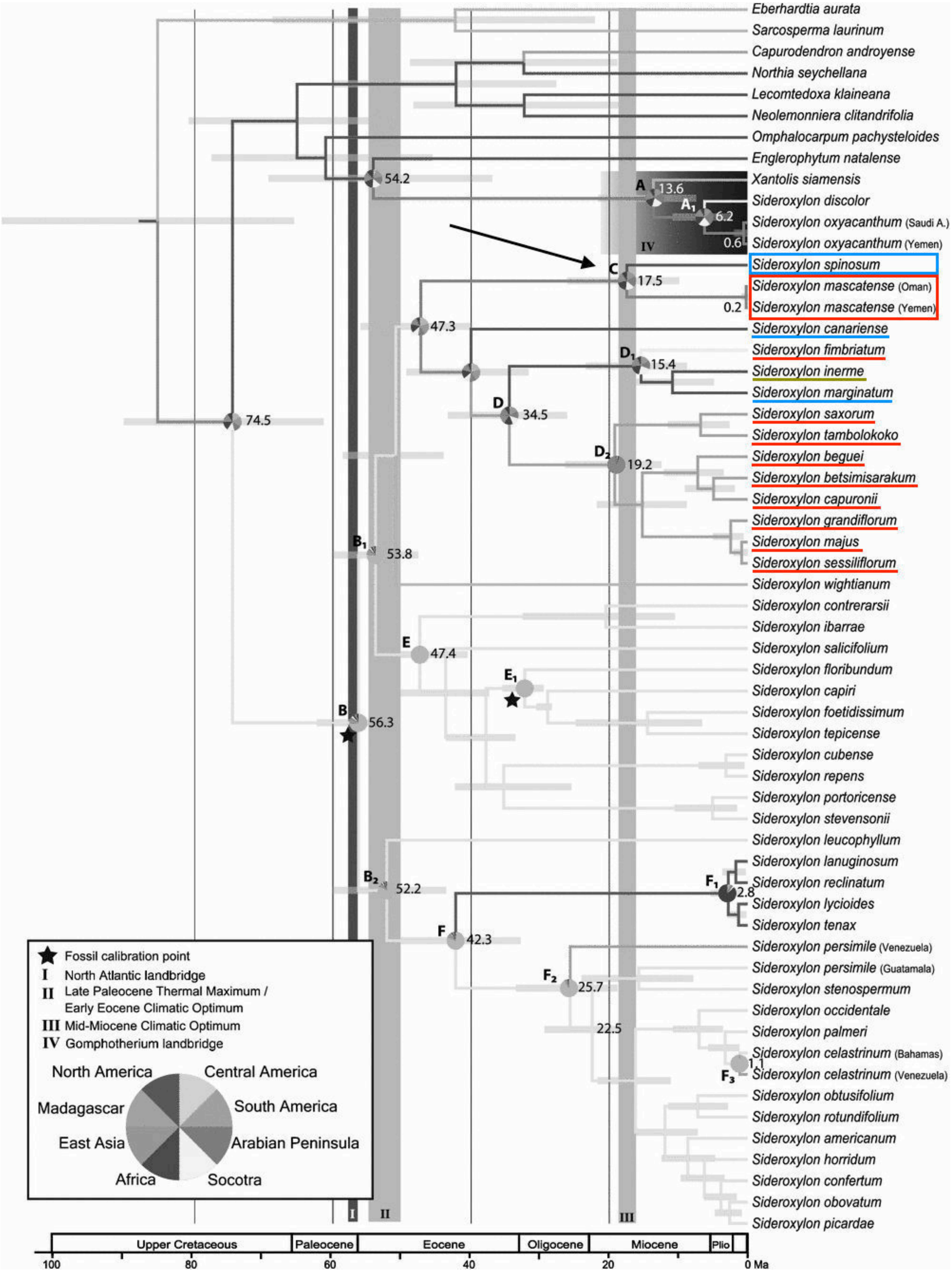

Eberhardtia aurata
Sarcosperma laurinum
Capurodendron androyense
Northia seychellana
Lecomtedoxa klaineana
Neolemonniera clitandrifolia
Omphalocarpum pachysteloides
Englerophytum natalense
Xantolis siamensis
Sideroxylon discolor
Sideroxylon oxyacanthum (Saudi A.)
Sideroxylon oxyacanthum (Yemen)
Sideroxylon spinosum
Sideroxylon mascatense (Oman)
Sideroxylon mascatense (Yemen)
Sideroxylon canariense
Sideroxylon fimbriatum
Sideroxylon inerme
Sideroxylon marginatum
Sideroxylon saxorum
Sideroxylon tambolokoko
Sideroxylon beguei
Sideroxylon betsimisarakum
Sideroxylon capuronii
Sideroxylon grandiflorum
Sideroxylon majus
Sideroxylon sessiliflorum
Sideroxylon wightianum
Sideroxylon contrerarsii
Sideroxylon ibarrae
Sideroxylon salicifolium
Sideroxylon floribundum
Sideroxylon capiri
Sideroxylon foetidissimum
Sideroxylon tepicense
Sideroxylon cubense
Sideroxylon repens
Sideroxylon portoricense
Sideroxylon stevensonii
Sideroxylon leucophyllum
Sideroxylon lanuginosum
Sideroxylon reclinatum
Sideroxylon lycioides
Sideroxylon tenax
Sideroxylon persimile (Venezuela)
Sideroxylon persimile (Guatamala)
Sideroxylon stenospermum
Sideroxylon occidentale
Sideroxylon palmeri
Sideroxylon celastrinum (Bahamas)
Sideroxylon celastrinum (Venezuela)
Sideroxylon obtusifolium
Sideroxylon rotundifolium
Sideroxylon americanum
Sideroxylon horridum
Sideroxylon confertum
Sideroxylon obovatum
Sideroxylon picardae
54.2
A
13.6
A₁
6.2
0.6
IV
C
17.5
0.2
47.3
D₁
15.4
D
34.5
74.5
D₂
19.2
B₁
53.8
E
47.4
E₁
B
56.3
B₂
52.2
F₁
2.8
F
42.3
F₂
25.7
22.5
1.1
F₃
Fossil calibration point
I North Atlantic landbridge
II Late Paleocene Thermal Maximum / Early Eocene Climatic Optimum
III Mid-Miocene Climatic Optimum
IV Gomphotherium landbridge
North America
Central America
Madagascar
South America
East Asia
Arabian Peninsula
Africa
Socotra
I
II
III
Upper Cretaceous
Paleocene
Eocene
Oligocene
Miocene
Plio
100
80
60
40
20
0 Ma

Figure S16

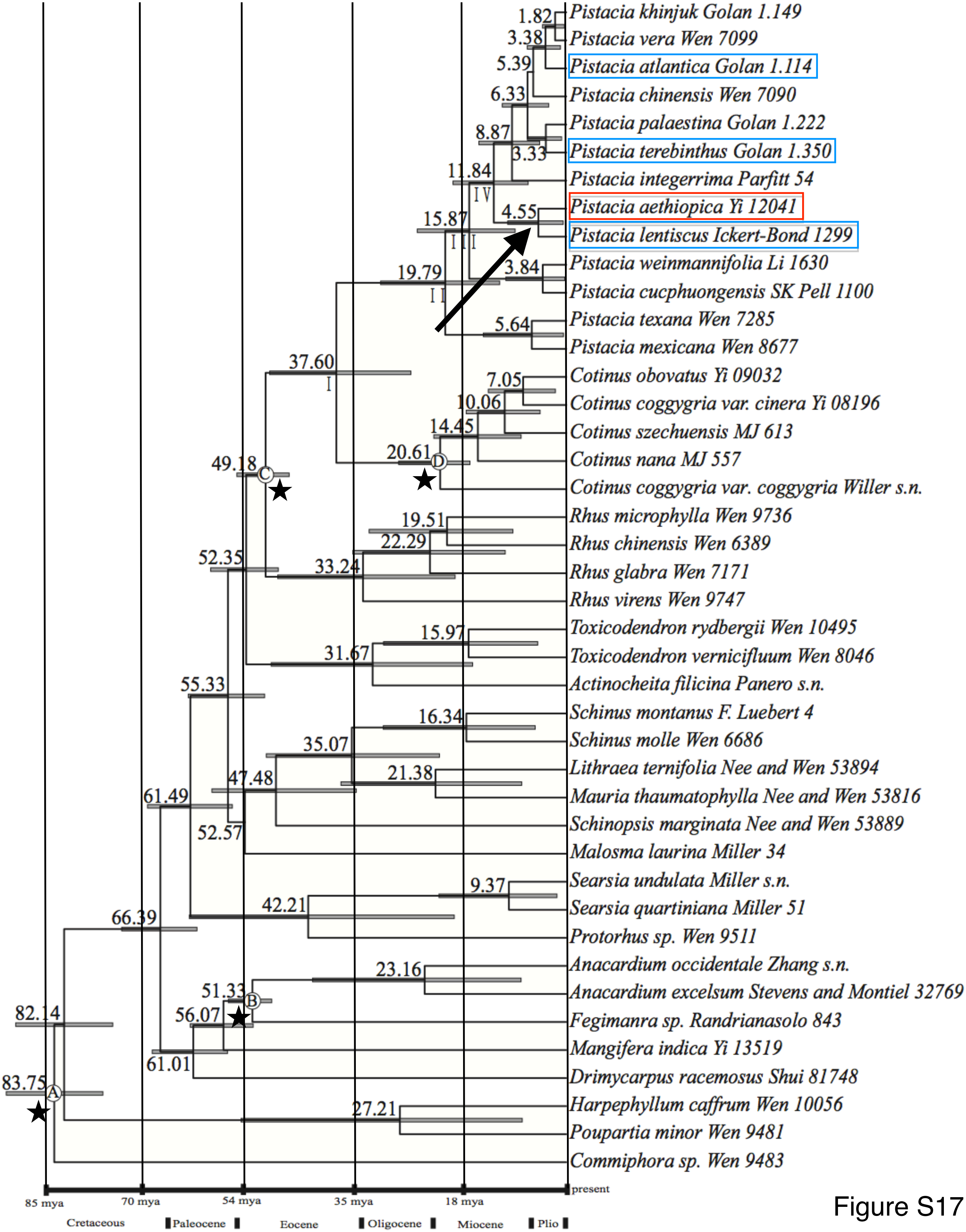

1.82
3.38
5.39
6.33
8.87
3.33
11.84
IV
4.55
15.87
III
3.84
19.79
II
5.64
37.60
I
7.05
10.06
14.45
20.61
D
49.18
C
19.51
22.29
52.35
33.24
15.97
31.67
55.33
16.34
35.07
47.48
21.38
61.49
52.57
9.37
42.21
66.39
23.16
51.33
B
56.07
82.14
61.01
83.75
A
27.21
Pistacia khinjuk Golan 1.149
Pistacia vera Wen 7099
Pistacia atlantica Golan 1.114
Pistacia chinensis Wen 7090
Pistacia palaestina Golan 1.222
Pistacia terebinthus Golan 1.350
Pistacia integerrima Parfitt 54
Pistacia aethiopica Yi 12041
Pistacia lentiscus Ickert-Bond 1299
Pistacia weinmannifolia Li 1630
Pistacia cucphuongensis SK Pell 1100
Pistacia texana Wen 7285
Pistacia mexicana Wen 8677
Cotinus obovatus Yi 09032
Cotinus coggygria var. cinera Yi 08196
Cotinus szechuensis MJ 613
Cotinus nana MJ 557
Cotinus coggygria var. coggygria Willer s.n.
Rhus microphylla Wen 9736
Rhus chinensis Wen 6389
Rhus glabra Wen 7171
Rhus virens Wen 9747
Toxicodendron rydbergii Wen 10495
Toxicodendron vernicifluum Wen 8046
Actinocheita filicina Panero s.n.
Schinus montanus F. Luebert 4
Schinus molle Wen 6686
Lithraea ternifolia Nee and Wen 53894
Mauria thaumatophylla Nee and Wen 53816
Schinopsis marginata Nee and Wen 53889
Malosma laurina Miller 34
Searsia undulata Miller s.n.
Searsia quartiniana Miller 51
Protorhus sp. Wen 9511
Anacardium occidentale Zhang s.n.
Anacardium excelsum Stevens and Montiel 32769
Fegimanra sp. Randrianasolo 843
Mangifera indica Yi 13519
Drimycarpus racemosus Shui 81748
Harpephyllum caffrum Wen 10056
Poupartia minor Wen 9481
Commiphora sp. Wen 9483
85 mya
70 mya
54 mya
35 mya
18 mya
present
Cretaceous
Paleocene
Eocene
Oligocene
Miocene
Plio

Figure S17

SUPPLEMENTARY DATA
CHAPTER 6

Lineage-specific climatic niche drives the tempo of vicariance in the Rand Flora

Mario Mairal, Isabel Sanmartín, Loïc Pellissier

Journal of Biogeography

SUPPORTING INFORMATION

Authors
Mario Mairal*[1], Isabel Sanmartín[1]† and Loïc Pellissier[2,3]†

Article title

Lineage-specific climatic niche drives the tempo of vicariance in the Rand flora

Appendix S1. Supplemental Figures

Figure S1.1. Phylograms of all the clades used in this study indicating the phylogenetic reconstructions, dating uncertainty and the disjunct node of interest (inside the green box). Adapted from Pokorny *et al.*, 2015. The figure of *Sideroxylon* is taken from Chen *et al.*, 2014.

Chen, C., Qi, Z.C., Xu, X.H., Comes, H.P, Koch, M.A., Jin, X.J., Fu C.X., Qiu Y-X (2014) Understanding the formation of Mediterranean-African-Asian disjunctions: evidence for Miocene climate-driven vicariance and recent long-distance dispersal in the Tertiary relict *Smilax aspera* (Smilacaceae). *New Phytologist*, **204**, 243-255.

