HISTORICAL BIOGEOGRAPHY OF ENDEMIC PLANTS IN THE CARIBBEAN AND *PODOCARPUS* AS A CASE STUDY

By

© María Esther Nieto-Blázquez

A dissertation submitted to the

School of Graduate Studies

In partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Biology Department, Faculty of Science

Memorial University of Newfoundland

July 2019

St. John's, Newfoundland and Labrador

Abstract

Despite numerous floristic studies of the Caribbean region, the large-scale evolutionary origins of its diversity remain relatively under-explored. Here, I add to the empirical knowledge of the origin and assembly of the Caribbean flora by large-scale molecular analysis of the biogeography and phylogenetics of its constituent genera, with special emphasis on the conifer genus *Podocarpus* L'Hér. ex Pers. and its endemic species in the Antillean island chains. Connection of the Greater Antilles to northern South America by a late Eocene / early Oligocene land bridge, GAARlandia, has been hypothesized to facilitate colonization of the Caribbean islands. Well-calibrated estimates from molecular data of the ages of extant Caribbean endemic plant genera show a combination of relatively recent (late Oligocene / early Miocene) and older (late Paleocene / early Eocene) lineages, such that the GAARlandia hypothesis is not necessary to explain floral colonization of the Caribbean. Ancestors of most endemic genera included in my study were of Antillean origin. I show that diversity in Caribbean *Podocarpus* is paraphyletic, the result of a single colonization of the Greater Antilles from South America in the late Oligocene, species diversification leading to endemism beginning in the early Miocene, and dispersal to the Lesser Antilles from a Greater Antillean ancestor later in the early Miocene. Detailed examination of the phylogeography of the two endemic species on Hispaniola, *P. buchii* and *P. hispaniolensis*, shows an initial colonization to the southern palaeo-island of Hispaniola by a *P. buchii* ancestor. Subsequent northward stepping-stone migration is evident in the localization of genetic clusters across the main cordillera systems. *P. hispaniolensis* arose by progenitor-derivative speciation in the Central

Cordillera. The assembly of Caribbean flora has been complex, and geology, vicariance, dispersal, and *in situ* speciation, together shaped the biotic assembly of the islands.

"Abunda la tierra aspera del Cibao (de Ciba, piedra) de pinos mui altos que no llevan piñas, por tal orden compuestos por naturaleza, que parecen azeytunos del Axarafe de Sevilla"

C. Columbus

Acknowledgments

There are many people I want to thank and who have been involved in one way or the other in the process of this PhD thesis.

I would like to start thanking my supervisor, Dr. Julissa Roncal for her support and guidance through this project. Thanks Julissa for being always available to help and helping me believe in my project when I had doubts about it. Also, thanks for your help with personal matters outside work, I will always be grateful for your mentorship and your help. Thanks to my co-supervisor Dr. Alex Antonelli for his help during this thesis and for hosting me at University of Gothenburg during the summer of 2015.

I would also like to thank my committee members, Dr. Dawn Marshal and Dr. Steve Carr for their valuable input during meetings and discussions. Your knowledge greatly improved my research.

I thank my funding sources, NSERC-Discovery grant to Julissa Roncal and MUN's President's Doctoral Student Investment Fund.

Thanks to Compute Canada for their assistance, in particular, thanks to Dr. Oliver Stueker for his advice and help on the use of the cluster.

Thanks to the Royal Botanic Garden Edinburgh, Royal Botanic Gardens Kew and Montgomery Botanical Center for supplying Podocarpaceae samples for my research.

Many thanks to colleagues from the Caribbean. Thanks to Francisco Jimenez for hosting me at Jardín Botánico Nacional Dr. Rafael Ma. Moscoso and providing everything I needed for my collections. Special thanks go to Teodoro Clase for the help during fieldwork and locating the podocarps in the field. Also thanks to Pedro, Angel and Yuley

for the assistance during fieldwork and at the Herbarium. In Cuba, I should sincerely thank Ramona Oviedo for helping me organizing the logistics of my trip to Guantanamo. I sure enjoyed your deep knowledge of the Cuban flora and your vitality. Thanks to Gerardo Begué from Unidad Provincial de Servicios Ambientales (UPSA) for his help during fieldwork in Ojito de Agua, Guantanamo.

Thanks to Dr. Paula Quiroga for sharing *Podocarpus* samples and helpful comments and thoughts during this thesis.

I would like to thank each of my present and past lab mates: Hugo Morales, Zobayer Alam, Fritz Pichardo, Sander Boisen and Nicolas Peñafiel, for many great moments in the lab and outside the lab. We sure have laughed a lot and those memories are priceless. Particular thanks to my dear friend Fritz, who was always happy to help me with technicalities when I had troubles, and with whom I shared many conversations about plants, herbaria and, "against my will", Linux and computing gibberish.

Thanks to all my St. John's friends for making my time in Newfoundland much more enjoyable and special.

I want to deeply thank my family for their support and encouragement not just during this dissertation. Thanks to my parents for their love and support throughout my life, and for teaching me the sense of determination.

Finally, thanks to Natalia, for her love, companion, patience, support and a huge of understanding during low times I had along the process. Sorry I tricked you to come here, but thanks for joining me in the Newfoundland adventure.

Table of Contents

List of Tables

List of Figures

List of Abbreviations and Symbols

AAR - Ancestral area reconstruction

AIC - Akaike information criteria

APG - Angiosperm phylogeny group

BAMM - Bayesian analysis of macroevolutionary mixtures

BEAST - Bayesian evolutionary analysis by sampling trees

BI - Bayesian inference

BiSSE - Binary state speciation and extinction model

bp - base pairs

BS - Maximum likelihood bootstrap values

DEC - Dispersal extinction cladogenesis

DIVA - Dispersal-Vicariance analysis

DIYABC - Do it yourself Approximate Bayesian Computation

ESS - Estimated sample size

GAARlandia - Greater Antilles + Aves Ridge landbridge

GBS - Genotyping by sequencing

GTR - General time reversible model

HPD - Highest posterior densities

IBD - Isolation by distance

IPNI - International plant names index

K - Number of genetic clusters

LDD - Long distance dispersal

Ma - Millions of years ago

MCC - Maximum clade credibility

MCMC - Markov chain Monte Carlo

ML - Maximum likelihood

MMCO - Middle Miocene climate optimum

mns - minimum number of samples

MRCA - most recent common ancestor

NCBI - National Center for Biotechnology Information

NGS - Next generation sequencing

PP - Bayesian posterior probabilities

RAD - Restriction site associated DNA

SNP - Single nucleotide polymorphism

UCLN - uncorrelated lognormal

Γ - GTR + I + Γ

Δ*K* - Delta *K*

λ - speciation rate

μ - extinction rate

q - transition rate

List of Appendices

CHAPTER 1

Introduction and Thesis Overview

1.1 Introduction

1.1.1. Historical biogeography

Biogeography is the discipline that aims to explain the geographical distribution and evolution of organisms in space and time (Ball, 1976). Traditionally, biogeography can be divided into two sub-disciplines: ecological biogeography and historical biogeography (Crisci, 2001). The most important difference between these two approaches are their temporal and spatial scales. Ecological biogeography aims to explain the distribution of organisms by examining interactions between organisms and the abiotic environment, often over relatively short temporal spans and small geographical scales. Historical biogeography aims to reconstruct the historical and biogeographic events that have led to the current distribution of biota (Posadas et al., 2006), usually at much longer temporal scales (e.g. geological epochs). Although the latter will be the focus of this thesis, it is important to acknowledge that the distribution of organisms is not only due to ecological or historical factors, but a combination of both (Santos & Amorim, 2007; Antonelli et al., 2018).

The way organisms are distributed around the globe was of major interest to early biologists and naturalists of the 18th and 19th centuries (Ebach et al., 2003). The work of Buffon (1761), established the first principle of biogeography, Buffon's Law. Based on his observations of mammals from the New and Old Worlds, Buffon postulated that geographically isolated regions that share similar environmental characteristics presented different assemblages of species. Later authors affirmed Buffon's Law for plant species and increased the understanding of underlying mechanisms by studying the relationships of biota distributions with climate and latitudinal and altitudinal ranges (Forster, 1778;

von Humboldt, 1905). Since the geographical world was seen as immutable and static (Posadas et al., 2006; Lomolino et al., 2010), early evolutionary biologists, such us Darwin (1859) and Wallace (1876, 1892), presumed that the biotic distributions were the result of dispersal from centers of origin and adaptation of organisms to new environments (Morrone & Crisci, 1995).

Two key advances in the second half of the 20th century shifted views of how species are distributed. Firstly, the acceptance of the Plate Tectonics Theory (Hammond, 1971) established a new paradigm to explain the distribution of organisms, especially those with disjunct distributions by means of vicariance. The breaking-up of continents and subsequent continental drift appeared to be a plausible and elegant way to explain the distribution of related taxa in geographically distant areas. For example, plate tectonics was considered the mechanism by which the distribution of numerous Southern Hemisphere plant families could be explained (Raven & Axelrod, 1974). Secondly, advances in molecular systematics and use of the molecular clock allowed researchers to correlate the timing of lineage divergences with vicariance events related to, for instance continental break-ups or climatological events, and to test hypotheses of biogeographic-event causality (Posadas et al., 2006). Numerous molecular dating analyses have, however, shifted the general scientific consensus away from vicariance in favor of long distance dispersal (LDD). For many plant groups in which biogeographic patterns were thought to be explained by vicariance, lineage divergence ages seem to be too young for plate tectonics to have played a role in their diversification (Yuan et al., 2005; Clayton et al., 2009; Pirie et al., 2015). The debate between vicariance and dispersal continues in the field of plant biogeography.

Several authors (Morrone & Crisci, 1995; Crisci et al., 2003; Santos & Amorim, 2007) have thoroughly reviewed the different approaches used in historical biogeography in the last century to explain the distributions of biota over time. In particular, molecular phylogenies have revolutionized the way we explain biogeographic patterns and have been extensively used in the last decades to reconstruct the evolutionary history of organisms in different spatio-temporal contexts (Ebach et al., 2003; Magallón, 2004; Posadas et al., 2006).

1.1.2. Testing biogeographic hypotheses

Vicariance and dispersal are often regarded as two distinct geographical processes for diversification and distribution shifts of species over time. They are not however mutually exclusive. Since historical events cannot be reproduced, researchers observe a pattern, and try to find explanations *a posteriori* about the processes causing the observed pattern (Platnick & Nelson, 1978; Crisp et al., 2011). In other words, we can test an unobserved process if a hypothesis is formulated and the observable outcome of the analysis can be contrasted with the alternative hypothesis (Penny & Phillips, 2004).

Crisp et al. (2011) presented a hypothetical example (Figure 1.1) of four different scenarios of divergence times for three lineages, which exemplifies how molecular dated phylogenies are used to test biogeographic hypotheses. One of the assumptions is that for a given vicariance event different lineages should show congruent patterns of distribution (Rosen, 1978; Crisp & Cook, 2007; Riddle et al., 2008), since the suggested barrier might potentially disrupt gene flow across diverse groups of organisms. When a vicariance hypothesis is rejected, dispersal is commonly substituted as the likely explanation of a

disjunct distribution (Crisp et al. 2011). The rarity and stochastic nature of dispersal events make them challenging to study and difficult to measure (Cain et al., 2000; Nathan et al., 2003). Despite this, dispersal hypotheses can also be formulated *a priori*.

Gillespie et al. (2012) showed how a good understanding of dispersal vectors, together with geological and environmental factors, could help in predicting patterns of dispersal. In their work, they examine different dispersal vectors, like winds, birds and ocean currents, to infer directionality, route, and potential arrival and establishment of propagules of terrestrial plants and animals and to formulate LDD predictions in the Pacific Ocean. Thus, previous knowledge on the dispersal ability of organisms could be integrated into the formulation of hypotheses that aim to test LDD as a process explaining disjunct distributions. However, the expected patterns might be obscured by the combination and interaction of the different dispersal vectors implicated in a LDD event (Nathan, 2008).