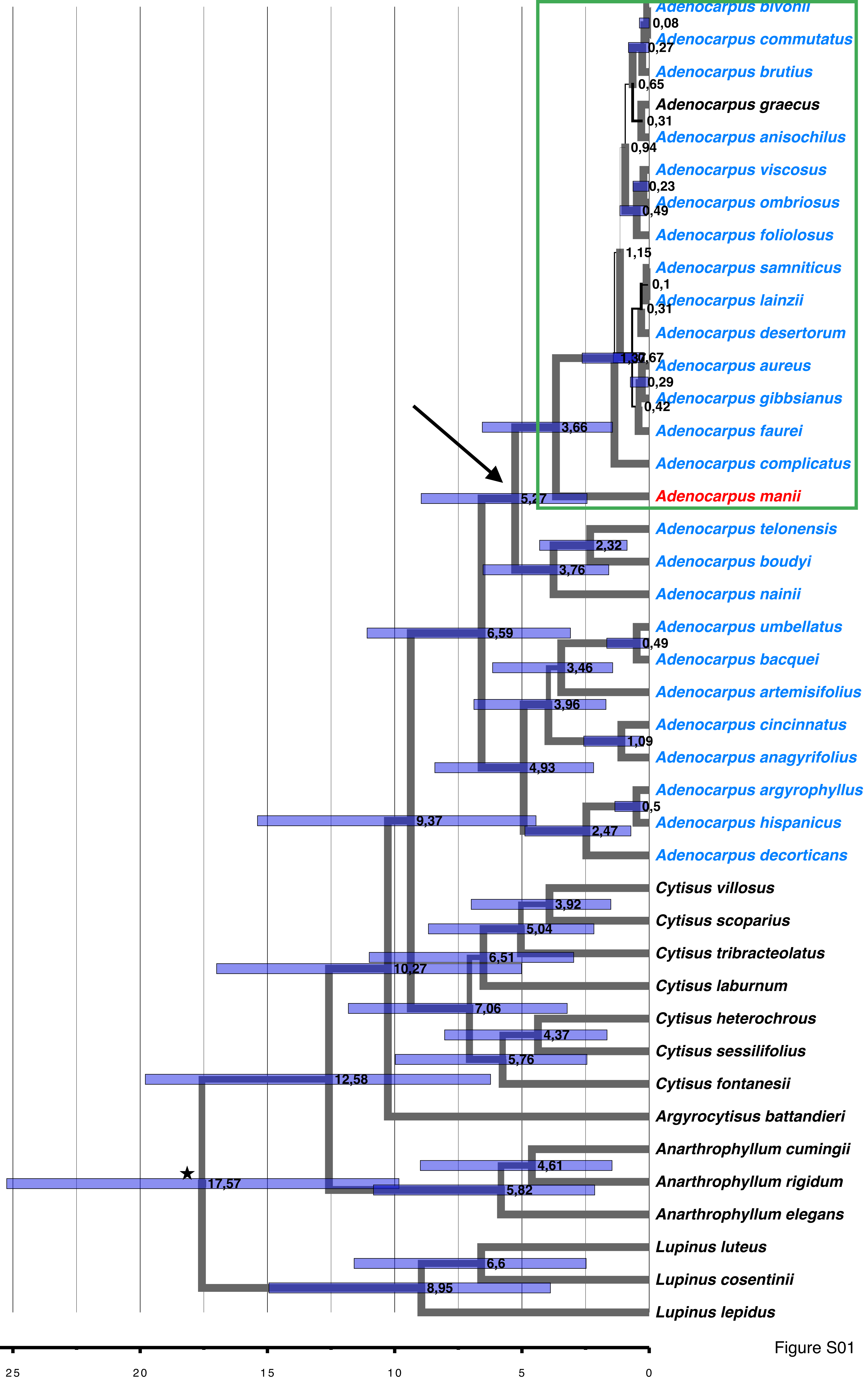

Adenocarpus bivonii
0,08
Adenocarpus commutatus
0,27
Adenocarpus brutius
0,65
Adenocarpus graecus
0,31
Adenocarpus anisochilus
0,94
Adenocarpus viscosus
0,23
Adenocarpus ombriosus
0,49
Adenocarpus foliolosus
1,15
Adenocarpus samniticus
0,1
Adenocarpus lainzii
0,31
Adenocarpus desertorum
1,37
0,67
Adenocarpus aureus
0,29
Adenocarpus gibbsianus
0,42
Adenocarpus faurei
3,66
Adenocarpus complicatus
5,27
Adenocarpus manii
Adenocarpus telonensis
2,32
Adenocarpus boudyi
3,76
Adenocarpus nainii
6,59
Adenocarpus umbellatus
0,49
Adenocarpus bacquei
3,46
Adenocarpus artemisifolius
3,96
Adenocarpus cincinnatus
1,09
Adenocarpus anagyrifolius
4,93
Adenocarpus argyrophyllus
9,37
0,5
Adenocarpus hispanicus
2,47
Adenocarpus decorticans
Cytisus villosus
3,92
Cytisus scoparius
5,04
6,51
Cytisus tribracteolatus
10,27
Cytisus laburnum
7,06
Cytisus heterochrous
4,37
Cytisus sessilifolius
5,76
12,58
Cytisus fontanesii
Argyrocytisus battandieri
Anarthrophyllum cumingii
4,61
★
17,57
Anarthrophyllum rigidum
5,82
Anarthrophyllum elegans
Lupinus luteus
6,6
Lupinus cosentinii
8,95
Lupinus lepidus
25
20
15
10
5
0

Figure S01

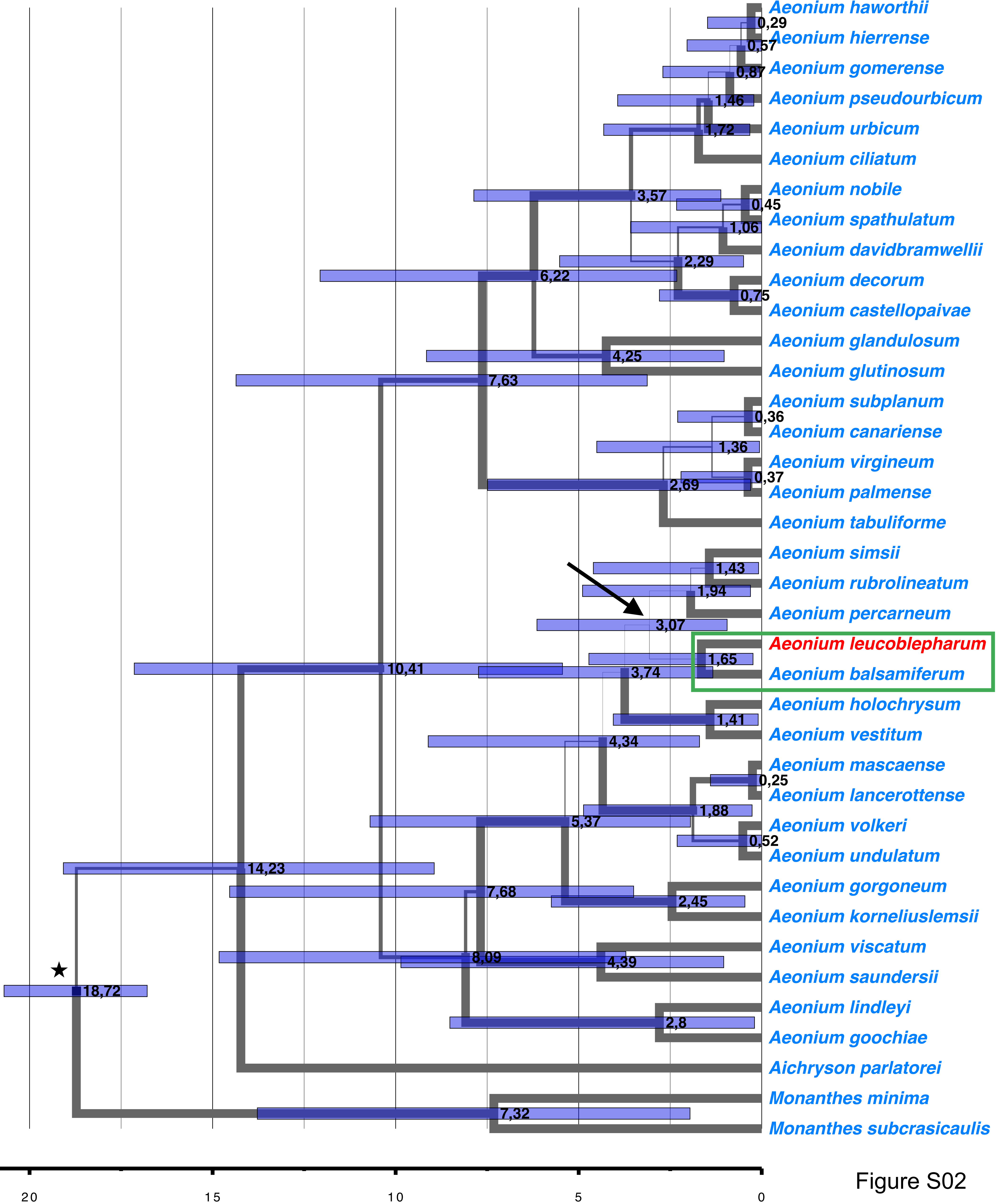
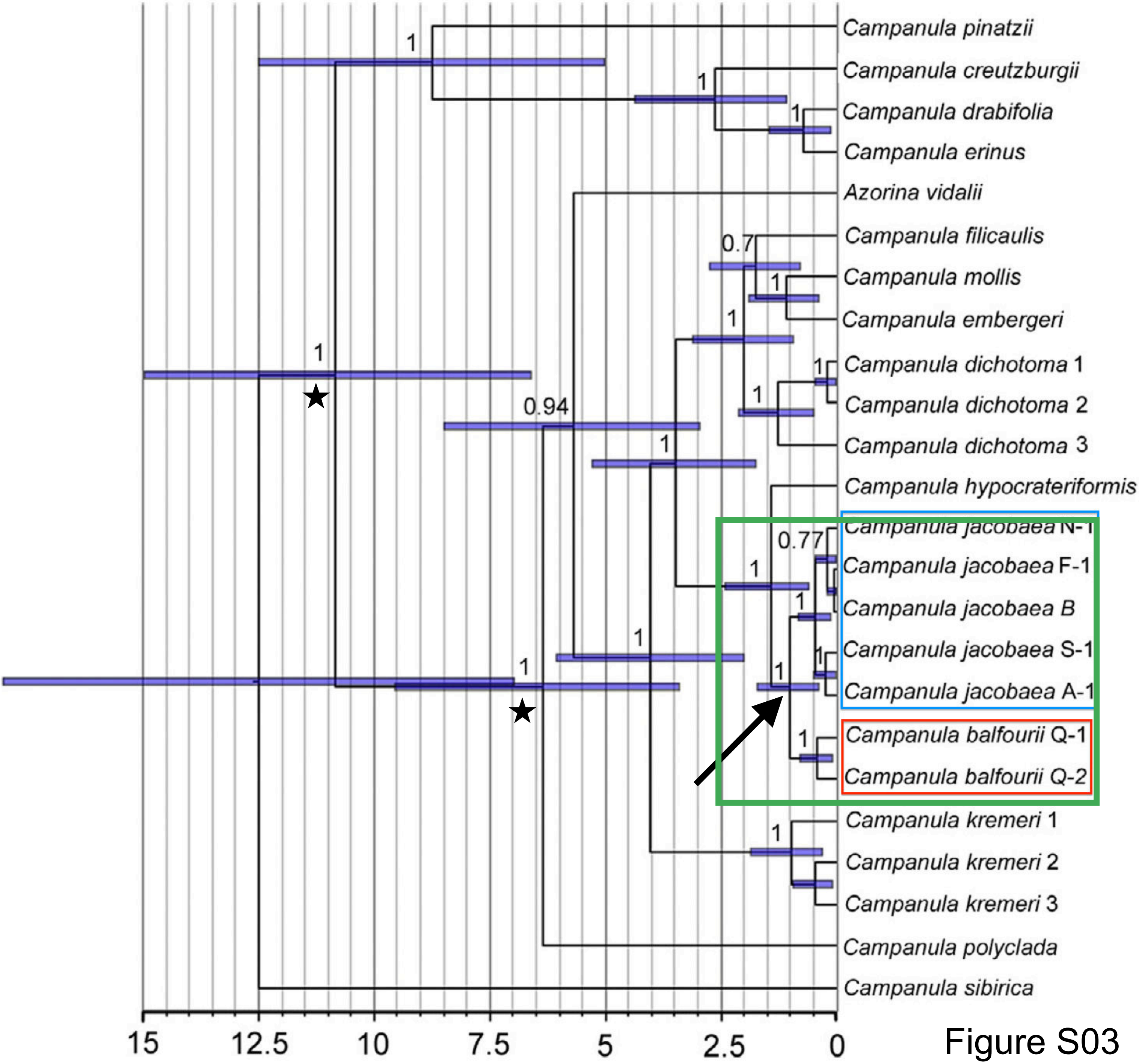

Aeonium haworthii
0,29
Aeonium hierrense
0,57
Aeonium gomerense
0,87
Aeonium pseudourbicum
1,46
Aeonium urbicum
1,72
Aeonium ciliatum
3,57
Aeonium nobile
0,45
Aeonium spathulatum
1,06
Aeonium davidbramwellii
2,29
6,22
Aeonium decorum
0,75
Aeonium castellopaivae
Aeonium glandulosum
4,25
7,63
Aeonium glutinosum
Aeonium subplanum
0,36
Aeonium canariense
1,36
Aeonium virgineum
0,37
2,69
Aeonium palmense
Aeonium tabuliforme
Aeonium simsii
1,43
1,94
Aeonium rubrolineatum
Aeonium percarneum
3,07
Aeonium leucoblepharum
1,65
10,41
3,74
Aeonium balsamiferum
Aeonium holochrysum
1,41
Aeonium vestitum
4,34
Aeonium mascaense
0,25
Aeonium lancerottense
1,88
5,37
Aeonium volkeri
0,52
Aeonium undulatum
14,23
Aeonium gorgoneum
7,68
2,45
Aeonium korneliuslemsii
Aeonium viscatum
8,09
4,39
★
18,72
Aeonium saundersii
Aeonium lindleyi
2,8
Aeonium goochiae
Aichryson parlatorei
Monanthes minima
7,32
Monanthes subcrasicaulis
20
15
10
5
0

Figure S02



Campanula pinatzii
Campanula creutzburgii
Campanula drabifolia
Campanula erinus
Azorina vidalii
Campanula filicaulis
Campanula mollis
Campanula embergeri
Campanula dichotoma 1
Campanula dichotoma 2
Campanula dichotoma 3
Campanula hypocrateriformis
Campanula jacobaea N-1
Campanula jacobaea F-1
Campanula jacobaea B
Campanula jacobaea S-1
Campanula jacobaea A-1
Campanula balfourii Q-1
Campanula balfourii Q-2
Campanula kremeri 1
Campanula kremeri 2
Campanula kremeri 3
Campanula polyclada
Campanula sibirica
1
1
1
0.7
1
1
1
1
1
0.94
1
0.77
1
1
1
1
1
1
1
15
12.5
10
7.5
5
2.5
0

Figure S03

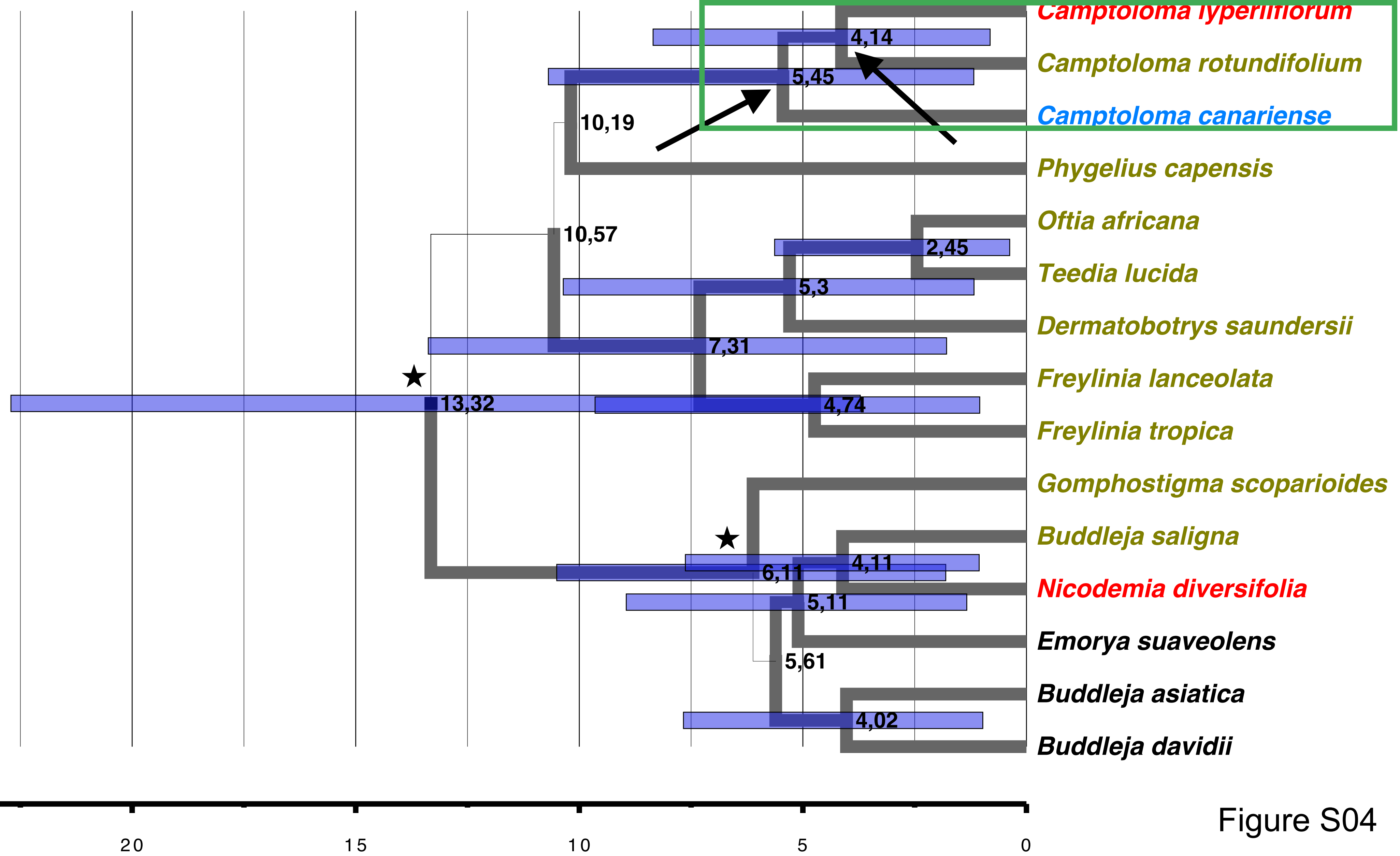

Camptoloma lyperiiflorum
4,14
Camptoloma rotundifolium
5,45
Camptoloma canariense
10,19
Phygelius capensis
Oftia africana
10,57
2,45
Teedia lucida
5,3
Dermatobotrys saundersii
7,31
Freylinia lanceolata
13,32
4,74
Freylinia tropica
Gomphostigma scoparioides
Buddleja saligna
4,11
6,11
Nicodemia diversifolia
5,11
Emorya suaveolens
5,61
Buddleja asiatica
4,02
Buddleja davidii
20
15
10
5
0

Figure S04

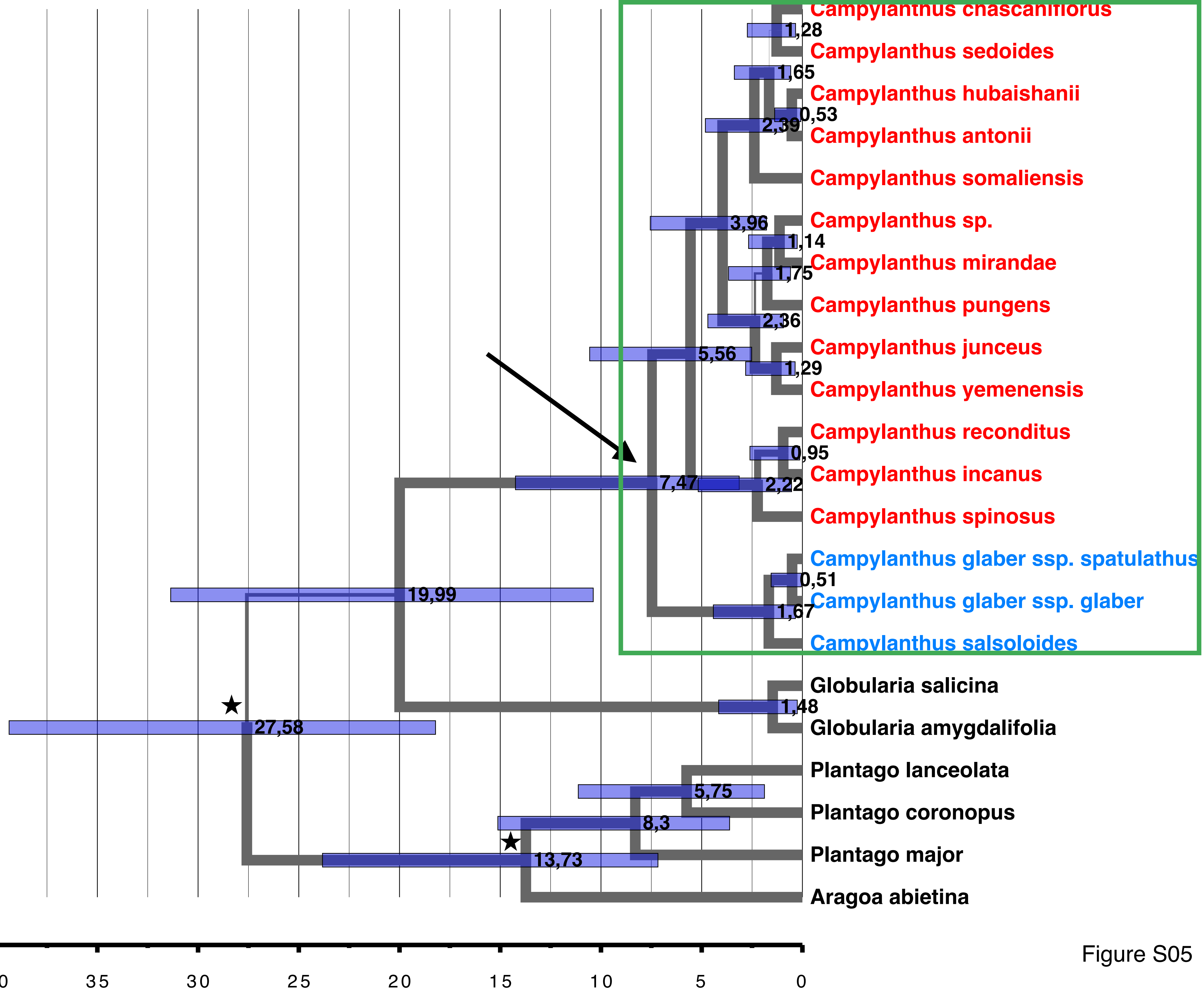

Campylanthus chascaniflorus
1,28
Campylanthus sedoides
1,65
Campylanthus hubaishanii
0,53
2,39
Campylanthus antonii
Campylanthus somaliensis
3,96
Campylanthus sp.
1,14
1,75
Campylanthus mirandae
Campylanthus pungens
2,36
Campylanthus junceus
5,56
1,29
Campylanthus yemenensis
Campylanthus reconditus
0,95
Campylanthus incanus
7,47
2,22
Campylanthus spinosus
Campylanthus glaber ssp. spatulathus
0,51
19,99
Campylanthus glaber ssp. glaber
1,67
Campylanthus salsoloides
Globularia salicina
★
1,48
27,58
Globularia amygdalifolia
Plantago lanceolata
5,75
Plantago coronopus
8,3
★
13,73
Plantago major
Aragoa abietina
0
35
30
25
20
15
10
5
0

Figure S05

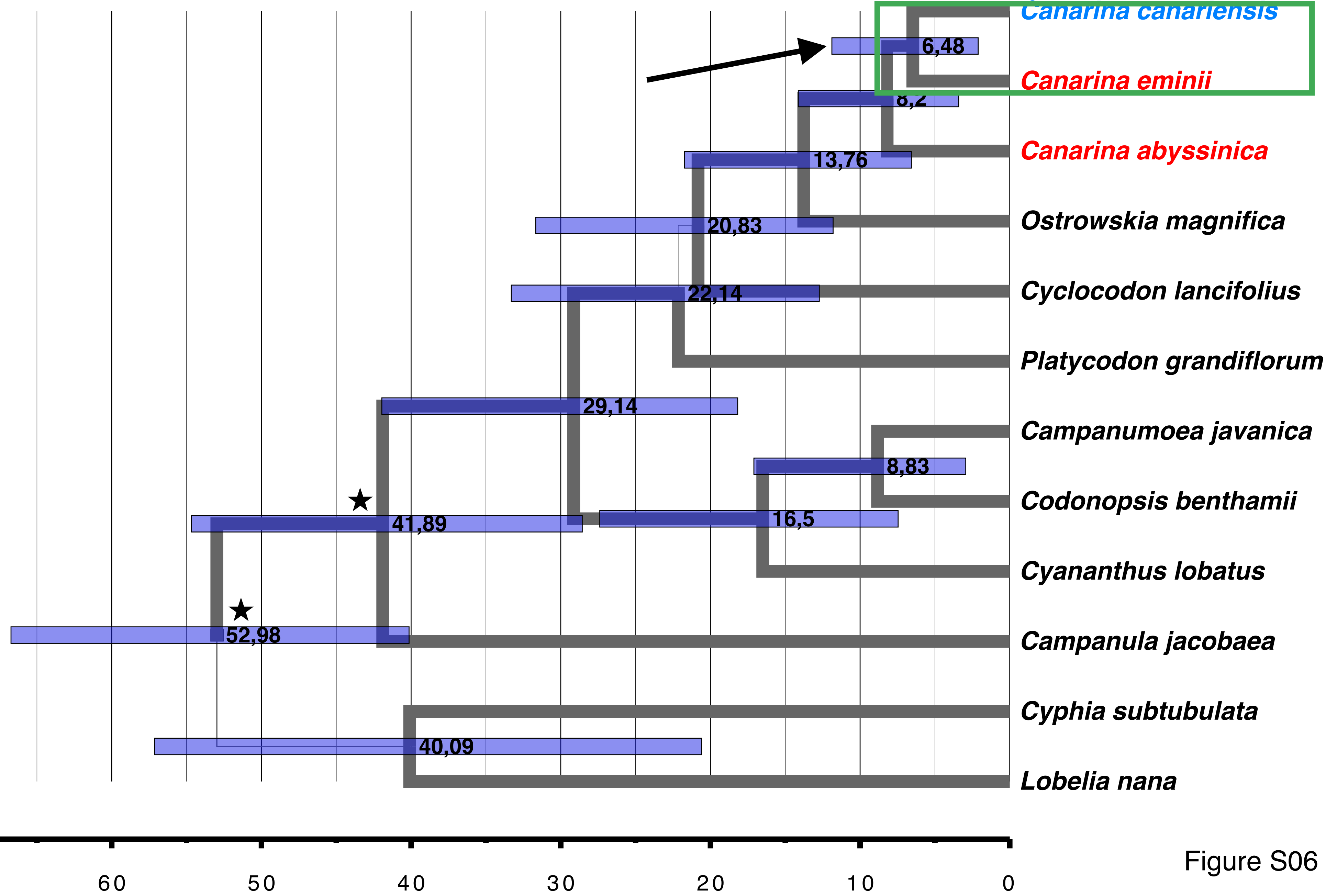

Canarina canariensis
6,48
Canarina eminii
8,2
Canarina abyssinica
13,76
Ostrowskia magnifica
20,83
Cyclocodon lancifolius
22,14
Platycodon grandiflorum
29,14
Campanumoea javanica
8,83
Codonopsis benthamii
★
41,89
16,5
Cyananthus lobatus
★
52,98
Campanula jacobaea
Cyphia subtubulata
40,09
Lobelia nana
60
50
40
30
20
10
0