1.1.3. Molecular phylogenetics and next generation sequencing in biogeography

Various authors have highlighted the importance of using phylogenies to test biogeographical hypotheses (Santos & Amorim, 2007; Crisp et al., 2011; Ronquist & Sanmartín, 2011). Phylogenies are depictions of the evolutionary relationships among an ancestor and all its descendants (Hennig, 1966), represented in a diagrammatic fashion. The phylogenetic method relies on Darwin's foundational concept that all living species are connected by lines of common descent with modification that show the evolution of organisms. The development of molecular phylogenetic methods coupled with methods to temporally calibrate phylogenies provides a suitable framework to estimate lineage

divergence and evolutionary processes on a timescale (Magallón, 2004; Ho, 2014; Bell, 2015).

The foundation of molecular dating (i.e. the molecular clock) was proposed by Zuckerkandl & Pauling (1962), who suggested that the divergence time between two species could be estimated from the number of changes in their DNA sequences. This required a constant rate of molecular evolution through time, and improvements of the molecular clock have been largely directed towards relaxing the assumption of such a constant rate. Sanderson (1997) proposed a new dating method, the *relaxed clock model*, which allows different rates of molecular evolution through time and among different taxa, which is arguably more plausible for most lineages (Sauquet, 2013). Molecular dating aims to date all nodes in a phylogenetic tree in order to make inferences about the timing of lineage divergence. A phylogenetic tree in which branches are proportional to time is called a chronogram. In order for a phylogenetic tree to represent absolute ages (e.g. millions of years), it must be temporally calibrated, either by placing an age constraint on at least one node (using ages of known geological events and fossils), or by using known rates of DNA substitution (Ho, 2014). By timing lineage divergences, we can test hypotheses about the influence of particular geological and (or) climatic events on such lineage divergences.

Evolution of geographic range within a phylogenetic context has also received attention in the last decades, particularly since the proliferation of new methods for biogeographic inference (Ree & Smith, 2008; Lamm & Redelings, 2009). Phylogenies are also the basis for ancestral area reconstructions (AARs) of lineages. In the same way, we study trait or character evolution through phylogenies, we can consider geographical

location as a heritable trait to study the evolution of species' range (Bremer, 1992; Lamm & Redelings, 2009; Crisp et al., 2011). Biogeographic reconstructions are crucial to link a species' range evolution to speciation events, especially allopatric speciation, either achieved by vicariance or dispersal (Lamm & Redealing, 2009). Different methods of biogeographic inference make different assumptions. For example, in reconstructions under a parsimony framework, the assumption is that the ancestral state represents one state or the other (presence in an area or not), but not both at the same time, so taxa with widespread distributions are not considered (Fitch, 1971). More complex parsimony and maximum likelihood methods can incorporate specific dispersal and vicariance scenarios into their models, and allow species to occur in more than one area (Pirie et al., 2012). Under the dispersal-vicariance (DIVA) approach, (implemented in DIVA; Ronquist, 1997), which works under a parsimony framework, dispersal events are penalized; thus, this approach is biased against dispersals (Ree et al., 2005). The dispersal-extinction-cladogenesis (DEC) approach (Ree et al., 2005; Ree & Smith, 2008) also allows occupation of multiple areas under a maximum likelihood approach. Under the DEC model the extension and contraction of species' ranges are due to dispersal to a previously unoccupied area and local extinction within areas, therefore this method is biased against vicariance (Lamm & Redealing, 2009). As for phylogenetic reconstruction, the model choice used in AARs will affect the interpretation of results (Crisp et al. 2011). Although molecular dating analyses add a temporal framework for studying ancestral lineage divergences, biogeographical reconstructions provide phylogenies with a spatial component to ancestral lineage divergences.

In recent years, there has been a great development in high-throughput sequencing technologies that have helped advance the field of molecular evolution. In comparison to the traditional Sanger sequencing, next generation sequencing (NGS) produces a much larger amount of genomic data for a larger number of samples in a time-efficient and a less expensive way (McCormack et al., 2013). Despite the advantages and potential for phylogenetics and phylogeographic research, NGS is still underused in these fields (Carstens et al., 2012; Eaton & Ree, 2013), in comparison to other fields such as population genetics, metagenomics, and disease genetics (Mardis, 2008). A problem in phylogenetics is the difficulty of resolving relationships amongst closely related species, or recently diverged lineages (Maddison & Knowles, 2006). The problem might be due to the lack of phylogenetic signal and/or the biological processes of incomplete lineage sorting and horizontal gene transfer (Eaton & Ree, 2013). These can lead to low support, short branches, and poor resolution of the phylogenetic reconstructions. NGS has the potential to address the lack of phylogenetic signal component.

NGS methods such as RAD sequencing (RAD-seq, Baird et al., 2008) and Genotyping by sequencing (GBS, Elshire et al., 2011) are particularly attractive in systematics. This is because they can produce genomic data for non-model organisms for which there is no full genome available (Rubin et al., 2012; Eaton & Ree, 2013). These methods can generate data for thousands of loci with phylogenetic informative markers, such as single nucleotides polymorphisms (SNPs). SNPs have been traditionally used in population genetic studies (Carstens et al., 2012; Leaché & Oaks, 2017) but in the last few years there has been an exponential growth in the number of studies that have used

SNP data to conduct phylogenetic and phylogeographic research (e.g. Card et al., 2016; Dupuis et al., 2017; Hamon et al., 2017; Klimova et al., 2018).

1.1.4. Island biogeography

Explaining the biological diversity and endemicity of island systems has been a major topic of study in biogeography (Darwin, 1859; Carlquist, 1965; Losos & Ricklefs, 2009). The physical isolation and discrete nature of these relatively small areas can be seen as ideal scenarios to test hypotheses and study the evolutionary processes that lead to the diversification and uniqueness of its biota (Emerson, 2002; Ricklefs & Bermingham, 2008). A key work to understand the dynamics of islands is *The Theory of Island Biogeography* by MacArthur & Wilson (1967), which addressed two patterns recognized on islands. The first was the relationship of species richness to island size (species richness increases with island size), and the second, species richness and isolation of islands (species richness decreases with island isolation).

The theory also proposed that there is a dynamic equilibrium between rates of immigration and extinction, which maintains a fairly constant number of species, despite the turnover of species composition on the island. Because of their nature and origin, oceanic islands are of special interest to evolutionary biologists because they have never been connected to the continent. Researchers can therefore study their evolutionary processes in "isolation" (Queiroz, 2005). Populations on oceanic islands are usually initiated by rare founder events, which result from a combination of both stochastic and selective processes (Lomolino et al., 2010). The degree of isolation of island systems from adjacent mainland areas affects not only the species richness, but also lineage

divergence. Islands close to the mainland act as a recipient of continental taxa, thus opportunities for divergence are limited by the continuous immigration of mainland relatives and substantial gene flow from the neighboring continental masses (Ricklefs & Bermingham, 2008). To the contrary, remote islands might present new and unexploited habitats, and new immigrants might have a high probability of *in situ* speciation, increasing the predisposition for adaptive radiations. This results in high numbers of novel elements (i.e. endemics) (Baldwin et al., 1991; Jorgensen & Olesen, 2001; Lomolino et al. 2010). The view of oceanic islands as mere recipients of biodiversity has been recently challenged in a study of Macaronesian bryophytes, where Patiño et al., (2015) show that oceanic elements act as reservoirs of novel genetic diversity for the assemblage of continental floras.

The use of molecular phylogenies has been essential to test hypotheses on the diversification events and evolutionary processes leading to island biotas. Molecular phylogenies are useful to study relationships between island taxa and their continental congeners, the origins of species groups on island systems, and the number and sequence of colonization events and possible recolonizations of continental masses (Emerson, 2002; Carine et al., 2004; Stuessy et al., 2014). A single colonization event should result in a monophyletic lineage in the phylogeny, whereas multiple colonization events would result in a paraphyletic or polyphyletic grouping of the insular species (Figure 1.2, from Emerson, 2002). Furthermore, dated phylogenies and biogeographic reconstructions allow us to discriminate between multiple colonization events versus *in situ* speciation occurring in island systems (Presgraves & Glor, 2010). Figure 1.3 shows two hypothetical phylogenies that tell different stories. For scenario A (top phylogeny), taxa (a) and (b) do

not appear to be sister taxa, thus indicating two different colonization events and not *in situ* speciation. Conversely, scenario B (bottom phylogeny) shows how (a) and (b) are sister taxa within the same island, and thus *in situ* speciation cannot be ruled out.

A hypothesis applied to oceanic islands is the progression rule hypothesis. This pattern refers to the concordance of island and lineages ages, where older lineages are found in older islands, and younger lineages in younger islands (Funk & Wagner, 1995). For progression to occur, early colonists must arrive at an older island before younger islands are formed, and subsequent colonists occupy newer islands as they are formed. This pattern is based on the foundation of Hennig's (1966) phylogenetic theory and the consequence of the speciation mechanism, where an ancestral species occupies an older geographic area, and a derived species a younger geographic area. Progression rule has been tested in oceanic archipelagos (e.g. Hawai'i, the Australs, the Marquesas, the Galapagos and Canary Islands) as reviewed by Shaw & Gillespie (2016).

1.1.5. Study region

My study region is the Caribbean archipelago. Several definitions have been used to establish biogeographic boundaries to the region (Myers et al., 2000; Smith et al., 2004). I follow the latest definition sensu Smith et al. (2004), which includes the Greater and Lesser Antilles, the Bahama archipelago, and the islands located off the northern coast of Venezuela. Throughout this thesis, I refer to this region as the Antilles, Caribbean, or West Indies, interchangeably. The archipelago lies amidst the Gulf of Mexico, the Caribbean Sea, the North Atlantic Ocean, and is surrounded by continental America.

Despite the relatively small land area of the Caribbean islands combined, the islands support about 11,000 species of plants (Acevedo-Rodríguez & Strong, 2008), that represent about 2% of all vascular plants on Earth (Santiago-Valentín & Olmstead, 2004). Work carried out by Francisco-Ortega et al. (2007) shows an outstanding level of endemism, with 185 genera and 8,000 species respectively. The Caribbean flora comprises a similar number of species as other important island systems of plant diversity, such us Madagascar or New Caledonia (Santiago-Valentín & Olmstead, 2004).

The vast plant diversity in the Caribbean can be explained not only by its proximity to continental America, which might have facilitated successful dispersal, but also by a very complex interaction of geological events that include volcanism, plate tectonic movements, and intervals of island emergence and submergence (Iturralde-Vinent & MacPhee, 1999; Fritsch & McDowell, 2003; Santiago-Valentin & Olmstead, 2004). Moreover, climatic change, through cooling or warming periods (Zachos et al., 2001; Weigelt et al., 2016) have greatly influenced the region since the Cretaceous (Fritsch & McDowell, 2003; Santiago-Valentin & Olmstead, 2004), and have had an impact on major sea-level changes. These sea-level changes have, in turn, an effect on the connectivity between the continent and the islands, creating further opportunities for migration (Weigelt et al., 2016).

Competing hypotheses have long been offered to explain the origins of Caribbean biota. The vicariance hypothesis (Rosen, 1975) proposes that the proto-Antilles, fragments of ancient islands, that were situated between North and South America, carried ancient biota as the Caribbean plate drifted eastward in the late Cretaceous.

Support for this hypothesis comes mainly from animal taxa, as shown by Hedges (2006) in his review of molecular clock analyses for ancient relictual groups.

Prior to acceptance of plate tectonics theory, over-water dispersal dominated the explanations on the origins of Caribbean taxa (Darlington, 1938). This view gained support again in the early 1990s based on vertebrate studies, which used as evidence the taxonomic composition and divergence time estimates between island lineages and their continental relatives (Hedges et al., 1992; Hedges, 2001). Some authors have pointed out the importance and potential role of tropical storms and hurricanes in the dispersion of organisms as flotsam (Hedges, 2001) and propagules (Borhidi, 1991). A recent meta-analysis that included 87 endemic Caribbean lineages showed that South America is a major source for Caribbean biota (Roncal et al., in press). In addition, the clockwise water current coming from the Atlantic that passes through northern South American waters, could have transported ancestors of Caribbean lineages (Hedges, 1996).

Another dispersal hypothesis that has been debated for many years in Caribbean biogeography is the GAARlandia hypothesis. Iturralde-Vinent and MacPhee (1999) proposed that for a period of 2 million years (35 to 33 Ma) a land span connected northern South America with the Greater Antilles. This connection would have facilitated colonization from the mainland into the Greater Antilles. A few molecular studies support GAARlandia as a colonization route for organisms based on divergence times (e.g. toads, Alonso et al., 2012; spiders, Tong et al., 2019; palms, Bacon et al., 2012). However, as pointed out by Ali (2012) stronger geological and palaeo-oceanographical data have not been produced to support this land bridge hypothesis.

1.1.6. Podocarpaceae: a key taxon in biogeography

The Podocarpaceae is the second largest family of conifers (Farjon, 2017), and in terms of morphology and ecology the most diverse (Kelch, 1998). Podocarps mainly have a tropical distribution with very few species representatives of different genera occurring in mountainous regions outside the tropics (e.g. *Podocarpus*, *Nageia*, *Dacrydium*, *Dacrycarpus*). The oldest accepted podocarp macrofossils date from the Middle Triassic of Gondwana (Townrow, 1967) from Natal. Fossil evidence shows that podocarps diversified during the Jurassic and modern genera originated during the Cretaceous (Dettmann, 1994). Despite the occurrence of pollen fossil records (Taggart, 1973) and wood fossils (Castañeda-Posadas et al., 2009) in North America, the family has remained mainly a southern or southern-derived family (Morley, 2011). There are several reasons why Podocarpaceae has been a key taxon to study biogeography. First, its southern hemisphere distribution in all continents, except Antarctica, has traditionally been explained by the breakup of Gondwana, becoming a key example to explain vicariance processes due to plate tectonics. Second, and in contrast to the Nothofagaceae, another key taxon in biogeography, Podocarpaceae extend into tropical and sub-tropical regions. Third, there is a rich fossil record for Podocarpacae, particularly for pollen (Hill & Brodribb, 1999). Because of their success and preference for cool and wet climates, tropical podocarps have been commonly used by palynologists to reconstruct past climates (Coomes & Bellingham, 2011) and have been found to be reliable indicators of past climatic change in Central and South America (Ledru et al., 2007; Cárdenas et al., 2011; Dalling et al., 2011).

In the Neotropics, the family is represented by five genera: *Podocarpus* (ca. 31 species), *Retrophyllum* (2 species), *Prumnopitys* (5 species), and the two monotypic genera *Lepidothamnus* and *Saxegothaea*. The oldest fossil assigned to *Podocarpus* dates from the early Eocene of Patagonia (Wilf et al., 2005; Wilf, 2012; Quiroga et al., 2016). Fossil evidence has also shown the presence of the family in northern South America since at least the late Eocene to the Oligocene (van der Hammen & Hooghiemstra, 2000). Species of *Podocarpus*, the most species-rich genus in the Neotropics within the family, generally have restricted distributions with disjunct populations, has been hypothesized to be an indication of refugial distribution or habitat specialization (Dalling et al., 2011). However, some *Podocarpus* species have a widespread distribution. For example, *P. parlatorei* (from Bolivia to southern Andes) and *P. oleifolius* (northern Andes of Colombia, Bolivia, Peru and Ecuador). Neotropical podocarps predominantly have a montane or lower-montane distribution, with a few exceptions, such as the Central American *P. guatemalensis* found at sea level, or the South American *P. celatus* found at ca. 130 m of altitude in Loreto, Peru.

Podocarpus female cones vary in color from green to red, to purplish. They are fleshy, drupe-like, and contain a single seed, sometimes two (Farjon, 2017). Seeds are probably dispersed by frugivorous birds and small mammals attracted by the color and (or) swollen bract (Mill, 2003a; Enright & Jaffré, 2011).

Systematic studies that include *Podocarpus* species have focused mainly on interfamilial relationships within conifers (Leslie et al., 2012, 2018) or intergeneric relationships within Podocarpaceae (Biffin et al., 2011; Knopf et al., 2012; Little et al., 2013). More recently, Quiroga et al. (2016) used chloroplast and nuclear DNA sequences

together with fossil data to examine the phylogeny and biogeography of Neotropical *Podocarpus*, however relationships amongst the Caribbean species and between continental taxa remain to be clarified, which motivated this thesis.

1.2. Thesis overview

I examine the evolutionary history and assemblage of endemic seed plants of the Caribbean plant taxa at broader and finer taxonomic scales, by means of fieldwork, published DNA sequences, and NGS data. I apply a phylo- biogeographical analysis, with special emphasis on the conifer genus *Podocarpus*.

In **Chapter 2**, I examine the historical biogeography of endemic seed plant genera in the Caribbean. The aim of this chapter was to elucidate the time and area of origin of endemic seed plant genera from the Caribbean. I used published DNA sequences from Zanne et al. (2014) to reconstruct a dated phylogenetic tree and infer the ancestral areas of endemic genera. My study is the most comprehensive study to date that has inferred divergence times, ancestral areas, and potential colonization events in the Caribbean for the largest number of endemic plant genera (32).

In **Chapter 3**, I present the case of Neotropical *Podocarpus* in order to elucidate not only the phylogenetic relationships but also the time and region of origin of Caribbean species. I investigated colonization events from the continent to the Antilles and among the Greater Antillean islands. I also tested the progression rule hypothesis, which has been unexplored for the Caribbean region. In addition, I investigated if island taxa

showed higher diversification rates than their continental relatives, as we would expect from the colonization of new niches in islands.

In **Chapter 4**, I study the evolutionary history of two endemic species of *Podocarpus* on Hispaniola island. I aim to elucidate the role of geological events (e.g. collision of north and south palaeo-islands and marine incursions) and dispersal barriers in the diversification of *Podocarpus* within the island. This contributes to the understanding of intra-island diversification in a geologically complex system.

In **Chapter 5**, I summarize the findings of each chapter and give general conclusions on Caribbean plant biogeography. I also discuss future directions for this line of research in the Caribbean.

1.3. References

Acevedo-Rodríguez P. & Strong M.T. (2008) Floristic richness and affinities in the West Indies. *Botanical Review*, **74**, 5–36.

Ali J.R. (2012) Colonizing the Caribbean: Is the GAARlandia land-bridge hypothesis gaining a foothold? *Journal of Biogeography*, **39**, 431–433.

Alonso R., Crawford A.J., & Bermingham E. (2012) Molecular phylogeny of an endemic radiation of Cuban toads (Bufonidae: *Peltophryne*) based on mitochondrial and nuclear genes. *Journal of Biogeography*, **39**, 434–451.

Antonelli A., Ariza M., Albert J., Andermann T., Azevedo J., Bacon C., Faurby S., Guedes T., Hoorn C., Lohmann L., Matos-Maraví P., Ritter C., Sanmartín I., Silvestro D., Tejedor M., ter Steege H., Tuomisto H., Werneck F., Zizka A., & Edwards S. (2018) Conceptual and empirical advances in Neotropical biodiversity research. *PeerJ Preprints*, **5**, e3074v1.

Bacon C.D., Baker W.J., & Simmons M.P. (2012) Miocene dispersal drives island radiations in the palm tribe Trachycarpeae (Arecaceae). *Systematic Biology*, **61**, 426–442.

Baird N.A., Etter P.D., Atwood T.S., Currey M.C., Shiver A.L., Lewis Z.A., Selker E.U., Cresko W.A., & Johnson E.A. (2008) Rapid SNP Discovery and Genetic Mapping Using Sequenced RAD Markers. *PLOS ONE*, **3**, e3376.

Baldwin B.G., Kyhos D.W., Dvorak J., & Carr G.D. (1991) Chloroplast DNA evidence for a North American origin of the Hawaiian silversword alliance (Asteraceae). *Proceedings of the National Academy of Sciences of the United States of America*, **88**, 1840–1843.

Ball I.R. (1976) Nature and Formulation of Biogeographical Hypotheses. *Systematic Zoology*, **24**, 407–430.

Bell C.D. (2015) Between a Rock and a Hard Place: Applications of the "Molecular Clock" in Systematic Biology. *Systematic Biology*, **40**, 6–13.

Biffin E., Conran J.G., & Lowe A.J. (2011) Podocarp Evolution: A Molecular Phylogenetic Perspective. *Ecology of the Podocarpaceae in Tropical Forests* (ed. by B.L. Turner and L.A. Cernusak), pp. 1–20. Smithsonian Institution Scholarly Press, Washington D.C.

Borhidi A. (1991) *Phytogeography and vegetation ecology of Cuba.* Akadémiai Kiadó, Budapest.

Bremer K. (1992) Ancestral Areas: A Cladistic Reinterpretation of the Center of Origin Concept. *Systematic Biology*, **41**, 436–445.

Buffon G.L.L. de C. (1761) (2004) Excerpts from Histoire Naturelle, Générale et Particulière. Translated into English. *Foundations of Biogeography: Classic Papers with Comments* (ed. by M.V. Lomolino, D.F. Sax, and J.H. Brown), pp. 16–18. The University of Chicago Press.

Cain M.L., Milligan B.G., & Strand A.E. (2000) Long-Distance seed dispersal in plant populations. *American Journal of Botany*, **87**, 1217–1227.

Card D.C., Schield D.R., Adams R.H., Corbin A.B., Perry B.W., Andrew A.L., Pasquesi G.I.M., Smith E.N., Jezkova T., Boback S.M., Booth W., & Castoe T.A. (2016)

Molecular Phylogenetics and Evolution Phylogeographic and population genetic analyses reveal multiple species of *Boa* and independent origins of insular dwarfism. *Molecular Phylogenetics and Evolution*, **102**, 104–116.
Cárdenas M.L., Gosling W.D., Sherlock S.C., Poole I., Pennington R.T., & Mothes P. (2011) The Response of Vegetation on the Andean Flank in Western Amazonia to Pleistocene Climate Change Macarena. *Science*, **331**, 1055–1058.
Carine M.A., Russell S.J., 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.
Carlquist S. (1965) *Island life; a natural history of the islands of the world.* Natural History Press, Garden City, NY.
Carstens B., Lemmon A.R., & Lemmon E.M. (2012) The Promises and Pitfalls of Next-Generation Sequencing Data in Phylogeography. *Systematic Biology*, **61**, 713–715.
Castañeda-Posadas C., Calvillo-Canadell L., & Cevallos-Ferriz S.R.S. (2009) Review of Palaeobotany and Palynology Woods from Miocene sediments in Panotla, Tlaxcala, Mexico. *Review of Palaeobotany and Palynology*, **156**, 494–506.
Clayton J.W., Soltis P.S., & Soltis D.E. (2009) Recent Long-Distance Dispersal Overshadows Ancient Biogeographical Patterns in a Pantropical Angiosperm Family (Simaroubaceae, Sapindales). *Systematic Biology*, **58**, 395–410.
Coomes D.A. & Bellingham P.J. (2011) Temperate and Tropical Podocarps: How Ecologically Alike Are They? *Ecology of the Podocarpaceae in Tropical Forests* (ed. by B. Turner and L.A. Cernusak), pp. 119–140. Smithsonian Institution
Crisci J. V (2001) The voice of Historical Biogeography. *Journal of Biogeography*, **28**, 157–168.
Crisci J. V, Katinas L., & Posadas P. (2003) *Historical Biogeography: An Introduction.* Harvard University Press, Cambridge.
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., Trewick S.A., & Cook L.G. (2011) Hypothesis testing in biogeography. *Trends in Ecology and Evolution*, **26**, 66–72.
Dalling J.W., Barkan P., Bellingham J., Healey J.R., & Tanner E.V.J. (2011) Ecology and Distribution of Neotropical Podocarpaceae. *Ecology of the Podocarpaceae in Tropical Forests* (ed. by B.L. Turner and L.A. Cernusak), pp. 43–56. Smithsonian Institution Scholarly Press, Washington D.C.
Darlington P.J.J. (1938) The origin of the fauna of the Greater Antilles, with discussion of dispersal of animals over water and through the air. *Q Rev Biol*, **13**, 274–300.
Darwin C. (1859) *On the origin of species by means of natural selection, or preservation of favoured races in the struggle for life.* London : John Murray, 1859.
Dettmann M.E. (1994) Cretaceous vegetation: The microfossil record. *History of the Australian Vegetation* (ed. by R.S. Hill), pp. 143–170. South Australia: University of Adelaide Press.
Dupuis J.R., Brunet B.M.T., Bird H.M., Lumley L.M., Fagua G., Boyle B., Levesque R., Cusson M., Powell J.A., & Sperling F.A.H. (2017) Molecular Phylogenetics and Evolution Genome-wide SNPs resolve phylogenetic relationships in the North