Figure S06

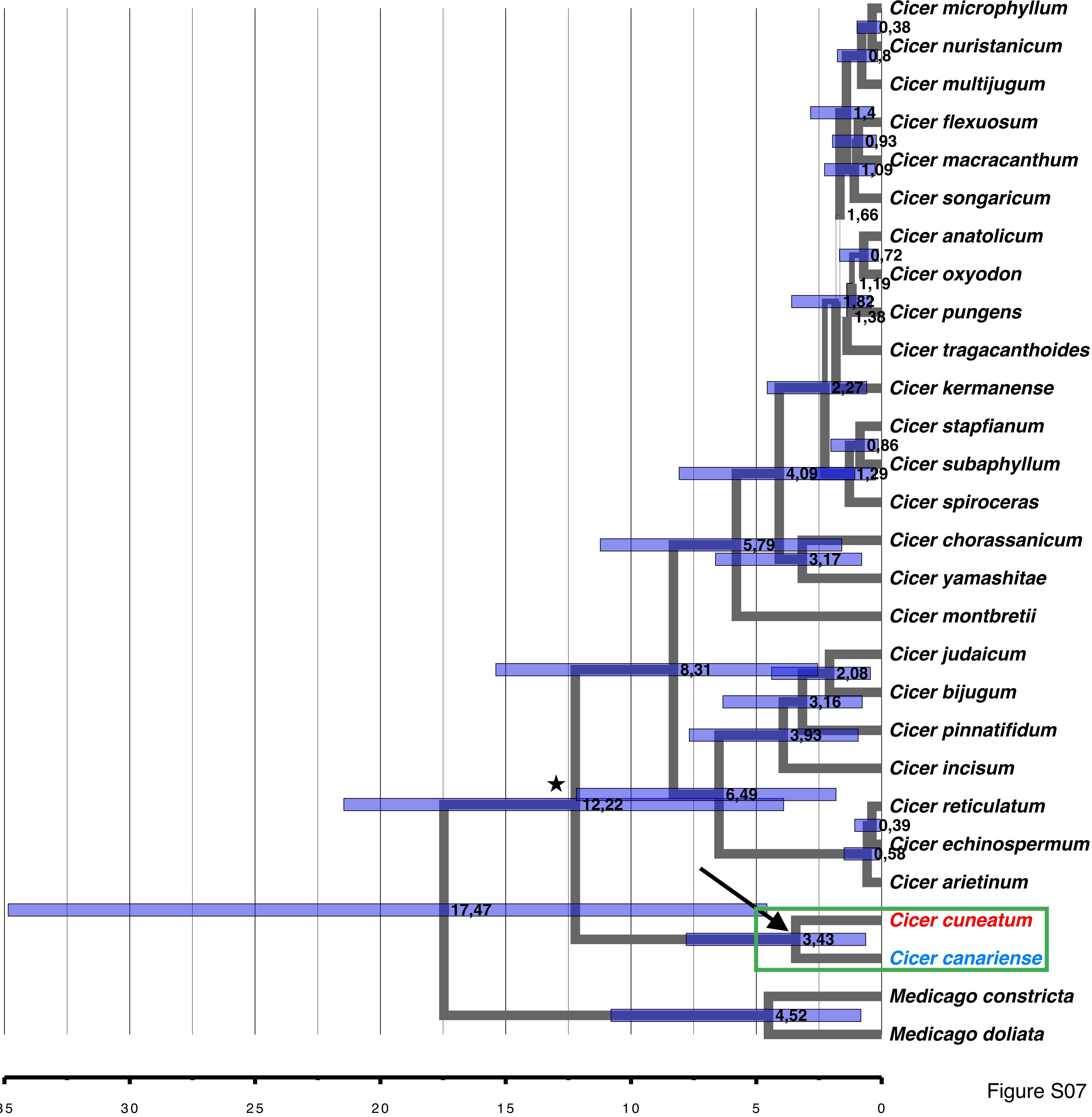

Cicer microphyllum
0,38
Cicer nuristanicum
0,8
Cicer multijugum
1,4
Cicer flexuosum
0,93
Cicer macracanthum
1,09
Cicer songaricum
1,66
Cicer anatolicum
0,72
Cicer oxyodon
1,19
1,82
Cicer pungens
1,38
Cicer tragacanthoides
2,27
Cicer kermanense
Cicer stapfianum
0,86
Cicer subaphyllum
4,09
1,29
Cicer spiroceras
Cicer chorassanicum
5,79
3,17
Cicer yamashitae
Cicer montbretii
Cicer judaicum
8,31
2,08
Cicer bijugum
3,16
Cicer pinnatifidum
3,93
Cicer incisum
★
6,49
12,22
Cicer reticulatum
0,39
Cicer echinospermum
0,58
Cicer arietinum
17,47
Cicer cuneatum
3,43
Cicer canariense
Medicago constricta
4,52
Medicago doliata
35
30
25
20
15
10
5
0

Figure S07

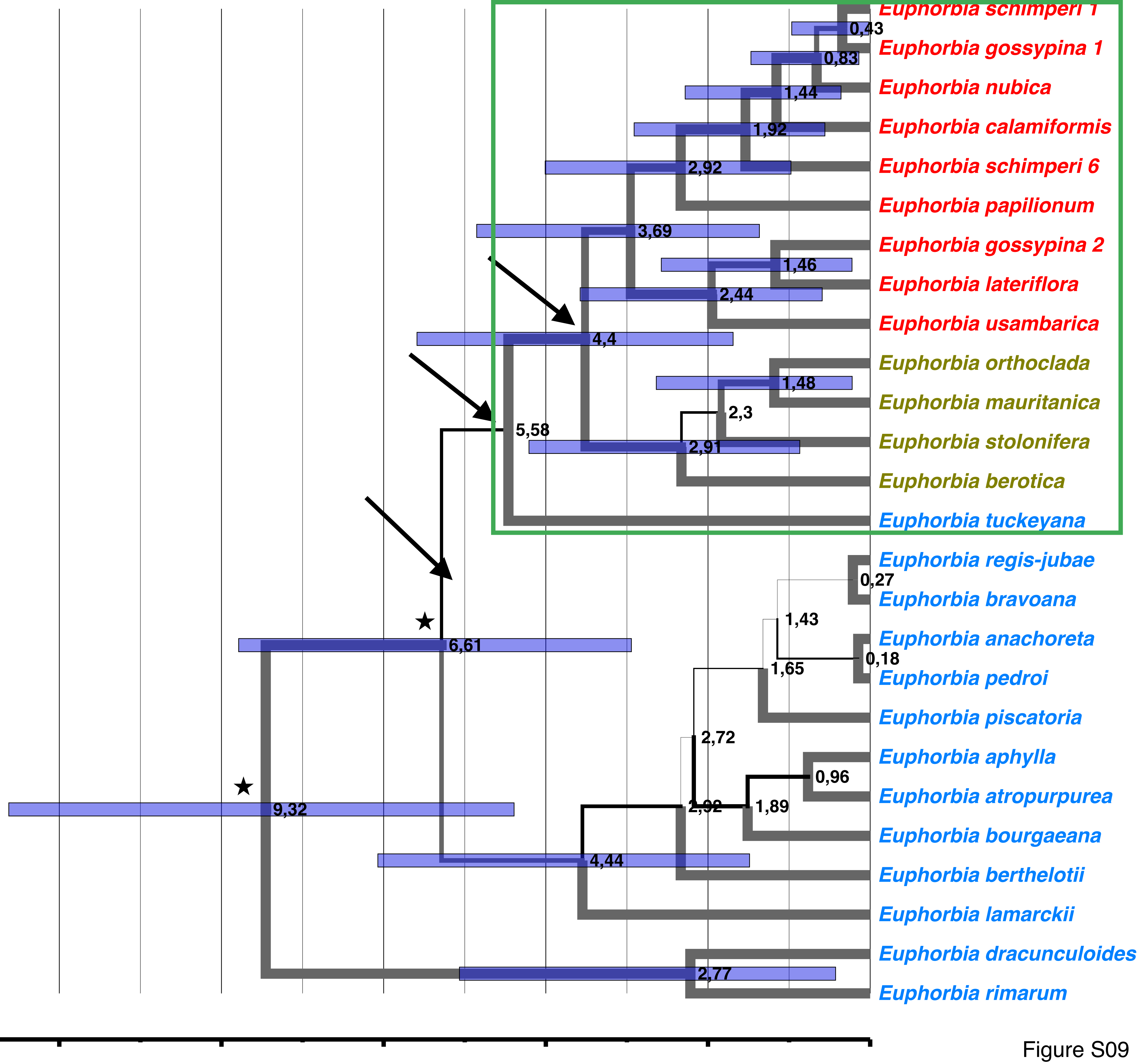

Euphorbia schimperi 1
Euphorbia gossypina 1
Euphorbia nubica
Euphorbia calamiformis
Euphorbia schimperi 6
Euphorbia papilionum
Euphorbia gossypina 2
Euphorbia lateriflora
Euphorbia usambarica
Euphorbia orthoclada
Euphorbia mauritanica
Euphorbia stolonifera
Euphorbia berotica
Euphorbia tuckeyana
Euphorbia regis-jubae
Euphorbia bravoana
Euphorbia anachoreta
Euphorbia pedroi
Euphorbia piscatoria
Euphorbia aphylla
Euphorbia atropurpurea
Euphorbia bourgaeana
Euphorbia berthelotii
Euphorbia lamarckii
Euphorbia dracunculoides
Euphorbia rimarum
0,43
0,83
1,44
1,92
2,92
3,69
1,46
2,44
4,4
1,48
2,3
5,58
2,91
0,27
1,43
0,18
1,65
6,61
2,72
0,96
9,32
2,92
1,89
4,44
2,77
★
★
12,5
10
7,5
5
2,5
0

Figure S09


Euphorbia brevirama
Euphorbia crassipes 2
Euphorbia friedrichiae
Euphorbia melanohydrata
Euphorbia clavarioides
Euphorbia esculenta
Euphorbia multiceps 2
Euphorbia decepta 1
Euphorbia davyi
Euphorbia crassipes 1
Euphorbia decepta 2
Euphorbia braunsii
Euphorbia procumbens
Euphorbia flanaganii
Euphorbia hypogaea 2
Euphorbia albipollinifera
Euphorbia hypogaea 1
Euphorbia namibensis
Euphorbia namaquensis 2
Euphorbia filiflora
Euphorbia caput-medusae 2
Euphorbia fasciculata
Euphorbia schoenlandii
Euphorbia restituta
Euphorbia caput-medusae 1
Euphorbia hallii
Euphorbia namaquensis 1
Euphorbia multiceps 1
Euphorbia arida
Euphorbia celata 1
Euphorbia quadrata
Euphorbia dregeana
Euphorbia pedemontana
Euphorbia hamata 1
Euphorbia celata 2
Euphorbia hamata 2
Euphorbia gariepina
Euphorbia indurescens
Euphorbia monteiroi
Euphorbia lignosa
Euphorbia patula
Euphorbia polycephala
Euphorbia bruynsii
Euphorbia globosa
Euphorbia wilmaniae
Euphorbia pseudotuberosa
Euphorbia trichadenia
Euphorbia mammillaris 1
Euphorbia cumulata
Euphorbia meloformis 4
Euphorbia pulvinata
Euphorbia pentagona
Euphorbia obesa ssp. obesa
Euphorbia obesa ssp. symmetrica
Euphorbia ferox
Euphorbia jansenvillensis 1
Euphorbia meloformis 1
Euphorbia polygona
Euphorbia stellispina
Euphorbia jansenvillensis 2
Euphorbia pillansii
Euphorbia pseudoglobosa 2
Euphorbia clandestina
Euphorbia pubiglans
Euphorbia pseudoglobosa 1
Euphorbia mammillaris 2
Euphorbia heptagona
Euphorbia susannae
Euphorbia silenifolia 2
Euphorbia multifolia 2
Euphorbia clava
Euphorbia bubalina
Euphorbia tuberosa
Euphorbia silenifolia 1
Euphorbia loricata 2
Euphorbia loricata 1
Euphorbia multifolia 1
Euphorbia oxystegia
Euphorbia bupleurifolia
Euphorbia omariana
Euphorbia platycephala
Euphorbia grantii
Euphorbia masirahensis
Euphorbia larica
Euphorbia rubriseminalis
Euphorbia meuleniana
Euphorbia balsamifera ssp. balsamifera
Euphorbia balsamifera ssp. adenensis
Euphorbia matabelensis
Euphorbia bongensis
Euphorbia oatesii
Euphorbia cuneata
Euphorbia smithii
Euphorbia caperonioides
Euphorbia insarmentosa
Euphorbia crotonoides
Euphorbia benthamii
Euphorbia marie-cladieae
Euphorbia socotrana
Euphorbia hamaderoensis
Euphorbia scheffleri
Euphorbia sp.
Euphorbia acalyphoides
Euphorbia hadramautica
Euphorbia longituberculosa
Euphorbia antso
0,37
1,39
2,08
3,51
4,13
1,15
3,51
1,4
4,47
0,85
1,99
2,55
1,33
2,74
1,76
3,17
5,24
0,46
1,08
1,66
2,58
1,03
3,3
1,44
4,17
1,66
3,17
1,02
6,72
2
3,27
3,44
0,95
3,94
1,19
4,74
0,6
2,69
2,12
5,73
0,49
0,88
1,77
1,24
2,46
3,84
0,9
1,02
1,72
1,03
2,41
8,21
0,63
1,21
3,19
1,19
4,01
1,72
2,54
0,88
1,19
0,31
1,72
5,15
0,27
2,67
1,2
3,07
1,52
5,44
10,68
4,53
0,74
6,18
1,72
1,06
1,84
7,43
4,99
2,06
2,43
5,43
14,07
0,48
1,61
6,18
7,24
3,79
2,2
3,47
6,89
3,4
16,49
9,47
4,39
5,88
2,95
11,41
0,57
1,21
19,59
3,17
1,11
2,99
8,04
25
20
15
10
5
0