American spruce budworm (*Choristoneura fumiferana*) species complex. *Molecular Phylogenetics and Evolution*, **111**, 158–168.
Eaton D.A.R. & Ree R.H. (2013) Inferring Phylogeny and Introgression using RADseq Data: An Example from Flowering Plants (Pedicularis: Orobanchaceae). *Systematic Biol*, **62**, 689–706.
Ebach M.C., Humphries C.J., & Williams D.M. (2003) Phylogenetic biogeography deconstructed. *Journal of Biogeography*, **30**, 1285–1296.
Elshire R.J., Glaubitz J.C., Sun Q., Poland J.A., Kawamoto K., Buckler E.S., & Mitchell S.E. (2011) A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species. *PLoS ONE*, **6**, e19379.
Emerson B.C. (2002) Evolution on oceanic islands: molecular phylogenetic approaches to understanding pattern and process. *Molecular Ecology*, **11**, 951–966.
Enright N.J. & Jaffré T. (2011) Ecology and Distribution of the Malesian Podocarps. *Ecology of the Podocarpaceae in Tropical Forests* (ed. by B.L. Turner and L.A. Cernusak), pp. 57–77. Smithsonian Institution Scholarly Press, Washington D.C.
Farjon A. (2017) *A handbook of the world's conifers.* Brill, Leiden.
Fitch W.M. (1971) Toward Defining the Course of Evolution: Minimum Change for a Specific Tree Topology. *Systematic Zoology*, **20**, 406–416.
Forster J.R. (1778. (1778) Observations made during a Voyage Round the World. *Foundations of Biogeography: Classic Papers with CommentsClassic Papers with Comments* (ed. by M.V. Lomolino, D.F. Sax, and J.H. Brown), pp. 19–27. The University of Chicago Press.
Francisco-Ortega J., Santiago-Valentín E., Acevedo-Rodríguez P., Lewis C., Pipoly III J., Meerow A.W., & Maunder M. (2007) Seed plant genera endemic to the Caribbean Island Biodiversity Hotspot: A review and a molecular phylogenetic perspective. *The Botanical Review*, **73**, 183–234.
Fritsch P.W. & McDowell T.D. (2003) Biogeography and phylogeny of Caribbean Plants-Introduction. *Systematic Botany*, **28**, 376–377.
Funk V.A. & Wagner W.L. (1995) Biogeography Patterns in the Hawaiian Islands. *Hawaiian Biogeography: Evolution on a hot spot archipelago* (ed. by W.L. Wagner and V.A. Funk), pp. 379–419. Smithsonian Institution Institution Press,
Gillespie R.G., Baldwin B.G., Waters J.M., Fraser C.I., Nikula R., & Roderick G.K. (2012) Long-distance dispersal: a framework for hypothesis testing. *Trends in Ecology & Evolution*, **27**, 47–56.
van der Hammen T. & Hooghiemstra H. (2000) Neogene and Quaternary history of vegetation, climate, and plant diversity in Amazonia. *Quaternary Science Review*, **19**, 725–742.
Hammond A.L. (1971) Plate Tectonics: The Geophysics of the Earth's Surface. *Science*, **173**, 40 LP-41.
Hamon P., Grover C.E., Davis A.P., Rakotomalala J., Raharimalala N.E., Albert V.A., Sreenath H.L., Stoffelen P., Mitchell S.E., Couturon E., Hamon S., De Kochko A., Crouzillat D., Rigoreau M., Sumirat U., Akaffou S., & Guyot R. (2017) Genotyping-by-sequencing provides the first well-resolved phylogeny for coffee (*Coffea*) and insights into the evolution of caffeine content in its species GBS coffee phylogeny and the evolution of caffeine content. *Molecular Phylogenetics and Evolution*, **109**,

351–361.
Hedges S.B. (1996) Historical Biogeography of West Indian Vertebrates. *Annual Review of Ecology and Systematics*, **27**, 163–196.
Hedges S.B. (2001) *Biogeography of the West Indies: An overview.* CRC Press, Boca Raton.
Hedges S.B. (2006) Paleogeography of the Antilles and origin of West Indian terrestrial vertebrates. *Annals of the Missouri Botanical Garden*, **93**, 231–244.
Hedges S.B., Hass C.A., & Maxson L.R. (1992) Caribbean biogeography: Molecular evidence for dispersal in West Indian terrestrial vertebrates. *Proceedings of the National Academy of Sciences of the United States of America*, **89**, 1909–1913.
Hennig W. (1966) *Phylogenetic Systematics.* University of Illinois Press.
Hill R.S. & Brodribb T.J. (1999) Southern Conifers in Time and Space. *Australian Journal of Botany*, **47**, 639–696.
Ho S.Y.W. (2014) The changing face of the molecular evolutionary clock. *Trends in Ecology & Evolution*, **29**, 496–503.
von Humboldt A. (1905) Essai Élémentaire de Géographie Botanique. Translated into English. *Foundations of Biogeography: Classic Papers with Comments* (ed. by M.V. Lomolino, D.F. Sax, and J.H. Brown), pp. 49–57. The University of Chicago Press.
Iturralde-Vinent M. & MacPhee R.D.. (1999) Paleogeography of the Caribbean Region: Implications for Cenozoic Biogeography. *Bulletin of the American Museum of Natural History*, **238**, 1–95.
Jorgensen T.H. & Olesen J.M. (2001) Adaptive radiation of island plants: evidence from Aeonium (Crassulaceae) of the Canary Islands. *Perspectives in Plant Ecology Evolution and Systematics*, **4**, 29–42.
Kelch D.G. (1998) Phylogeny of Podocarpaceae: Comparison of evidence from morphology and 18s rDNA. *American Journal of Botany*, **85**, 986–996.
Klimova A., Ortega-Rubio A., Vendrami D.L.J., & Hoffman J.I. (2018) Genotyping by sequencing reveals contrasting patterns of population structure, ecologically mediated divergence, and long-distance dispersal in North American palms. *Ecology and Evolution*, **8**, 5873– 5890.
Knopf P., Schulz C., Little D.P., Stutzel T., & Stevenson D.W. (2012) Relationships within Podocarpaceae based on DNA sequence, anatomical, morphological, and biogeographical data. *Cladistics*, **28**, 271–299.
Lamm K.S. & Redelings B.D. (2009) Reconstructing ancestral ranges in historical biogeography: properties and prospects. *Journal of Systematics and Evolution*, **47**, 369–382.
Leaché A.D. & Oaks J.R. (2017) The Utility of Single Nucleotide Polymorphism (SNP) Data in Phylogenetics. *Annual Review of Ecology and Systematics*, **48**, 69–84.
Ledru M.-P., Luisa M., Salatino F., Ceccantini G., Salatino A., Pinheiro F., & Pintaud J. (2007) Regional assessment of the impact of climatic change on the distribution of a tropical conifer in the lowlands of South America. *Diversity and Distributions*, **13**, 761–771.
Leslie A.B., Beaulieu J., Holman G., Campbell C.S., Mei W., Raubeson L.R., & Mathews S. (2018) An overview of extant conifer evolution from the perspective of the fossil record -. *American Journal of Botany*, **105**, 1–14.

Leslie A.B., Beaulieu J.M., Rai H.S., Crane P.R., & Donoghue M.J. (2012) Hemisphere-scale differences in conifer evolutionary dynamics. *Proceedings of the National Academy of Sciences*, **109**, 16217–16221.
Little D.P., Knopf P., & Schulz C. (2013) DNA Barcode Identification of Podocarpaceae — The Second Largest Conifer Family. *PLoS ONE*, **8**, e81008.
Lomolino M.V., Riddle B.R., Whittaker R.J., & Brown J.H. (2010) *Biogeography.* Sinauer Associates, Inc., Sunderland, Massachusetts.
Losos J.B. & Ricklefs R.E. (2009) Adaptation and diversification on islands. *Nature*, **457**, 830–836.
MacArthur R.H. & Wilson E.O. (1967) *The theory of island biogeography.* Princeton University Press, Princeton, N.J.
Maddison W.P. & Knowles L.L. (2006) Inferring Phylogeny Despite Incomplete Lineage Sorting. *Systematic Biology*, **55**, 21–30.
Magallón S.A. (2004) Dating Lineages: Molecular and Paleontological Approaches to the Temporal Framework of Clades. *International Journal of Plant Sciences*, **165**, S7–S21.
Mardis E.R. (2008) The impact of next-generation sequencing technology on genetics. *Trends in Genetics*, **24**, 133–141.
McCormack J.E., Hird S.M., Zellmer A.J., Carstens B.C., & Brumfield R.T. (2013) Applications of next-generation sequencing to phylogeography and phylogenetics. *Molecular Phylogenetics and Evolution*, **66**, 526–538.
Mill R.R. (2003a) Towards a Biogeography of the Podocarpaceae. *Acta Horticulturae*, **615**, 137–147.
Mill R.R. (2003b) Towards a Biogeography of the Podocarpaceae. *Acta Hortic.*, **65**, 137–147.
Morley R.J. (2011) Dispersal and Paleoecology of Tropical Podocarps. *Ecology of the Podocarpaceae in Tropical Forests* (ed. by B.L. Turner and L.A. Cernusak), pp. 21–41. Smithsonian Institution Scholarly Press, Washington D.C.
Morrone J.J. & Crisci J. V (1995) HISTORICAL BIOGEOGRAPHY: Introduction to Methods. *Annual Review of Ecology and Systematics*, **26**, 373–401.
Myers N., Mittermeier R.A., Mittermeier C.G., Fonseca G.A.B., & Kent J. (2000) Biodiversity hotspots for conservation priorities. *Nature*, **403**, 853–858.
Nathan R. (2008) Long-Distance Dispersal of Plants. *Science*, **313**, 786–788.
Nathan R., Perry G., Cronin J.T., Strand A.E., & Cain M.L. (2003) Methods for estimating long-distance dispersal. *Oikos*, **103**, 261–273.
Patiño J.P., Carine M., Mardulyn P., Devos N., Mateo R., González-Mancebo J.M., Shaw A.J., & Vanderpoorten A. (2015) Approximate Bayesian computation reveals the rucial role of oceanic islands for the assembly of continental biodiversity. *Systematic Biology*, **64**, 579–589.
Penny D. & Phillips M.J. (2004) The rise of birds and mammals: are microevolutionary processes sufficient for macroevolution? *Trends in Ecology and Evolution*, **19**, 516–522.
Pirie M.D., Humphreys A.M., Antonelli A., Galley C., & Linder H.P. (2012) Model uncertainty in ancestral area reconstruction: A parsimonious solution ? *TAXON*, **61**, 652–664.