Figure S11

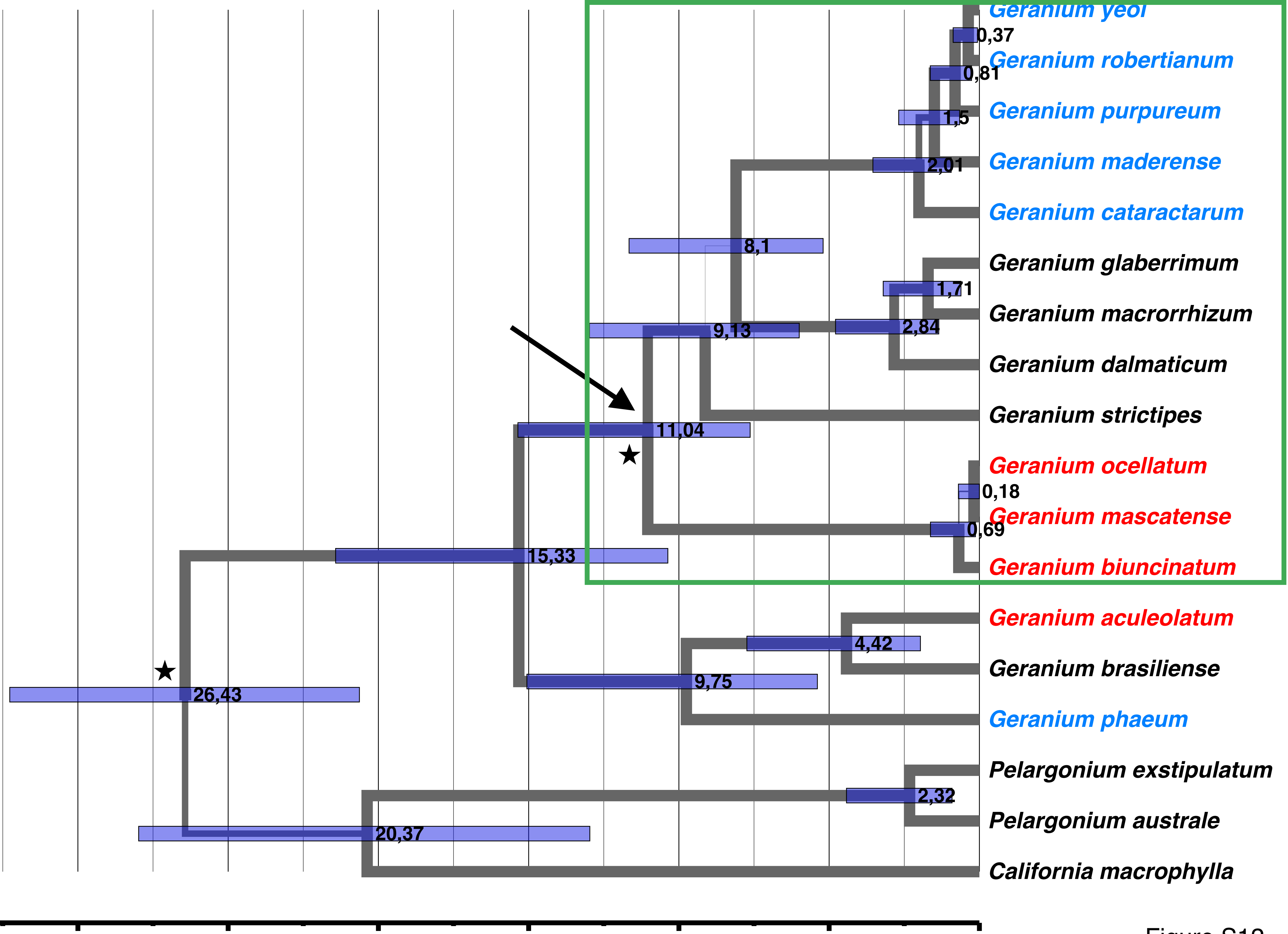

Geranium yeoi
0,37
Geranium robertianum
0,81
Geranium purpureum
1,5
Geranium maderense
2,01
Geranium cataractarum
8,1
Geranium glaberrimum
1,71
Geranium macrorrhizum
9,13
2,84
Geranium dalmaticum
Geranium strictipes
11,04
★
Geranium ocellatum
0,18
Geranium mascatense
0,69
15,33
Geranium biuncinatum
Geranium aculeolatum
4,42
Geranium brasiliense
★
9,75
26,43
Geranium phaeum
Pelargonium exstipulatum
2,32
Pelargonium australe
20,37
California macrophylla
30
25
20
15
10
5
0

Figure S12


Hypericum scioanum 1
0,13
Hypericum scioanum 2
2,98
Hypericum peplidifolium 1
0,56
3,85
Hypericum peplidifolium 2
Hypericum aethiopicum
5,23
Hypericum conjungens
3,32
Hypericum elegans
5,87
Hypericum kiboense 2
2,01
Hypericum kiboense 3
3,07
Hypericum kiboense 1
6,77
Hypericum tetrapterum
2,02
Hypericum undulatum
3,94
Hypericum maculatum 1
0,45
Hypericum maculatum 2
6,17
7,87
Hypericum perforatum 1
1,18
Hypericum perforatum 2
3,95
Hypericum perforatum 3
0,25
9,21
Hypericum perforatum 4
Hypericum aethiopicum ssp. sonderi
Hypericum olympicum 1
0,61
Hypericum olympicum 2
Hypericum linariifolium 1
1,06
Hypericum linariifolium 2
1,72
10,84
Hypericum humifusum 2
2,62
Hypericum humifusum 1
8,37
Hypericum papuanum
6,72
Hypericum perfoliatum
9,48
Hypericum scouleri
3,99
Hypericum formosum
9,88
6,93
Hypericum erectum
Hypericum richeri ssp. burseri
5,35
Hypericum montbretii
Hypericum glandulosum 1
0,19
Hypericum glandulosum 2
1,43
13,31
Hypericum reflexum
4,46
Hypericum coadunatum
1,87
Hypericum naudinianum
6,52
Hypericum caprifolium 1
1,22
Hypericum caprifolium 2
Hypericum tomentosum 1
10,27
1,85
Hypericum tomentosum 2
3,73
Hypericum somaliense
6,67
Hypericum pubescens
15,79
1,56
9,42
Hypericum psilophytum
11,4
Hypericum nummularioides
Hypericum aucheri
1,77
Hypericum triquetrifolium
5,67
Hypericum athoum
3,42
7,96
Hypericum annulatum
Hypericum montanum 1
0,35
Hypericum montanum 2
Hypericum cerastoides
18,86
Hypericum empetrifolium var. oliganthum
0,54
Hypericum xylosteifolium
0,94
Hypericum wightianum
6,37
Hypericum coris 1
0,94
9,74
Hypericum coris 2
Hypericum nummularium 1
10,88
0,83
Hypericum nummularium 2
Hypericum ericoides
Hypericum callithyrsum
5,68
Hypericum hyssopifolium
14,84
7,46
Hypericum helianthemoides
10,06
4,67
Hypericum scabrum
10,58
Hypericum confertum
12,22
Hypericum hirsutum
Hypericum linarioides
Hypericum choisianum
0,29
Hypericum patulum
1,92
24,4
Hypericum henryi
3,72
Hypericum acmosepalum
0,74
Hypericum lancasteri 2
2,18
4,4
Hypericum pseudohenryi
Hypericum curvisepalum
1,49
7,08
Hypericum lancasteri 1
2,79
Hypericum hookerianum
9,43
Hypericum synstylum
Hypericum socotranum
3,17
Hypericum balfourii
11,29
7,62
Hypericum longistylum
Hypericum ascyron
6,48
Hypericum elatoides
12,34
Hypericum balearicum 1
1,08
Hypericum balearicum 2
6,18
Hypericum calycinum
9,38
27,16
Hypericum dogonbadanicum
Hypericum bequaertii 1
15,99
0,12
Hypericum revolutum ssp. keniense
0,62
Hypericum revolutum ssp. revolutum
0,12
2,45
Hypericum bequaertii 2
Hypericum roeperianum 1
0,16
4,02
Hypericum roeperianum 2
Hypericum quartinianum 1
1,29
Hypericum quartinianum 2
Hypericum foliosum
0,9
Hypericum hircinum ssp. metroi
3,8
Hypericum grandifolium 1
0,94
5,21
Hypericum grandifolium 2
10,84
Hypericum androsaemum
Hypericum canariense 1
17,3
0,17
Hypericum canariense 2
21,02
Hypericum pamphylicum
Hypericum scopulorum
0,4
Hypericum tortuosum
Hypericum brachyphyllum
1,76
Hypericum fasciculatum 1
3
Hypericum galioides
33,75
4,83
Hypericum tenuifolium
2,87
Hypericum lissophloeus
4,06
6,89
Hypericum fasciculatum 2
Hypericum prolificum 1
1,06
11,62
Hypericum prolificum 2
3,49
Hypericum densiflorum
Hypericum ellipticum
13,6
Hypericum hypericoides 1
0,79
Hypericum hypericoides 2
7,41
Hypericum cistifolium 1
0,93
9,35
Hypericum cistifolium 2
Hypericum crux-andreae
5,88
Hypericum tetrapetalum
Hypericum laricifolium 1
0,45
Hypericum laricifolium 2
1,65
★
34,93
21,93
2,47
Hypericum mexicanum
Hypericum strictum
3,86
Hypericum brevistylum 1
0,24
Hypericum brevistylum 2
7,27
2,31
Hypericum quitense
8,77
Hypericum gentianoides
Hypericum drummondii
29,3
Hypericum mutilum ssp. mutilum
13,69
0,92
Hypericum mutilum ssp. boreale
3,29
Hypericum boreale
4,75
49,9
Hypericum lalandii
6,75
Hypericum canadense
2,42
Hypericum majus
Triadenum fraseri
Hypericum elodes 1
0,16
★
53,88
27,72
Hypericum elodes 2
Hypericum aegypticum ssp. webbii
1,33
Hypericum aegypticum ssp. aegypticum
Harungana madagascarensis
11,97
Vismia glaziovii
24,66
Psorospermum senegalense
Eliea articulata
60
50
40
30
20
10
0

Figure S13

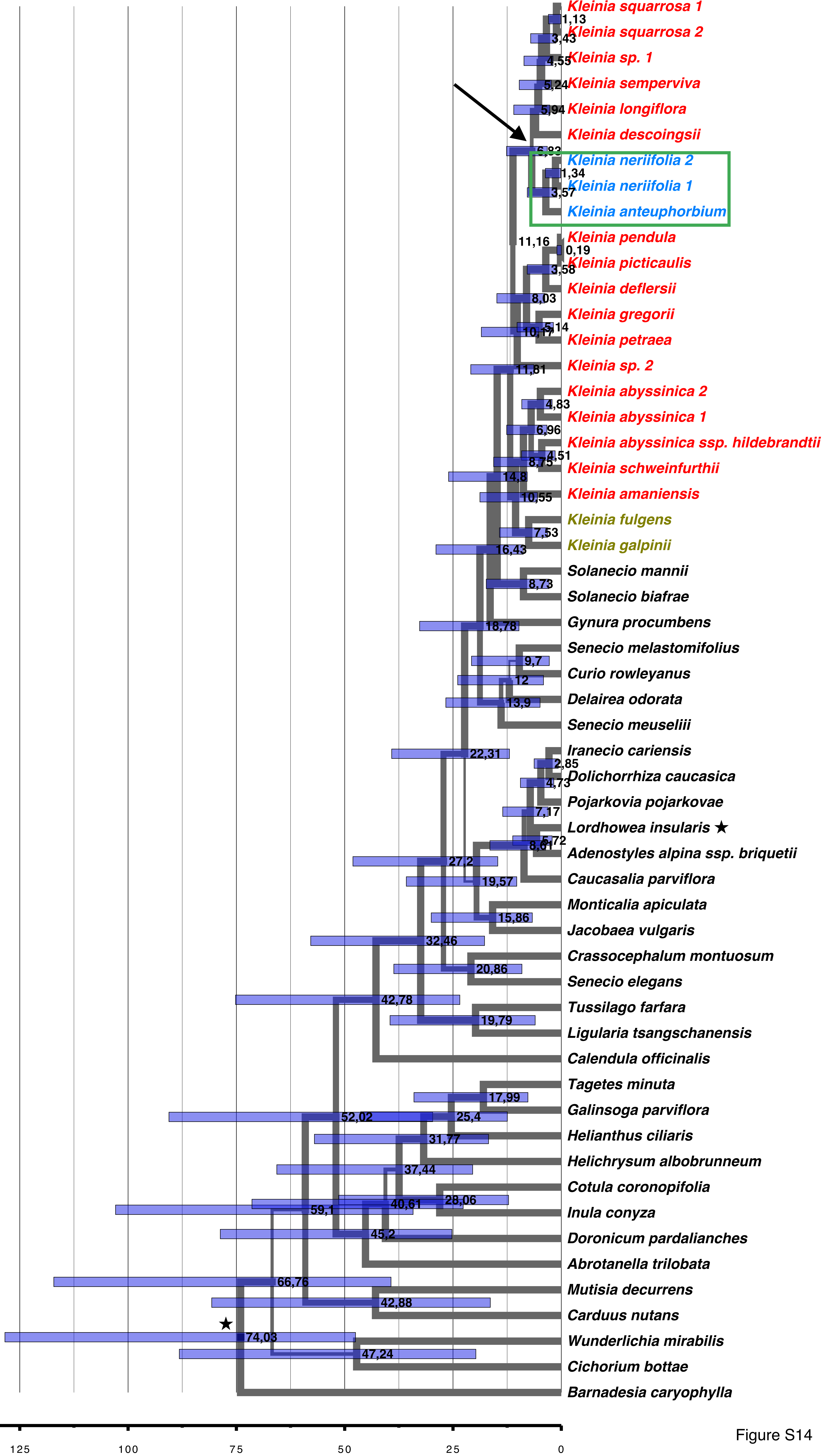

Kleinia squarrosa 1
1,13
Kleinia squarrosa 2
3,43
Kleinia sp. 1
4,55
Kleinia semperviva
5,24
Kleinia longiflora
5,94
Kleinia descoingsii
6,83
Kleinia neriifolia 2
1,34
Kleinia neriifolia 1
3,57
Kleinia anteuphorbium
Kleinia pendula
11,16
0,19
Kleinia picticaulis
3,58
Kleinia deflersii
8,03
Kleinia gregorii
5,14
10,17
Kleinia petraea
Kleinia sp. 2
11,81
Kleinia abyssinica 2
4,83
Kleinia abyssinica 1
6,96
Kleinia abyssinica ssp. hildebrandtii
4,51
8,75
Kleinia schweinfurthii
14,8
Kleinia amaniensis
10,55
Kleinia fulgens
7,53
Kleinia galpinii
16,43
Solanecio mannii
8,73
Solanecio biafrae
Gynura procumbens
18,78
Senecio melastomifolius
9,7
Curio rowleyanus
12
Delairea odorata
13,9
Senecio meuseliii
Iranecio cariensis
22,31
2,85
Dolichorrhiza caucasica
4,73
Pojarkovia pojarkovae
7,17
Lordhowea insularis ★
5,72
8,61
Adenostyles alpina ssp. briquetii
27,2
Caucasalia parviflora
19,57
Monticalia apiculata
15,86
Jacobaea vulgaris
32,46
Crassocephalum montuosum
20,86
Senecio elegans
42,78
Tussilago farfara
19,79
Ligularia tsangschanensis
Calendula officinalis
Tagetes minuta
17,99
Galinsoga parviflora
52,02
25,4
Helianthus ciliaris
31,77
Helichrysum albobrunneum
37,44
Cotula coronopifolia
28,06
40,61
59,1
Inula conyza
45,2
Doronicum pardalianches
Abrotanella trilobata
66,76
Mutisia decurrens
42,88
Carduus nutans
★
74,03
Wunderlichia mirabilis
47,24
Cichorium bottae
Barnadesia caryophylla
125
100
75
50
25
0

Figure S14

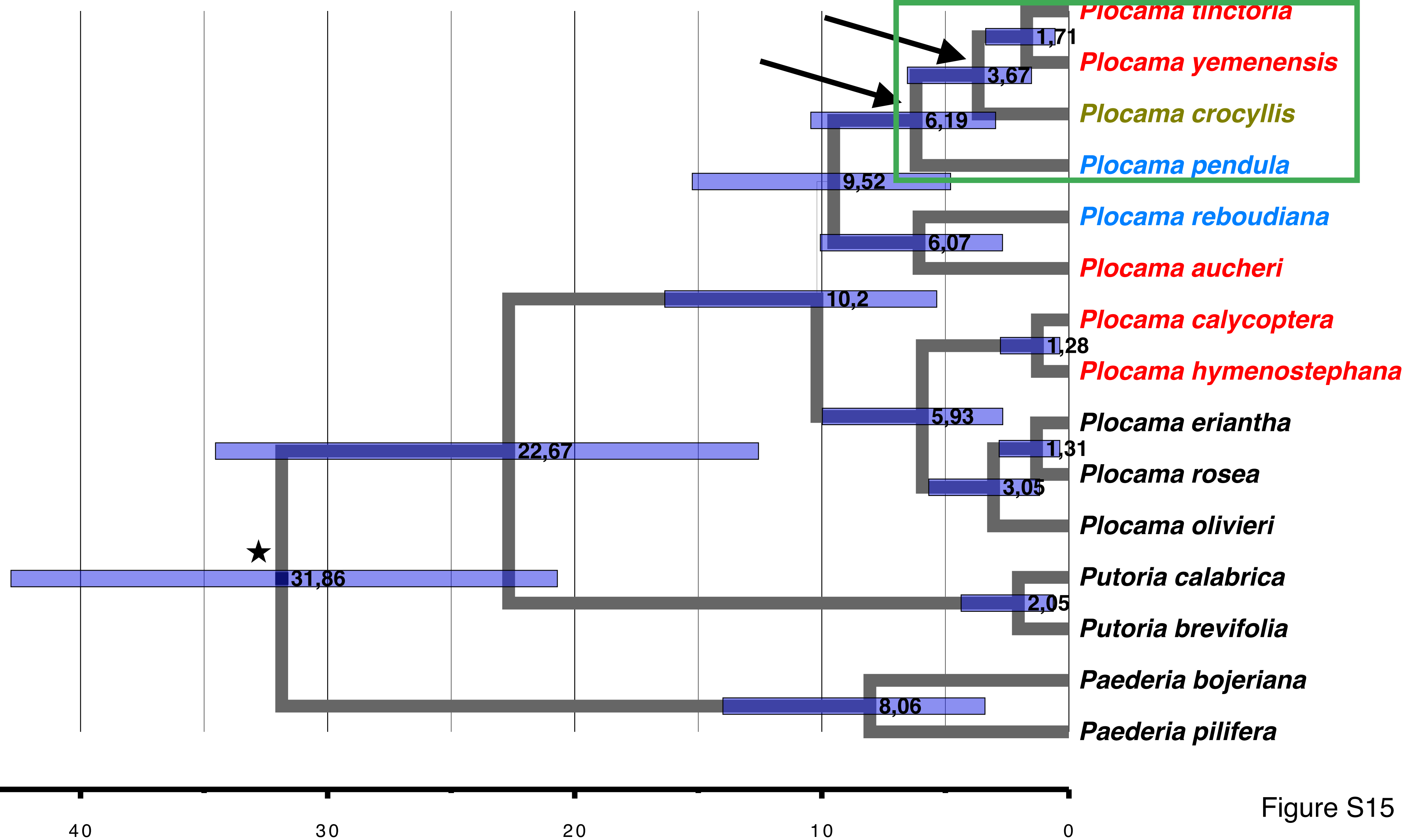

Plocama tinctoria
1,71
Plocama yemenensis
3,67
Plocama crocyllis
6,19
Plocama pendula
9,52
Plocama reboudiana
6,07
Plocama aucheri
10,2
Plocama calycoptera
1,28
Plocama hymenostephana
5,93
Plocama eriantha
1,31
Plocama rosea
22,67
3,05
Plocama olivieri
★
31,86
Putoria calabrica
2,05
Putoria brevifolia
Paederia bojeriana
8,06
Paederia pilifera
40
30
20
10
0

Figure S15

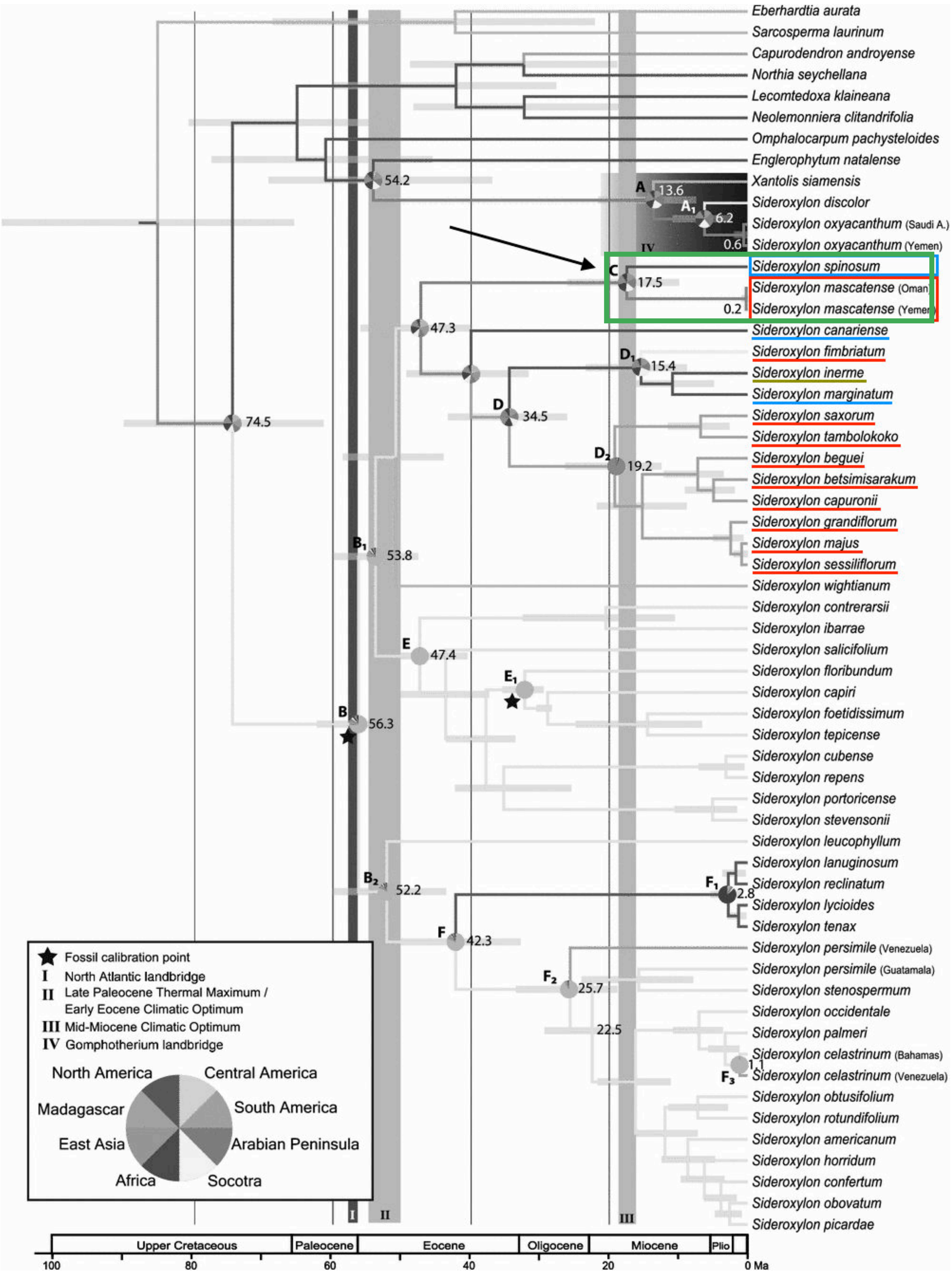
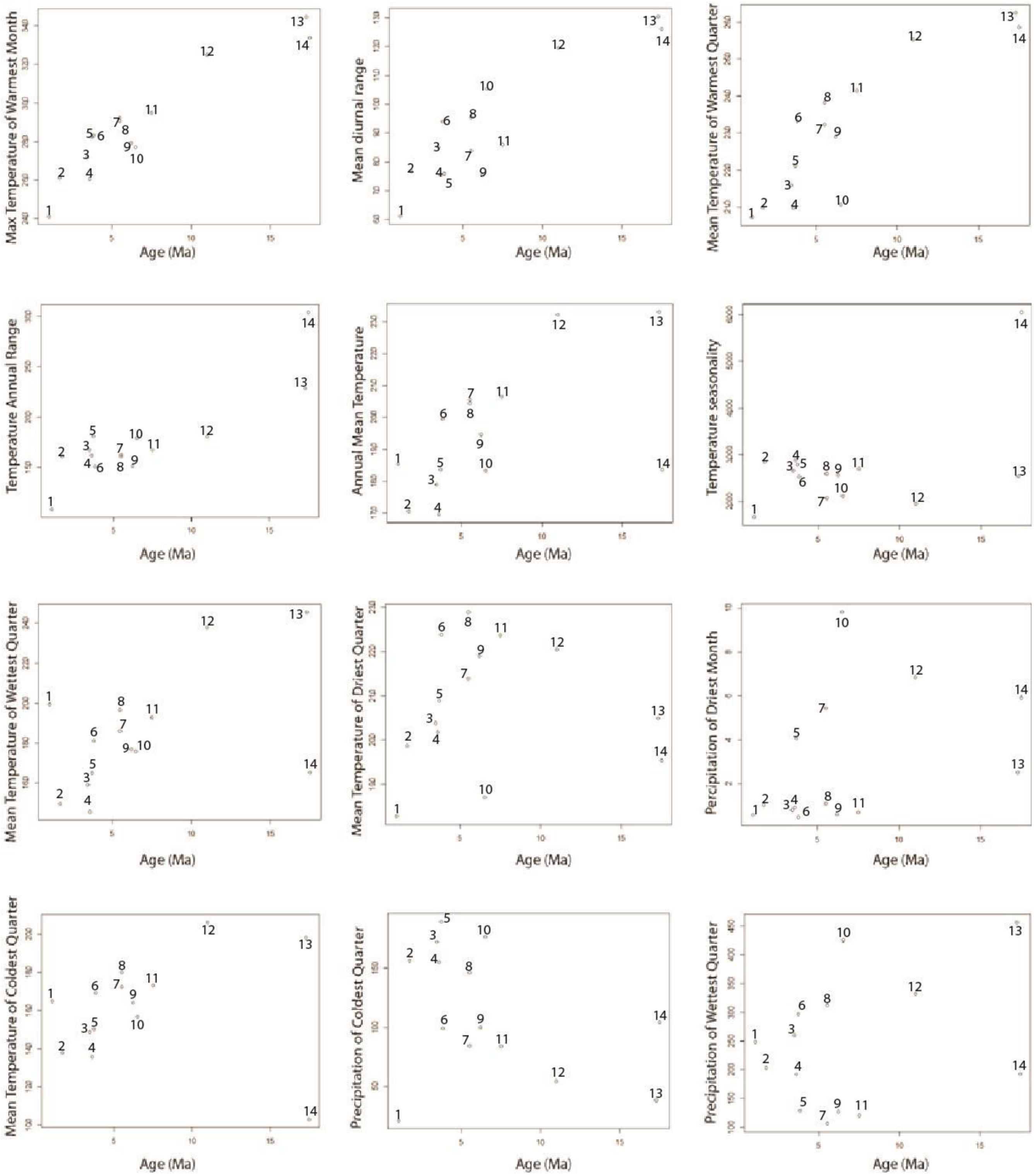