Pirie M.D., Litsios G., Bellstedt D.U., Salamin N., & Kissling J. (2015) Back to Gondwanaland: can ancient vicariance explain (some) Indian Ocean disjunct plant distributions? *Biology Letters*, **11**, 20150086.
Platnick N.I. & Nelson G. (1978) A Method of Analysis for Historical Biogeography. *Sistematic Zoology*, **27**, 1–16.
Posadas P., Crisci J. V, & Katinas L. (2006) Historical biogeography: A review of its basic concepts and critical issues. *Journal of Arid Environments*, **66**, 389–403.
Presgraves D.C. & Glor R.E. (2009) Evolutionary Biology: Speciation on Islands. *Current Biology*, **20**, 440–442.
Queiroz A. De (2005) The resurrection of oceanic dispersal in historical biogeography. *Trends in Ecology and Evolution*, **20**, 68–73.
Quiroga M.P., Mathiasen P., Iglesias A., Mill R.R., & Premoli A.C. (2016) Molecular and fossil evidence disentangle the biogeographical history of *Podocarpus*, a key genus in plant geography. *Journal of Biogeography*, **43**, 372–383.
Raven P.H. & Axelrod D.I. (1974) Angiosperm Biogeography and Past Continental Movements. *Annals of the Missouri Botanical Garden*, **61**, 539–673.
Ree R.H., Moore B.R., Webb C.O., & Donoghue M.J. (2005) A likelihood framework for inferring the evolution of geographic range on phylogenetic trees. *Evolution*, **59**, 2299–2311.
Ree R.H. & Smith S.A. (2008) Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. *Systematic Biology*, **57**, 4–14.
Ricklefs R. & Bermingham E. (2008) The West Indies as a laboratory of biogeography and evolution. *Philosophical Transactions of the Royal Society B*, **363**, 2393–2413.
Riddle B.R., Dawson M.N., Hadly A., Hafner D.J., Hickerson M.J., Mantooth S.J., & Yoder A.D. (2008) The role of molecular genetics in sculpting the future of integrative biogeography. *Progress in Physical Geography*, **32**, 173–202.
Ronquist F. (1997a) Dispersal-Vicariance Analysis: A new approach to the quantification of Historical Biogeography. *Systematic Biology*, **46**, 195–203.
Ronquist F. (1997b) Dispersal-Vicariance Analysis: A New Approach to the Quantification of Historical Biogeography. *Systematic Biology*, **46**, 195–203.
Ronquist F. & Sanmartín I. (2011) Phylogenetic Methods in Biogeography. *Annual Review of Ecology, Evolution and Systematics*, **42**, 441–464.
Rosen D.E. (1975) A Vicariance Model of Caribbean Biogeography. *Systematic Zoology*, **24**, 431–464.
Rosen D.E. (1978) Vicariant Patterns and Historical Explanation in Biogeography. *Systematic Zoology*, **27**, 159–188.
Rubin B.E.R., Ree R.H., & Moreau C.S. (2012) Inferring Phylogenies from RAD Sequence Data. *PLOS ONE*, **7**, e33394.
Sanderson M.J. (1997) A Nonparametric Approach to Estimating Divergence Times in the Absence of Rate Constancy. *Molecular Biology and Evolution*, **14**, 1218–1231.
Santiago-Valentín E. & Olmstead R.G. (2004) Historical Biogeography of Caribbean Plants: Introduction to current knowledge and possibilities from a phylogenetic perspective. *Taxon*, **53**, 299–319.
Santos C.M.D. & Amorim D.S. (2007) Why biogeographical hypotheses need a well

supported phylogenetic framework: A conceptual evaluation. *Papéis Avulsos de Zoologia*, **47**, 63–73.
Sauquet H. (2013) A practical guide to molecular dating Guide pratique de la datation moléculaire. *Comptes rendus Palevol*, **12**, 355–367.
archipelagos: Hotspots for inferences of evolutionary process. *Proceedings of the National Academy of Sciences*, **113**, 7986–7993.
Smith M.L., Hedges S.B., Buck W., Hemphill A., Inchaustegui S., Ivie M.A., Martina D., Francisco-Ortega J., & Maunder M. (2004) Caribbean Islands. *Hotspots revisited: Earth's biologically richest and most threatened terrestrial ecoregions* (ed. by R.A. Mittermeier, R.R. Gil, H. M., J. Pilgrim, T. Brooks, C.G. Mittermeier, J.F. Lamoreux, and G.A.B. da Fonseca), CEMEX, Mexico, D.F.
Stuessy T.F., Takayama K., López-Sepúlveda P., & Crawford D.J. (2014) Interpretation of patterns of genetic variation in endemic plant species of oceanic islands. *Botanical Journal of the Linnean Society*, **174**, 276–288.
Taggart R.E. (1973) Additions to the Miocene Sucker Creek Flora of Oregon and Idaho. *American Journal of Botany*, **60**, 923–928.
Tong Y., Binford G., Rheims C.A., & Kuntner M. (2019) Huntsmen of the Caribbean: Multiple tests of the GAARlandia hypothesis. *Molecular Phylogenetics and Evolution*, **130**, 259–268.
Townrow J.A. (1967) On *Rissikia* and *Mataia* podocarpaceous conifers from the lower Mesozoic of Southern Lands. *Papers and Proceedings of the Royal Society of Tasmania*, **101**, 103–136.
Wallace A.R. (1876) *The geographical distribution of animals. With a study of the relations of living and extinct faunas as elucidating the past changes of the earth's surface.* . Harper and brothers, New York.
Wallace A.R. (1892) *Island life: or, the phenomena and causes of insular faunas and floras, including a revision and attempted solution of the problem of geological climates.* London: Macmillan.
Weigelt P., Steinbauer M.J., Cabral J.S., & Kreft H. (2016) Late Quaternary climate change shapes island biodiversity. *Nature*, **532**, 99–102.
Wilf P. (2012) Rainforest conifers of Eocene Patagonia: attached cones and foliage of the extant southeast Asian and Australasian genus *Dacrycarpus* (Podocarpaceae). *American Journal of Botany*, **99**, 562–584.
Wilf P., Johnson K.R., Cúneo N.R., Smith M.E., Singer B.S., & Gandolfo M.A. (2005) Eocene Plant Diversity at Laguna del Hunco and Río Pichileufú, Patagonia, Argentina. *The American Naturalist*, **165**, 634–650.
Yuan Y.-M., Wohlhauser S., Möller M., Klackenberg J., Callmander M.W., & Küpfer P. (2005) Phylogeny and Biogeography of *Exacum* (Gentianaceae): A Disjunctive Distribution in the Indian Ocean Basin Resulted from Long Distance Dispersal and Extensive Radiation. *Systematic Biology*, **54**, 21–34.
Zachos J., Pagani M., Sloan L., Thomas E., & Billups K. (2001) Trends, rhythms, and aberrations in global climate 65 Ma to present. *Science*, **292**, 686–693.
Zuckerkandl E. & Pauling L. (1962) Molecular Disease, Evolution, and Genic Heterogeneity. *Horizons in Biochemistry* (ed. by M. Kasha and B. Pullman), pp. 189–225. Academic Press, New York.

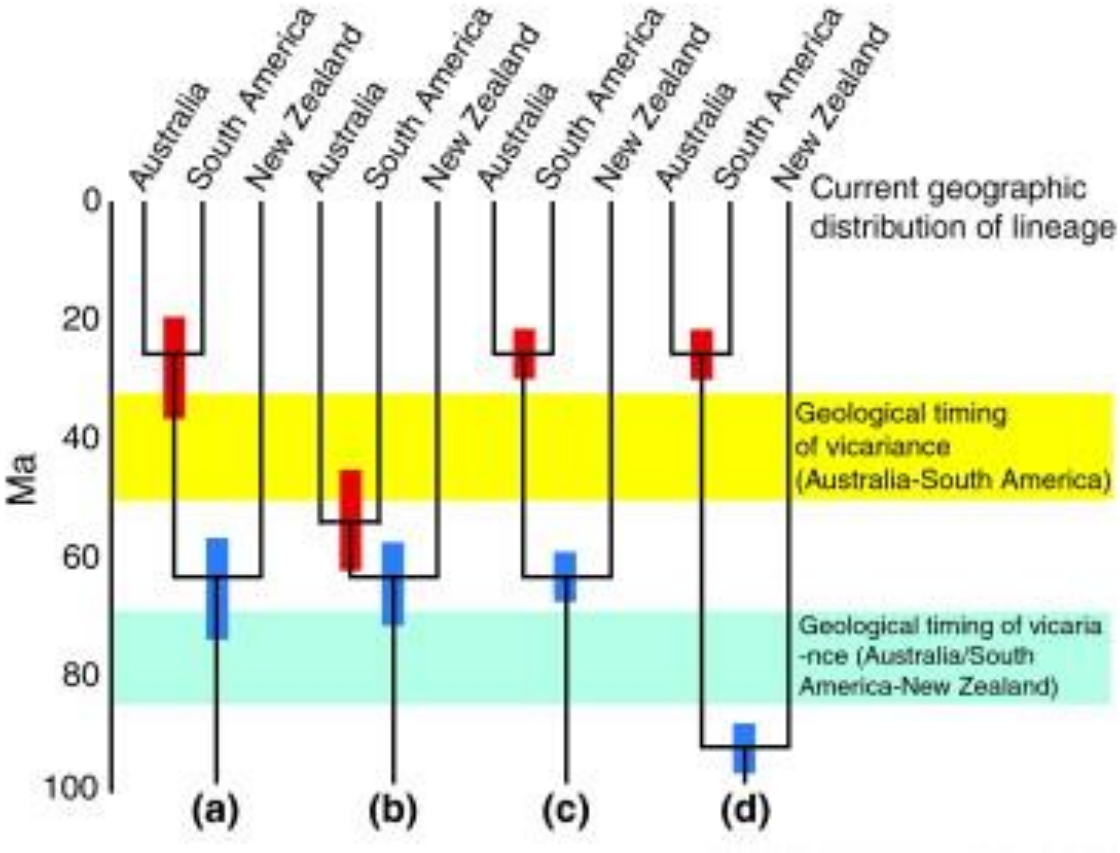


Figure 1.1. Four hypothesized scenarios of lineage divergence in the Austral and Neotropical zones. In (a) and (b) all confidence intervals fall within the timeframes of the two vicariance events. In (c) and (d) none of the confidence intervals fall within the timeframes for the two vicariance events, thus the vicariance hypothesis explaining lineage divergence should be rejected. Figure from Crisp et al. (2011).

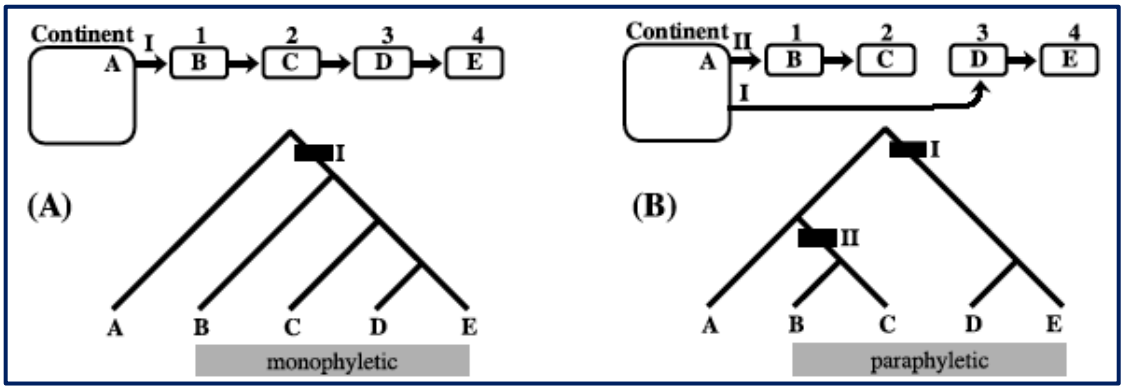


Figure 1.2. Interpretation of monophyly and paraphyly/polyphyly in a colonization context. Scenario (A) illustrates a single colonization event (denoted by black horizontal bar), thus a molecular phylogenetic reconstruction will result in a monophyletic clustering of island taxa. Scenario (B) illustrates two independent colonization events, thus a molecular phylogenetic reconstruction will result on a paraphyletic or polyphyletic clustering of island taxa. Figure from Emerson (2002).