Eberhardtia aurata
Sarcosperma laurinum
Capurodendron androyense
Northia seychellana
Lecomtedoxa klaineana
Neolemonniera clitandrifolia
Omphalocarpum pachysteloides
Englerophytum natalense
Xantolis siamensis
Sideroxylon discolor
Sideroxylon oxyacanthum (Saudi A.)
Sideroxylon oxyacanthum (Yemen)
Sideroxylon spinosum
Sideroxylon mascatense (Oman)
Sideroxylon mascatense (Yemen)
Sideroxylon canariense
Sideroxylon fimbriatum
Sideroxylon inerme
Sideroxylon marginatum
Sideroxylon saxorum
Sideroxylon tambolokoko
Sideroxylon beguei
Sideroxylon betsimisarakum
Sideroxylon capuronii
Sideroxylon grandiflorum
Sideroxylon majus
Sideroxylon sessiliflorum
Sideroxylon wightianum
Sideroxylon contrerarsii
Sideroxylon ibarrae
Sideroxylon salicifolium
Sideroxylon floribundum
Sideroxylon capiri
Sideroxylon foetidissimum
Sideroxylon tepicense
Sideroxylon cubense
Sideroxylon repens
Sideroxylon portoricense
Sideroxylon stevensonii
Sideroxylon leucophyllum
Sideroxylon lanuginosum
Sideroxylon reclinatum
Sideroxylon lycioides
Sideroxylon tenax
Sideroxylon persimile (Venezuela)
Sideroxylon persimile (Guatamala)
Sideroxylon stenospermum
Sideroxylon occidentale
Sideroxylon palmeri
Sideroxylon celastrinum (Bahamas)
Sideroxylon celastrinum (Venezuela)
Sideroxylon obtusifolium
Sideroxylon rotundifolium
Sideroxylon americanum
Sideroxylon horridum
Sideroxylon confertum
Sideroxylon obovatum
Sideroxylon picardae
54.2
A
13.6
A_1
6.2
0.6
IV
C
17.5
0.2
47.3
D_1
15.4
D
34.5
D_2
19.2
74.5
B_1
53.8
E
47.4
E_1
B
56.3
B_2
52.2
F
42.3
F_1
2.8
F_2
25.7
22.5
F_3
1.1
I
II
III
Fossil calibration point
I North Atlantic landbridge
II Late Paleocene Thermal Maximum / Early Eocene Climatic Optimum
III Mid-Miocene Climatic Optimum
IV Gomphotherium landbridge
North America
Central America
Madagascar
South America
East Asia
Arabian Peninsula
Africa
Socotra
Upper Cretaceous
Paleocene
Eocene
Oligocene
Miocene
Plio
100
80
60
40
20
0 Ma

Figure S16

Figure S1.2. Relationships between age of divergence and the 19 climate variables considered in this study. Numbers in each graph follow the same legend as in Figure 2.



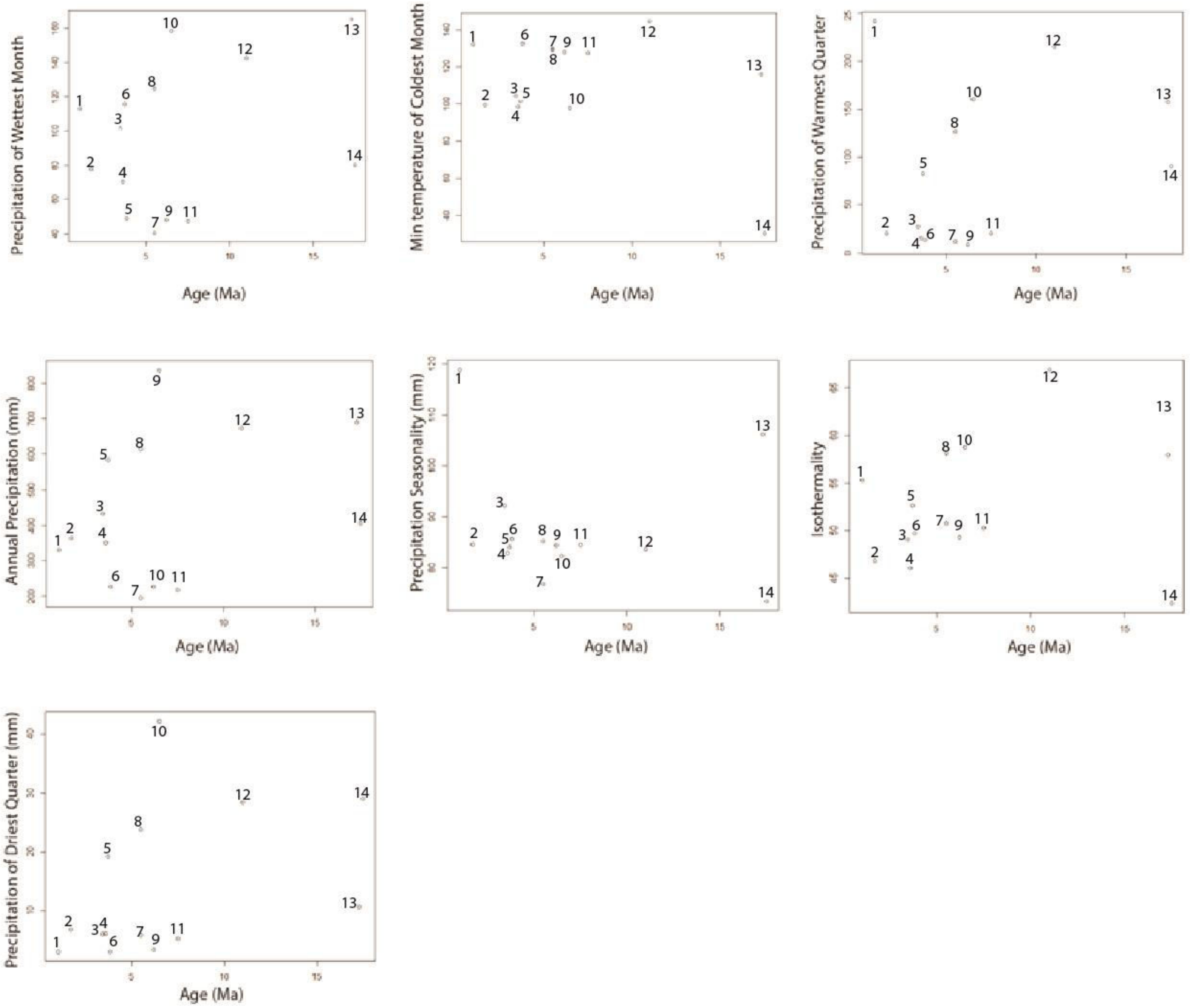

Precipitation of Wettest Month
Min temperature of Coldest Month
Precipitation of Warmest Quarter
Annual Precipitation (mm)
Precipitation Seasonality (mm)
Isothermality
Precipitation of Driest Quarter (mm)
Age (Ma)
1
2
3
4
5
6
7
8
9
10
11
12
13
14

Figure S1.3. Box-whisker-plots comparing the within-lineage variance versus between-lineage variance in climatic values for a set of uncorrelated variables that show a significant association to age of divergence. A lower within- than between-lineage variance is interpreted as supporting the assumption of niche conservatism within the studied lineages.

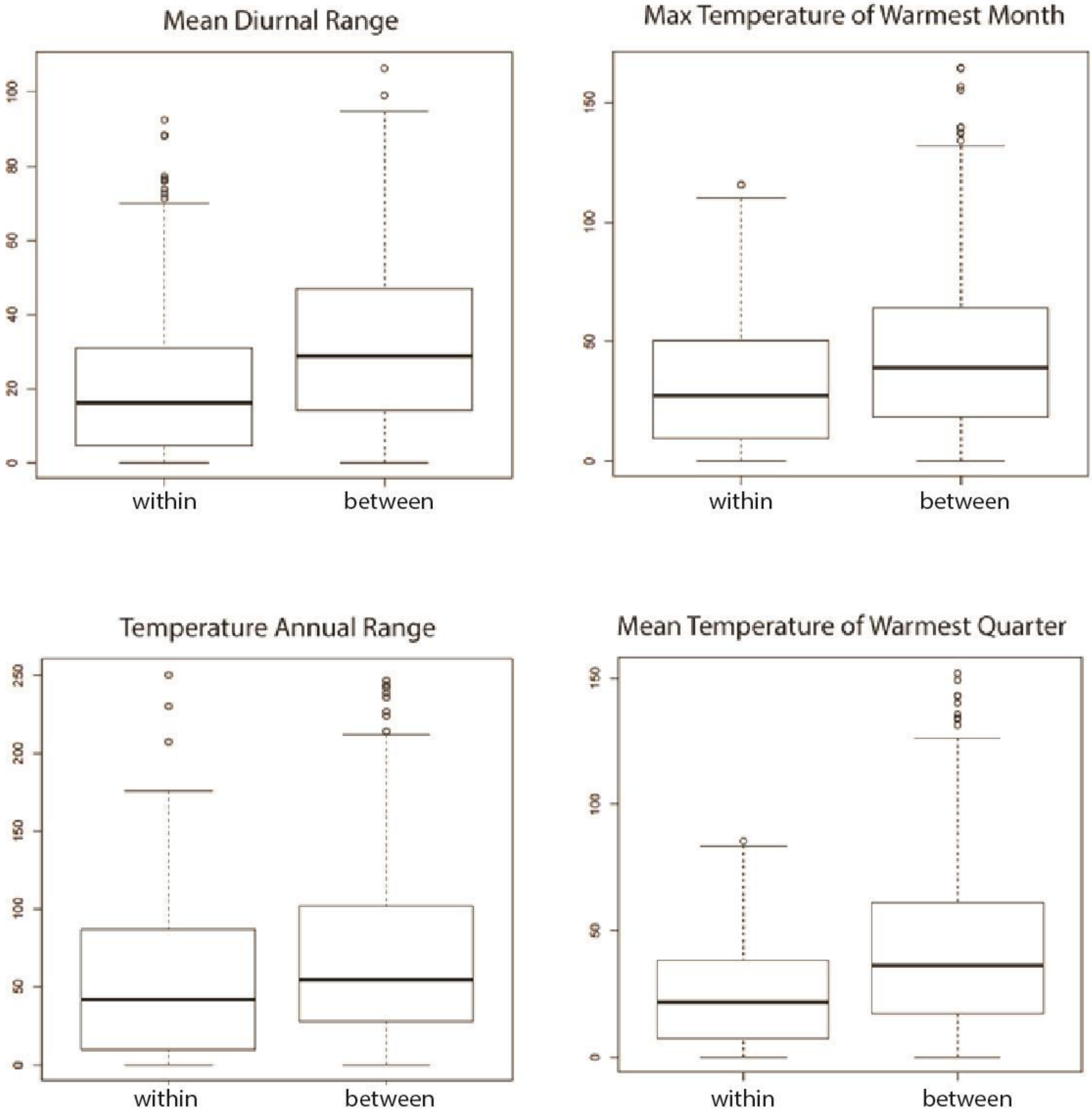
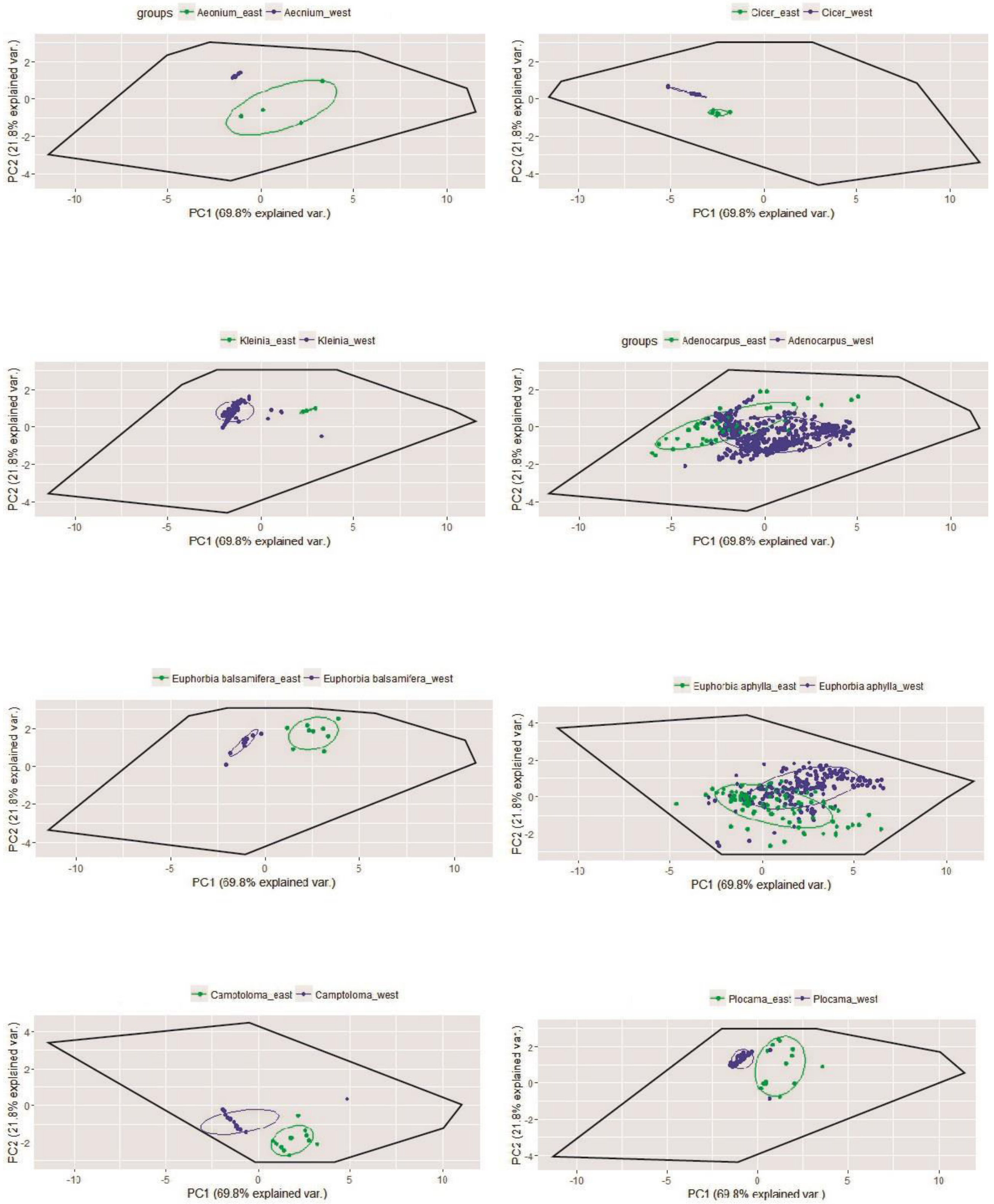
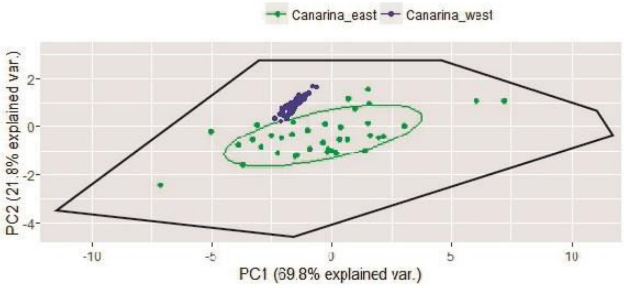
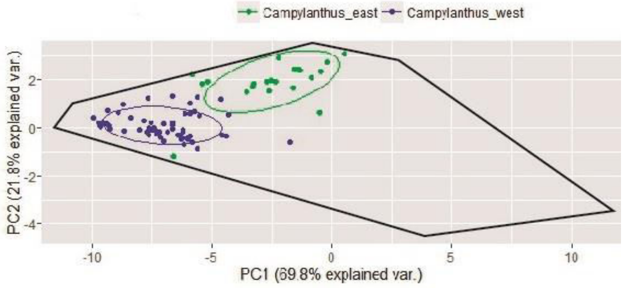
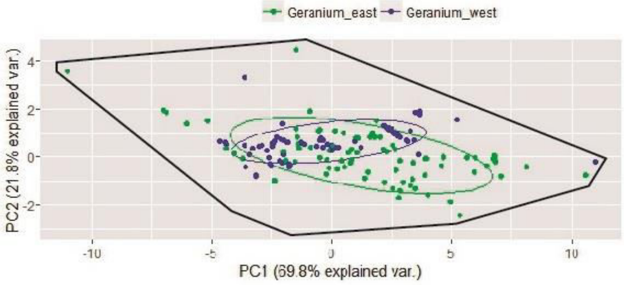