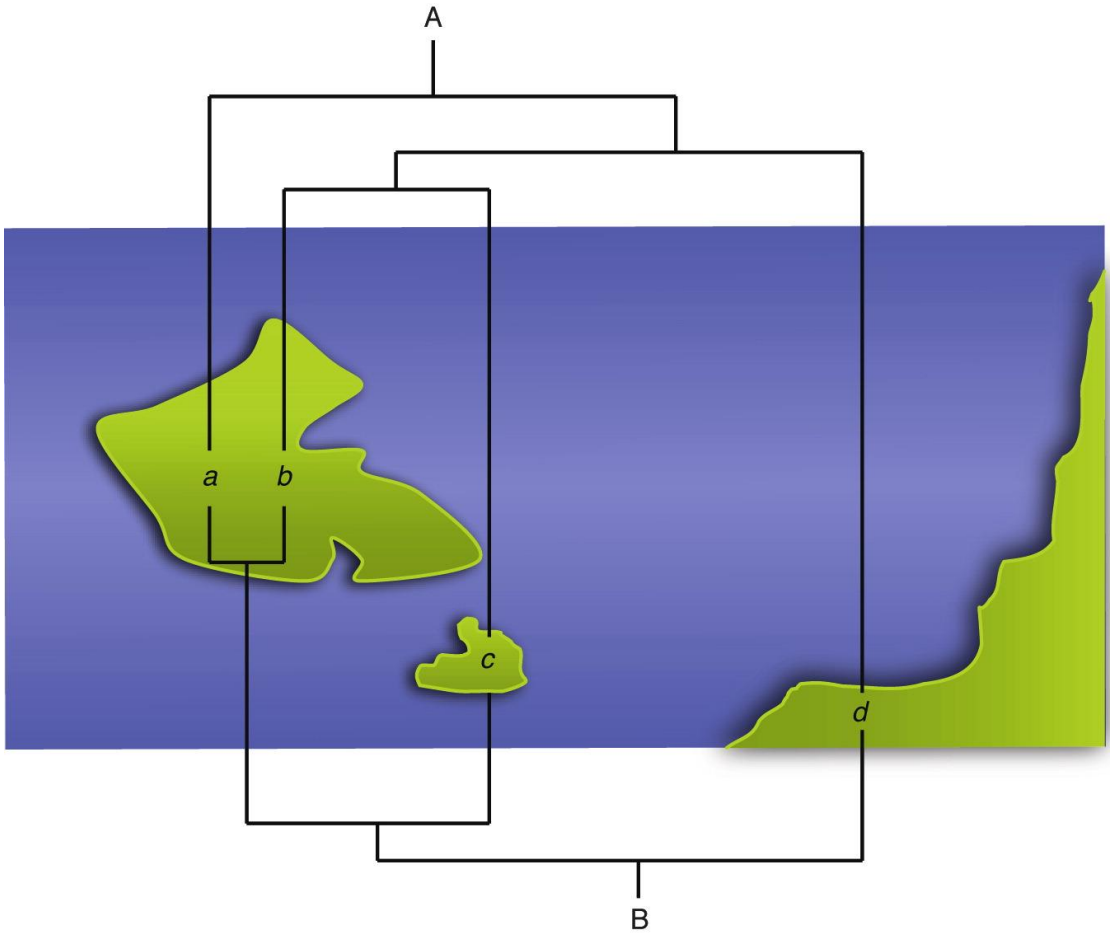


Figure 1.3. Alternative scenarios for island colonization. In Scenario B, species ***a*** and ***b*** are sister taxa, and along with species ***c*** part of a monophyletic island colonization. In Scenario A, island colonization is polyphyletic, with species ***a*** and ***b*** in separate lineages. Note that the latter is more closely related to species ***c*** on the other island. Figure from Presgraves & Glor (2010).

Co-authorship statement

The thesis is the result of my independent research, under the supervision of Julissa Roncal and co-supervision of Alexandre Antonelli. Contributions by each author for each chapter is as follows:

Chapter 2: Historical Biogeography of endemic seed plant genera in the Caribbean: Did GAARlandia play a role?

This chapter is a version of a manuscript published in the journal *Ecology & Evolution* (Nieto-Blázquez, M.E., Antonelli, A., Roncal, J., 2017. Historical biogeography of endemic seed plant genera in the Caribbean: Did GAARlandia play a role? Ecol. Evol. 7, 10158–10174. https://doi.org/10.1002/ece3.3521). M.E. Nieto-Blázquez, A. Antonelli and J. Roncal designed the study. M.E. Nieto-Blázquez compiled and analyzed the data. M.E. Nieto-Blázquez wrote the article with contributions from J. Roncal and A. Antonelli.

Chapter 3: Historical Biogeography of the conifer genus *Podocarpus* in the Caribbean

This chapter is a version of a manuscript in preparation for the *Journal of Biogeography*. M.E. Nieto-Blázquez and J. Roncal designed the study. M.E. Nieto-Blázquez compiled and analyzed the data. Lourdes Peña Castillo advised on the preparation of data and use of NGS pipeline. M.E. Nieto-Blázquez wrote the chapter with contributions from J. Roncal.

Chapter 4: Evolutionary history of *Podocarpus* in Hispaniola

This chapter is a version of a manuscript to be submitted to *Molecular Phylogenetics and Evolution*. Contributions to this chapter were made by M.E. Nieto-Blázquez and J. Roncal. M.E. Nieto-Blázquez and J. Roncal designed the study. M.E. Nieto-Blázquez compiled and analyzed the data. M.E. Nieto-Blázquez wrote the chapter with contributions from J. Roncal.

CHAPTER 2

Historical Biogeography of endemic seed plant genera in the Caribbean: Did GAARlandia play a role?

This manuscript can be found in *Ecology and Evolution*:

Nieto-Blázquez, M. E., Antonelli, A., & Roncal, J. (2017). Historical biogeography of endemic seed plant genera in the Caribbean: Did GAARlandia play a role? *Ecology and Evolution*, 7, 10158–10174. https://doi.org/10.1002/ece3.3521

2.1. Abstract

The Caribbean archipelago is a region with an extremely complex geological history and an outstanding plant diversity with high levels of endemism. The aim of this study was to better understand the historical assembly and evolution of endemic seed plant genera in the Caribbean, by first determining divergence times of endemic genera to test whether the hypothesized Greater Antilles and Aves Ridge (GAARlandia) land bridge played a role in the archipelago colonization and second by testing South America as the main colonization source as expected by the position of landmasses and recent evidence of an asymmetrical biotic interchange. I reconstructed a dated molecular phylogenetic tree for 625 seed plants including 32 Caribbean endemic genera using Bayesian inference and ten calibrations. To estimate the geographic range of the ancestors of endemic genera, I performed a model selection between a null and two complex biogeographic models that included timeframes based on geological information, dispersal probabilities, and directionality among regions. Crown ages for endemic genera ranged from early Eocene (53.1 Ma) to late Pliocene (3.4 Ma). Confidence intervals for divergence times (crown and/or stem ages) of 22 endemic genera occurred within the GAARlandia time frame. Contrary to expectations, the Antilles appears as the main ancestral area for endemic seed plant genera and only five genera had a South American origin. In contrast to patterns shown for vertebrates and other organisms and based on my sampling, I conclude that GAARlandia did not act as a colonization route for plants between South America and the Antilles. Further studies on Caribbean plant dispersal at the species and population levels will be required to reveal finer-scale biogeographic patterns and mechanisms.

2.2. Introduction

Although islands cover only about 5% of the Earth's surface, they contain about a quarter of all terrestrial plant species (Caujapé-Castells, 2011). Explaining the high biological diversity and endemicity of islands has been a topic of study in the last three centuries, pioneered by Darwin (1876) and Wallace (1892). The structure of insular communities is the result of the interaction among three fundamental biological processes: immigration, speciation, and extinction (Whittaker et al., 2008, 2017; Lomolino et al., 2010). According to the new synthesis in island biogeography theory (Lomolino et al., 2010), these three fundamental processes are scale dependent affecting different levels of biological organization from individuals, to populations or communities, and biotas. Evolutionary and geological dynamics have been identified to affect the biotic level of insular organization (Haila, 1990).

The Caribbean archipelago (i.e., Greater and Lesser Antilles, and the Bahamas) is one of the world's 34 biodiversity hotspots (Mittermeier et al., 2004) and represents the most important insular system in the Neotropics (Maunder et al., 2008). Despite the relatively small land area of this archipelago, there are nearly 13,000 seed plant species, of which almost 8,000 are endemic (Acevedo-Rodríguez & Strong, 2008). This alpha-diversity is similar to that of Madagascar, and three times larger than that of New Caledonia (Myers et al., 2000). There are 180 seed plant genera endemic to the Caribbean (Francisco-Ortega et al., 2007) which represents 13.2% of the total number of genera on the islands, and 86 of the 180 endemic genera (47.7%) are monotypic. Endemic genera are concentrated in the Greater Antilles, especially in Cuba and Hispaniola, the largest and most heterogeneous islands (Santiago-Valentín & Olmstead, 2004).

The vast flora diversity in the Caribbean can be explained not only by its proximity to the American continent, which might have facilitated successful dispersal from an outstandingly rich biota, but also by the very complex interaction of geological events, which include volcanism, plate tectonic movements, and intervals of island emergence and submergence (Iturralde-Vinent & MacPhee, 1999; Graham, 2003). Moreover, climatic change, through cooling or warming periods (Zachos et al., 2001), has greatly influenced the region since the Cretaceous (Fritsch & McDowell, 2003) and has had an impact on major sea-level changes. These sea-level changes had in turn an effect on the connectivity between the continent and the islands, creating further migration opportunities (Weigelt et al., 2016). The geological history of the Greater and Lesser Antilles are quite distinct from one another, and the main sequence of events is described in detail in Graham (2003). The Greater Antilles originated in the Cretaceous [c. 130 Million years ago (Ma)], forming a volcanic chain of sea mountains between North and South America (Pindell & Kennan, 2009). This chain of islands, known as Proto-Antilles, moved northeastward until they collided first with the Yucatan Peninsula (c. 84 Ma) and then with the Bahamas Platform in the early Eocene (c. 56 Ma). The Lesser Antilles were formed subsequently between the middle Eocene (c. 47–38 Ma, in the north) and the Oligocene (c. 34–23 Ma, in the south) as a result of the subduction of the South American Plate under the Caribbean Plate. By the middle Eocene (c. 49 Ma), most of the Greater and Lesser Antilles were above water. This geological activity for the last 100 Ma (Burke, 1988) might have presented significant opportunities for speciation, colonization, and vicariance (Hedges, 2001).

In 1999, Iturralde-Vinent and McPhee introduced a controversial hypothesis, the "GAARlandia (Greater Antilles + Aves Ridge) land bridge." They proposed that colonization of the Antilles was possible from northeast South America through a quasi-continuous land bridge or island chain that lasted for a period of 1–2 Ma, close to the Eocene–Oligocene boundary, c. 34 Ma. The Eocene–Oligocene boundary coincides with a major drop in temperature and sea level that might have affected connectivity between regions exposing land areas (Hedges, 2001). Studies that support the colonization role of GAARlandia are primarily based on molecular dating estimates, and comprise amphibians (Alonso et al., 2012), invertebrates (Binford et al., 2008; Chamberland et al., 2018; Tong et al., 2019), vertebrates (Hulsey et al., 2011), and also plants, as shown for the genus *Styrax* (Styracaceae, (Fritsch, 2003), *Moacroton* (Euphorbiaceae, Van Ee et al., 2008), and *Copernicia* (Arecaceae, Bacon et al., 2012). Despite this evidence, the existence of GAARlandia is still a debatable hypothesis to explain lineage colonization and diversification in the Caribbean (Ali, 2012), due to limited geological and paleoceanographical evidence supporting its existence and because molecular and biogeographic evidence is still incomplete for the Caribbean biota.

While floristic studies have shown strong links between the Caribbean flora and that of the surrounding continental landmasses (Acevedo-Rodríguez & Strong, 2008), little is known regarding the precise timing and geographic origin of the flora as a whole. Most insight on Caribbean historical biogeography results from molecular phylogenies of vertebrates (Dávalos, 2004; Hedges, 2006; Hulsey et al., 2011; Monceau et al., 2013), which suggest a combination of dispersal and vicariance for the Antillean fauna. North (NA) and Central America (CA) have been identified as colonization sources for active

dispersers, such as birds, bats, and freshwater fishes (Hedges, 1996, 2006) into the Caribbean region. In contrast, South America (SA) has been suggested as the main source for passive dispersers (nonvolant fauna), which would require floating mechanisms (Hedges, 1996, 2006), and for vertebrates using potential land bridges for island colonization (Alonso et al., 2012; Dávalos, 2004).

Francisco-Ortega et al. (2007) provided a checklist of Caribbean endemic seed plant genera and a review of molecular phylogenetic studies of these plants. Their review highlighted that DNA phylogenies were available for only 35% of the Antillean genera. Since then, several molecular phylogenies that include Caribbean endemic genera have been published (e.g. Jestrow et al., 2010, 2012a,b; Appelhans et al., 2012), revealing a complex biogeographic history (Roncal et al., 2008). Some endemic genera have sister taxa that are widely distributed in continental America (Lavin et al., 2001a, 2003; Rova et al., 2002; Wurdack et al., 2005), others have relatives with a more restricted continental distribution (Lavin et al., 2001a; Baldwin et al., 2002; Wojciechowski et al., 2004), and a few are sister to taxa that are native to regions outside the Neotropics, such as Africa (Lavin et al., 2001a, 2001b), Polynesia (Kimball & Crawford, 2004), and New Caledonia (Motley et al., 2005).