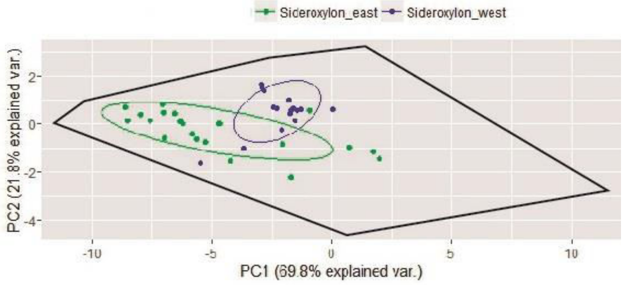
Figure S1.4. PCA analysis showing the degree of overlap in climatic niche space between the east (green dots) / west (blue dots) disjunct sister taxa, compared with the rest of species (convex hull in black).

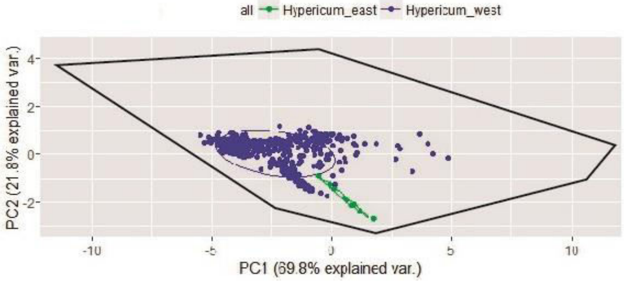



Canarina_east
Canarina_west
PC2 (21.8% explained var.)
PC1 (69.8% explained var.)
2
0
-2
-4
-10
-5
0
5
10


Campylanthus_east
Campylanthus_west
PC2 (21.8% explained var.)
PC1 (69.8% explained var.)
2
0
-2
-4
-10
-5
0
5
10


Geranium_east
Geranium_west
PC2 (21.8% explained var.)
PC1 (69.8% explained var.)
4
2
0
-2
-10
-5
0
5
10


Sideroxylon_east
Sideroxylon_west
PC2 (21.8% explained var.)
PC1 (69.8% explained var.)
2
0
-2
-4
-10
-5
0
5
10


all
Hypericum_east
Hypericum_west
PC2 (21.8% explained var.)
PC1 (69.8% explained var.)
4
2
0
-2
-10
-5
0
5
10

EUROPEAN

CERTIFICATES

IRD
Institut de recherche
pour le développement

Dr. Thomas L.P. Couvreur
Institut de Recherche pour le Développement
911, Av. Agropolis
34090 Montpellier

30 July 2015, Montpellier

To whom it may concern,

I write this report as a researcher belonging to the "Institut de Recherche pour le Dévoloppement (IRD)", Montpellier, France. I undertake research in plant systematics and evolution, especially in rain forest biodiversity and evolution. I have worked on phylogenetics, molecular dating and more recently on phylogeography of Central African taxa. I have never met Mr. Pisa and so I consider myself as impartial in relation to his work. I deliver this report as an external reviewer for the purposes of fulfilling the requirements to obtain the European Doctorate mention.

This PhD investigates the evolutionary history and phylogeography of a small genus (*Canarina*, three species) within the Campanulaceae family. *Canarina* shows an interesting "continental wide" disjunction between the Canary Islands (one species) and East Africa (2 species), a typical "Rand Flora type of distribution.

After an extensive introduction in Spanish, there are an impressive six multi authored chapters, 5 of which are first authored by the defendant. These chapters are either published/in press (1st, 3rd and 5th), submitted or in preparation for submission in top international peer reviewed journals. The first chapter (published in Molecular Ecology) naturally investigates the origin and causes of this disjunction within the genus. The defendant uses an extensive genetic dataset with sampling between and within species as well as species distribution modeling approaches to explore the timing and ecological evolution of the genus. The next two chapters (one accepted in Molecular Ecology and one to be submitted to Annals of Botany) explore the phylogeography of two species in two contrasting habitats: the Afromontane biome in East African (species *Canarina eminii*) and the Canary Islands (species *Canarina canariensis*). In both cases, population level sampling and state of the art phylogeographical methods are employed to unravel the evolutionary histories of these fascinating areas. Chapter IV (to be submitted to Conservation Biology) explores the impact of recent human actions of the genetic diversity of both previous species. This chapter has a strong applied conservation turn to it, nicely complementing the first three ones. Again, the use of two different species in two contrasting environments (islands and mountains) proves an interesting dual approach to such questions. Chapter V (published in Frontiers in Genetics) takes a step back and undertakes a comparative temporal analysis on the origin of the disjunct patterns common to the rand flora in a wide range of species (17 lineages). This chapter uses molecular dating methods and diversification rates analyses to show a common pattern of divergences between most species, coinciding with the onset of the African aridification during the Miocene and Pliocene. The last chapter (to be submitted to Ecological Letters) explores the ecological side of these disjunctions again based on a wide sampling of lineages (16). Using ecological niche modeling coupled with paleoclimatic data and temporal information (from chapter five), this chapter underlines the existence of a southern Sahara wet corridor allowing connection between populations during the Late Miocene. This became interrupted during the Mid Pliocene leading to the observed disjunctions.

This PhD explored a number of fascinating and long standing questions in African biogeography. The whole PhD is hugely diverse at numerous levels, showing that the candidate

IRD
Institut de recherche
pour le développement

Dr. Thomas L.P. Couvreur
Institut de Recherche pour le Développement
911, Av. Agropolis
34090 Montpellier

has mastered a wide range of methods, approaches and settings, going from molecular phylogenetics to population genetics to species distribution modeling, using microsatellite data, AFLP or sequence data all of this applied to very contrasting regions in the Canary Islands and Afromontane region of East Africa. The information one gathers from this PhD from the extensive introduction to the set of 6 perfectly complementary chapters is very impressive. And therefore I certify that the present dissertation meets the standards of quality, novelty and contributions necessary for its defense as a European Doctorate".

Thomas Couvreur

UMR DIADE
Centre IRD
911, Avenue Agropolis - BP 64501
F - 34394 MONTPELLIER CEDEX 5
Tél. 33 (0)4 67 41 62 21 - Fax: 33 (0)4 67 41 62 22

Dr. Luis Valente
+49 (0)157 581 895 90
University of Potsdam,
Karl-Liebknecht-Str. 24-25,
14476 Potsdam, Germany
valente@uni-potsdam.de

To whom it may concern,

I write this report as an Alexander von Humboldt Postdoctoral fellow at the Unit of Evolutionary Biology/Systematic Zoology of the Institute of Biochemistry and Biology of the University of Potsdam. My expertise lies in phylogenetics and island biogeography modeling. In particular, I develop dynamic stochastic models that allow the estimation of biogeographical rates using data on times of colonisation and speciation obtained from molecular phylogenies. I have also worked extensively in analyzing diversification rates, historical biogeography and trait evolution in large plant radiations. I have not met Mr. Mario Mairal and I have not collaborated scientifically with him, therefore I consider myself impartial in relation to his work. I deliver this report as an external reviewer for the purposes of fulfilling the requirements to obtain the European Doctorate mention.

The main topic of the thesis is to understand the general causes of the striking continent-wide disjunctions associated with Rand flora lineages and to reconstruct the evolutionary and biogeographical history of species that exhibit this fascinating pattern. To this aim, the thesis uses case studies of historical biogeography and ecological niches of genus *Canarina* (an excellent example of a lineage exhibiting a Rand flora distribution), as well as meta-analyses encompassing multiple Rand flora lineages. The thesis begins with a comprehensive introduction that summarizes the state-of-the-art and methods used to study the biogeography of the Rand Flora, Afromontane sky-islands and Canary island biogeography. The thesis is then divided into several research chapters, three of which have already been accepted for publication in international peer-reviewed journals.

Chapter 1 is devoted to reconstructing the continent-wide disjunction of the three species of genus *Canarina*. Using a combination of phylogenetic divergence-dating analyses (including a new method for calibrating heterogenous datasets comprising both above-species and population-level sampling) and ecological niche modeling, the results of this chapter reveal that the divergence between East African and Canary island lineages occurred in the Miocene and was likely associated with aridification-driven extinction in a vicariance context. Chapters 2 and 3 conduct detailed biogeographical and evolutionary reconstructions for two *Canarina* species, respectively *C. eminii* and *C. canariensis*, using data from both chloroplast and nuclear genomes. These two studies highlight the importance of the formation of recent geographical barriers (Rift Valley in East Africa, and volcanic eruptions and landslides in the Canaries) in shaping current species distribution and genetic affinities. Chapter 4 uses microsatellite data and ABC simulations to study the demography of *C. canariensis* and *C. eminii*, revealing recent population declines in the two species and highlighting the need for conservation action in the few remaining fragments of Afro-Macaronesian forests. Chapter 5 presents a meta-analysis of phylogenetic patterns and divergence times of 17 plant lineages that display Rand Flora-type disjunct distributions, and its results show that the majority of disjunctions coincide with the onset of aridification in Africa. Finally, Chapter 6 estimates the extant climatic tolerances of 16 Rand flora lineages and projects them back in time to infer ancestral distributions, in order to investigate the relationship between species climatic tolerance and the age of the Rand Flora disjunction. The results of this Chapter reveal an interesting relationship between age of divergence of Rand Flora lineages and the climatic niches of those lineages (particularly related with temperature), suggesting that climate change in the Miocene and Pliocene in this region drove lineage separation at a rate that matches climatic niches.

The thesis is both coherent and diverse in its scope. In order to produce this dissertation, Mr. Mairal has clearly learnt and employed an impressive array of scientific skills. This is demonstrated by the fact that the thesis has involved fieldwork in East Africa and Canaries, molecular work in laboratory, extensive literature surveys and advanced statistical and phylogenetic analyses. In all chapters the student reveals that he has acquired very good writing and analytical skills. The bibliography is thorough and up-to-date, with the most relevant previous work being cited. Analyses throughout were conducted rigorously using appropriate and often very original protocols. One aspect that I particularly appreciated in this thesis is the fact Mr. Mairal never loses sight of the bigger picture and the general significance of the questions being addressed. For example, even when conducting very specific and detailed analyses of phylogeography of single species (*C. eminii* in chapter 2 or of *C. canariensis* in chapter 3), an effort was made to compile detailed tables with information on the patterns previously found in different lineages of plants and animals occurring in the same regions, in order to place the results into a wider context.

In sum, this dissertation presents a valuable synthesis of the research conducted on the Rand flora in recent years, and, most importantly, it significantly moves this field of research forward by providing exciting new results and ideas. It also presents novel rigorous methodologies that will certainly provide a useful framework for further evolutionary and biogeographical studies of this fascinating biogeographical pattern. In my opinion, the thesis clearly meets the standards of quality, novelty and contributions necessary for its defense as a European Doctorate.

Potsdam, 27-07-2015

Luis Valente

Luis Valente, PhD