With the aim of providing insights into the origin and evolution of the Caribbean flora, I targeted endemic seed plant genera. I focused on genera because most plant phylogenies are still poorly sampled at the species level, rendering the inference of range evolution problematic and biased by the inclusion of common, widespread species with island and continental distributions, and fewer island endemics. It was also beyond the scope of this study to analyze the biogeographic history of individual endemic species

within non endemic genera. Even though higher taxa (e.g., genera and families) may not be as intercomparable as biological species, processes normally considered in the context of speciation like divergent selection and geographic isolation can generate evolutionary significant units above the species level (Barraclough, 2010; Barraclough & Humphreys, 2015). Plant genera can therefore also be used as units of biodiversity.

I reconstructed a dated phylogenetic tree and tested different biogeographic scenarios to address the following questions: (1) When did endemic seed plant genera diverge from their sister taxa, and (2) what were the most likely regions that ancestors of endemic genera occupied? My hypotheses are (1) GAARlandia played a major role as a migration route in the colonization of the Caribbean Islands. Under this hypothesis, I expect to find the origin of endemic genera (i.e., mean stem to crown ages) contemporaneous with the hypothesized presence of GAARlandia. (2) Endemic genera descended from South American ancestors because of their proximity to GAARlandia, which facilitated colonization from SA more than from CA or NA, and considering the asymmetry in dispersal or migration directionality during large part of the Neogene observed in birds, plants and mammals (Weir et al., 2009; Bacon et al., 2015). Through a taxon sampling of 32 endemic seed plant genera, this study provides a comprehensive evolutionary and biogeographic framework to understand the historical assembly of the Caribbean flora at the genus level.

2.3. Material and Methods

2.3.1. Taxon sampling selection

I searched for sequences from all endemic plant genera following the compilation by Francisco-Ortega et al. (2007) on the data matrix of Zanne et al. (2014) who reconstructed a dated phylogeny for 32,223 plant species. Zanne et al. (2014) used the International Plant Names Index (IPNI), Tropicos, The Plant List and Angiosperm Phylogeny Group (APG) to verify taxonomic nomenclature. I found 56 species within 41 endemic genera in Zanne et al. (2014). Of these 41 endemic genera, 33 are included in the 35% of Caribbean endemic genera included in molecular phylogenies stated by Francisco-Ortega et al. (2007). Therefore, I have included 52% endemic genera for which there were molecular phylogenies available at the time of the publication. I used the NCBI taxonomy facility (Federhen, 2012) to select up to 10 species for every genus within the suprageneric rank to which the endemic genera belong. When genera contained more than 10 species, I selected species that represented the entire distributional range of the genus, and with complete sequences available in the Zanne et al. (2014) matrix.

2.3.2. DNA sequence selection and alignment

Of the seven gene regions available in Zanne et al. (2014), I selected four (18S rDNA, *atpB*, *matK*, and *rbcL*) for my alignment. I excluded the 26S rDNA region because it was not well represented (only 17 sequences were available for my taxon sampling). The ITS and *trnL-trnF* gene regions were available for a fair number of species (511 and 594, respectively) but were also excluded because sequences were difficult to align and DNA homology could not be confirmed. Each of the four-gene regions was aligned

independently using MAFFT v. 7.187 on XSEDE (Katoh & Standley, 2013) via the CIPRES Science Gateway (Miller et al., 2010). Manual trimming and concatenation of gene regions were performed in GENEIOUS v. 7.1.9 (Kearse et al., 2012). The final four-gene concatenated matrix had a total length of 5,462 bp and contained 625 seed plant species (Spermatophyta) within 319 genera in 20 families, including 41 Caribbean endemic genera (Table 2.1). I had 37% missing nucleotide data in this final alignment.

2.3.3. Phylogenetic reconstruction and dating

I performed tree searches using the four-gene concatenated matrix under a maximum likelihood (ML) approach. Phylogeny reconstruction was performed on RAxML-HPC2 version 8.2.8 (Stamatakis, 2006) via the CIPRES Science Gateway using the rapid bootstrap algorithm with 500 replicates. I selected six gymnosperms in the Zamiaceae family to root the tree, which included the monotypic endemic Caribbean genus *Microcycas* and five *Zamia* species. I used JMODELTEST2 v.0.1.1 (Darriba et al., 2012) via the CIPRES Science Gateway to select the best nucleotide substitution model for the four-gene alignment under the Akaike information criterion (AIC, Akaike, 1974). The best-fit model was GTR + I + Γ, which was selected for subsequent analyses.

In order to estimate absolute divergence times, I inferred a time-calibrated phylogenetic tree using a Bayesian inference (BI) approach as implemented in BEAST v2.3.1 (Bouckaert et al., 2014). Analysis on the concatenated matrix used the uncorrelated lognormal (UCLN) relaxed clock (Drummond et al., 2006). The tree prior was set to the Yule model, which models a constant lineage birth rate for each branch in the tree. Ten calibration points were applied to the dating analysis. In order to avoid overestimation of

divergence ages, I chose the oldest fossil found to constrain the stem of each particular clade (Table 2.2).

The BI analysis was run on Westgrid's "Parallel" cluster (Compute Canada Services) for a total of 891 million generations of Markov chain Monte Carlo (MCMC), with parameters sampled every 30,000 generations and discarded 25% as burn-in using TREEANNOTATOR 2.3.1 (Bouckaert et al., 2014). Availability of time at the cluster determined the number of generations. The resulting log file was checked in TRACER v1.6 (Rambaut et al., 2014) to assess convergence using effective sample size (ESS) values, and the log likelihood versus the generation number plots. The final number of trees used to generate the maximum clade credibility (MCC) tree was 19,079.

2.3.4. Ancestral area estimation

I used the ML method implemented in the R package BioGeoBEARS v.0.2.1 (Matzke, 2013) to estimate the evolution of geographic ranges in endemic genera. BioGeoBEARS allows estimating the ancestral range of taxa using several inference models, such as dispersal, extinction, and cladogenesis (DEC, Ree et al., 2005; Ree & Smith, 2008), dispersal–vicariance (DIVA, Ronquist, 1997), and Bayesian biogeographic inference (BayArea, Landis et al., 2013). BioGeoBEARS requires an ultrametric tree (I used the MCC tree from BEAST) and a matrix of geographic distributions in presence–absence format. As BioGeoBEARS requires positive branch lengths (Matzke, 2013), I manually edited the only negative branch length by adding 0.3 nucleotide substitution per site. To prepare the presence–absence matrix, I obtained species distributions from the Global Biodiversity Information Facility (GBIF, https://www. gbif.org/, accessed 12 June, 2015).

I defined five biogeographic operational areas: (A) Antilles; (B) Central America; (C) South America; (D) North America; and (E) rest of the world (Figure 2.1). Species distributions were coded using the R implementation in the software package SpeciesGeoCoder v.1.0-4 (Töpel et al., 2016). The output presence–absence matrix was visually inspected and corrected manually for erroneous assignments.

I first ran a null analysis with no time frames and equal rates of dispersal among operational areas for each of the six biogeographic models (DEC; DEC*j*; BAYEAREALIKE; BAYAREALIKE*j*; DIVALIKE; DIVALIKE*j*). A second and more complex stratified model was run in order to reflect more realistically the paleogeographic framework of the Caribbean allowing different dispersal rates among operational areas at six different time frames: (1) 0–15 Ma; (2) 15–33 Ma; (3) 33–35 Ma; (4) 35–50 Ma; (5) 50–130 Ma; and (6) 130–378 Ma. Equal dispersal probabilities between regions were scaled from 0 (e.g., when areas were not yet formed) to 1 (e.g., when a land bridge or continuous landmass is proposed to have connected operational areas). I used intermediate values (i.e., 0.01, 0.1, and 0.5) to constrain dispersal events reflecting the presumed biotic connectivity between areas. To test the hypothesized directionality of dispersal events from south to north (Weir et al., 2009; Bacon et al., 2015), a third complex stratified model was run using unequal dispersal probabilities for the period 0–15 Ma. I allowed an extra 0.25 with respect to the previous complex model for the dispersal probabilities from SA to the Antilles and from SA to NA. The complete dispersal matrices used in the ancestral area reconstruction analysis are shown in Table 2.3 and the detailed explanation of the paleogeographic context under each time frame is as follows:

1. 0–15 Ma: From the middle Eocene to the Holocene, landmasses had approximately occupied their current position. The Central American Seaway between South America and the Panama Bloc was fully closed by 15–13 Ma (Montes et al., 2015; Jaramillo et al., 2017) facilitating biotic interchange between North and South America as shown for wide range of taxonomic groups in Bacon et al. (2015). I therefore set up a dispersal constraint of 0.5 between North and South America and gave the maximum dispersal score of 1 between possible dispersal events between Central and North America and between Central and South America reflecting connectivity between landmasses (De Baets et al., 2016). During this period, the Antilles were already above water; therefore, dispersal from/to the Antilles and the surrounding landmasses was possible; I therefore set a constraint of 0.5 for dispersal from/to the Antilles and North America, Central America, and South America, and a minimal constraint of 0.1 from/to the rest of the world. Northern Hemisphere landmasses were at least partially connected through land bridges that increased the connectivity among regions. The Beringia land bridge connected Eurasia to North America and was interrupted around ca. 5.5 Ma (Gladenkov et al., 2002). The North Atlantic land bridge connecting Europe to North America was hypothesized to have existed between the regions up to the Eocene (Tiffney, 1985); however, studies based on ocean microfauna and ocean circulation patterns suggest that the land bridge might have existed until as late as 15 Ma (Schnitker, 1980; Poole & Vorren, 1993). Therefore, I set a minimal constraint of 0.1 to reflect potential dispersal from/to North America and the rest of the world, and the same 0.1 constraint from/to South America and the rest of the world for potential long distance dispersal events;

2. 15–33 Ma: I reduced the dispersal probabilities from/to North and South America to 0.1 in order to reflect the preclosure of the Panama Isthmus (Montes et al., 2015);

3. 33–35 Ma: For this time frame, I kept the same dispersal probabilities as in time frame 1 but allowed a higher dispersal probability of 1 from/to South America and the Antilles to reflect the hypothesized GAARlandia land bridge;

4. 35–50 Ma: From early to late Eocene. I set a dispersal probability of 0.5 from/to Antilles and North America, South America, and Central America. As Central America was not fully formed, I set a probability of 0.1 for potential dispersal events from/to Central and North America, and Central and South America. I also set a probability of 0.01 from/to Central America and the rest of the world. A probability of 0.1 was given for dispersal events from/to North and South America;

5. 50–130 Ma: From late Eocene to lower Cretaceous. Central America was not fully formed, restricting the possibility of dispersal between North America and South America (Montes et al., 2015). To reflect potential long distance dispersal events, I set a constraint of 0.01 from/to North and South America, and also from/to North America and the rest of the world, and from/to South America and the rest of the world. I imposed a dispersal constraint of 0.01 for migrations from/to Central America and the Antilles and also 0.01 from/to Central America and the rest of the world. Same minimal dispersal probability of 0.01 from/to Central America and South America and from/to Central America and North America. Greater Antilles were above water, and Lesser Antilles started forming (Graham, 2003; Pindell & Kennan, 2009), thus, I imposed a minimal dispersal probability of 0.1 from/to Antilles to South and North America;

6. 130–378 Ma: From the middle Devonian to late Jurassic, landmasses were mostly conglomerated, and the Pangea supercontinent started to break up at about 200 Ma, forming Gondwana and Laurasia. Gondwana started to break up about 150 Ma. I reduced all dispersal constrains in this time frame to 0.01 as the lower bound of this time frame is contemporaneous to the estimated origin of Angiosperms (Bell et al., 2010; Silvestro et al., 2015), and it is prior to the origin of Zamiaceae (Salas-Leiva et al., 2013). I did not allow any dispersal event from/to the Antilles as those islands had not formed yet.

2.3.5. A compilation of independent evolutionary and biogeographic studies on a subset of Caribbean endemic genera

An uneven and/or limited taxon sampling and lack of phylogenetic resolution resulting from a few sampled genes can bias estimates of divergence times and ancestral areas in the broad-scale dated phylogenetic analysis (Linder et al., 2005; Pirie & Doyle, 2012). I therefore conducted a second approach to contrast and validate my broad-scale results using multiple independently dated phylogenetic trees. I compiled crown and stem ages from published trees that comprised the Caribbean endemic genera included in the broad-scale analysis. I found information for 24 of the 41 endemic genera. Six of these studies (covering 11 endemic genera) investigated the ancestral areas for such endemic genera (Table 2.4).

2.4. Results

2.4.1. Phylogenetic reconstruction and molecular dating

The ML and BI analyses recovered congruent tree topologies for the higher relationships of taxa. Figure 2.2 shows the phylogenetic relationships among families and the distribution of endemic genera across the BI tree. ESS values were above 200, except for the treeLikelihood (ESS of 103), and the *Trithrinax* and Solanaceae calibration points (ESS of 39 and 23, respectively).

Most family relationships were congruent with the latest APG III, (2009). In one exception, the BI, but not the ML analysis, recovered a clade of 15 Euphorbiaceae species within the Orchidaceae clade, which I therefore removed from the tree using the '*drop.tip*()' function in the R package 'ape' (Paradis et al., 2004). Consequently, the endemic genera *Moacroton* and *Acidocroton* were removed from the dated tree and will not be further discussed. In addition, due to my taxon sampling criteria and DNA marker selection, some clades containing endemic genera had very few (<3) species, did not include the sister genus, or the suprageneric rank was poorly represented (Table 2.1). Therefore, divergence times and ancestral areas for the endemic genera *Doerpfeldia*, *Espadaea*, *Fuertesia*, *Goetzea*, *Haenianthus*, *Petitia*, and *Synapsis* could not be estimated accurately, and results are not shown. After these exclusions, the total number of endemic genera for which I present results is 32.

The Bayesian dating analysis showed that divergence between Angiosperms and Gymnosperms occurred at 370 Ma ([95% HPD (higher posterior density) 366–374 Ma]). The mean crown age for the Angiosperms was estimated at 191 Ma (95% HPD 162–220 Ma), and 50.8 Ma (95% HPD 26.2–79.1) for the Zamiaceae. Mean crown ages of

endemic genera dated from the early Eocene [*Hebestigma*, Leguminosae: 53.1 (95% HPD 33.1–73.0) Ma] to the Pliocene [*Stahlia*, Leguminosae: 3.40 (95% HPD 0.0078–8.50) Ma], whereas mean stem ages ranged from the late Cretaceous [*Hebestigma*, Leguminosae: 106 (95% HPD 88.6–123) Ma] to the middle–late Miocene [*Stahlia*, Leguminosae: 8.64 (95% HPD 1.86–15.9) Ma]. Eleven of the 32 endemic genera had stem and crown node 95% HPD ages younger than the GAARlandia time frame (<33 Ma), while 22 genera had stem and/or crown 95% HPD ages during the hypothesized land bridge (Figure 2.3). *Hebestigma* was probably the only genus that diverged before GAARlandia as the lowest 95% HPD bound of its crown age was estimated at 33 Ma. The mean crown ages of three endemic genera occurred within the GAARlandia time frame (*Acidoton* [Euphorbiaceae] at 31.8 Ma; *Arcoa* [Leguminosae] at 34.1 Ma; and *Chacotheca* [Phyllanthaceae] at 32.9 Ma). The mean stem age of *Neobracea* [Apocynaceae, 35.6 Ma] also fell within GAARlandia. Divergence time estimations at crown and stem nodes for endemic genera can be found in Table 2.5 (see also Figures A1 and A2 for BEAST MCC tree and node numbers in MCC tree in Appendix A). The four-gene concatenated matrix, and BEAST MCC tree are available in Dryad (http://dx.doi.org/10.5061/dryad.gq93s).

2.4.2. Ancestral area estimation

The likelihood values for the null and the two complex stratified models can be found in Table 2.6 for the 18 models ran in BioGeoBEARS. Model selection did not support the hypothesized directionality of dispersal from south to north (complex model 2). The DEC*j* model from the complex model 1 (with founder effect, time stratification and

symmetrical dispersal constraints) was selected as the most appropriate for my data set (Lnl = −1221.2 and AIC = 2,448.3), while the second-best model was the DEC*j* from the complex model 2 (Lnl = −1235.1 and AIC = 2,476.3).

The complete ancestral area estimation using the DEC*j* model (complex model 1) is shown in Figure A3 in Appendix A. To plot the most likely ancestral distribution of endemic genera, I selected their corresponding stem nodes. When an endemic genus was sister to another endemic genus, I treated both genera as a unit (i.e., endemic clade) and selected the stem node of the endemic clade to plot the results. This was the case of *Lasiocroton–Leucocroton*, *Dilomilis–Neocogniauxia*, and *Broughtonia–Psychilis–Quisqueya–Tetramicra*. I considered *Bonania–Grimmeodendron* also as an endemic clade, because the sister relationship of *Grimmeodendron eglandulosum* and *Sebastiania bilocularis* (a nonendemic species) was not well supported in the tree [posterior probability (PP) of .43], and *Grimmeodendron–Sebastiana* was sister to the endemic genus *Bonania* with strong support (PP=.99).

My results show that nine endemic genera or clades and their sister groups had ancestors distributed in the Antilles (i.e., highest probability values for the Antilles, Figure 2.4). This corresponds to a total of 16 of the 32 sampled endemic genera. Four endemic genera or clades (*Lasiocroton–Leucocroton*, *Penelopeia*, *Stahlia*, and *Dendropemon*) colonized the Antilles from CA. Five genera (*Brya*, *Calycogonium*, *Hebestigma*, *Leptocereus*, and *Rhodopis*) had ancestors distributed in SA, and five genera (*Acidoton*, *Arcoa*, *Chascotheca*, *Ditta*, and *Picrodendron*) had ancestors widely distributed in the rest of the world. The ancestral area of *Neobracea* was estimated in the Antilles and rest of the world (Figure 2.4). Furthermore, none of the endemic genera

surveyed had ancestors solely distributed in NA. Nine endemic genera or clades had highest area probabilities of <50%, illustrating the degree of uncertainty in the analysis (see Table 2.5 for ancestral reconstruction probabilities for each genus).

My analysis also recovered ten instances in which continental taxa appeared nested within a Caribbean clade, suggesting potential island to mainland colonization. More specifically, I detected five colonization events from the Antilles to CA (in Orchidaceae, Leguminosae, Euphorbiaceae, Rubiaceae, and Arecaceae) and one to SA (in Cucurbitaceae). In three instances, the nested continental taxa had widespread distributions in continental America (in Zamiaceae and Phyllanthaceae (CA and SA), Rubiaceae (AN and CA), and Leguminosae (CA and NA)).

2.4.3. A compilation of independent evolutionary and biogeographic studies on a subset of Caribbean endemic genera

I found 10 studies that reported crown and stem ages for 24 endemic genera (Table 2.4). In these studies, divergence times ranged from 1.17 (in *Lasiocroton*) to 95 Ma (in *Ditta*) for the crown ages and from 2.6 (in *Acidoton*) to 105 Ma (in *Ditta*) for the stem ages. Five genera (*Acidoton*, *Bonania*, *Ditta*, *Lasiocroton*, and *Leucocroton*) had stem and crown ages outside the 95% HPD interval recovered in the broad-scale analysis, the rest (18) had crown and/or stem ages inside my 95% HPD interval, and thus, I consider them in agreement with my broad-scale approach.

A biogeographic origin was proposed across six studies for eleven endemic genera based on several methods including DEC, RASP, and DIVA-GIS. Five genera were hypothesized to have reached the Antilles from SA (*Acidoton*, *Anacaona*, *Cubanola*,

Penelopeia, and *Leptocereus*), while four had ancestors distributed in the Antilles (*Lasiocroton*, *Leucocroton*, *Hemithrinax*, and *Zombia*). The endemic cycad genus *Microcycas* was reported to have an African Caribbean ancestor, and an ancestor distributed in parts of Mexico, Mesoamerica, and SA was recovered for *Bonania* (Table 2.4).

2.5. Discussion

The origin of endemic genera exhibited a mixed pattern of colonization from continental masses and in situ radiations within the islands, where all continental surrounding masses except for NA appeared to be sources for island colonization. Twenty-two of the 32 genera had crown and/or stem node 95% HPD ages within the hypothesized GAARlandia time span. However, based on the range evolution analysis, I found no support for the hypothesized facilitative role of GAARlandia for SA colonizers, as crown and stem ages for endemic genera with ancestors distributed in SA did not fall within the GAARlandia period.

2.5.1. Oligocene to Miocene origin of Caribbean endemic plant genera

The age for the split between Angiosperms and Gymnosperms in Silvestro et al. (2015, 95% HPD 367.2–382.3 Ma) is congruent with the one I found in my study. The crown age I recovered for the Angiosperms was also concordant with earlier studies (Bell et al., 2010, 95; % HPD 167–199 Ma; Smith et al., 2010, 95; % HPD 182–257 Ma; Magallón et al., 2013; 95% HPD 171.48–257.86 Ma) but older than that reported in Silvestro et al.

(2015, 95% HPD 133.0–151.8 Ma) and Magallón et al. (2015, 95% HPD 136–139.95 Ma).

My results showed that endemic plant genera of the Caribbean originated (mean stem to crown node ages) from the late Cretaceous (ca. 105 Ma) until the Pliocene (ca. 3 Ma), a period during which the Caribbean islands reached their current position with respect to the surrounding continental masses. These divergence times support the perception that at least some endemic Caribbean biota reflects the ancient geological history of the archipelago (Ricklefs & Bermingham, 2008), and a model of continuous assembly of generic diversity in the Caribbean based on Dominican amber deposits from the early Eocene to early Miocene (Iturralde-Vinent & MacPhee, 1996). The inferred mean crown and stem ages were generally congruent with previous studies that have included endemic Caribbean genera, except for five genera in the Euphorbiaceae, for which the literature reports stem and crown ages outside the age confidence intervals (Figure 2.3). Divergence times for *Acidoton*, *Leucocroton*, and *Lasiocroton* estimated here were older than those reported in Cervantes et al. (2016). This is expected even for cases where fossil calibrations are correctly implemented in a molecular dating analysis (including their phylogenetic placement, age, and implementation). This is because fossils only provide minimum ages, and some fossils should just by chance be far too young in relation to the taxon they represent. By performing a single molecular dating analysis with a large supermatrix and several fossils for calibration, my results should reduce such stochastic errors and provide a more consistent estimation for all internal clade ages (see also Antonelli et al., 2017). The divergence time for *Ditta* inferred here was much younger than the estimate of Van Ee et al. (2008). Their estimate should be taken with

caution, however, as *Ditta* was sampled as part of the out-group in a species-level phylogeny focused on *Croton* subgenus *Moacroton* which belongs to a different tribe.

I found 22 of the 32 genera for which the 95% HPD ages at the stem and/or crown nodes overlap with the GAARlandia time frame. As divergence times alone cannot unequivocally support the GAARlandia hypothesis, I discuss below their relevance in light of the range evolution analysis.

2.5.2. Colonization from the continent and in situ speciation of Caribbean endemic plant genera

The Caribbean archipelago is considered to be sufficiently isolated from continental masses to allow allopatric divergence, but relatively close to maintain a dynamic island–continental interaction of biota (Ricklefs & Bermingham, 2008). As the Antilles is surrounded by continental landmasses, one might expect a great proportion of Caribbean ancestors to have occurred in them, and my results support this. About 32% of the endemic genera had ancestors distributed in continental America. My results support overseas dispersal as an important factor to explain the distribution of endemic genera in the Antilles, in contrast to earlier works based on vicariance biogeography, which proposed that the Caribbean biota reflects the early geological history of the Proto-Antilles arc (Rosen, 1975, 1985).

My biogeographic analysis recovered CA, SA, and areas from the Old World as colonization sources for endemic plant genera in the Caribbean. Central American ancestors most probably reached the islands via the Central American Seaway, which is inferred by simulation models to have had a west-to-east direction prior to the closure of

the Isthmus of Panama (Sepulchre et al., 2014). However, wind dispersal cannot be ruled out as it has been documented for sister genera (Renner, 2004; Cervantes et al., 2016), and hurricanes occur frequently in the region (Hedges, 2001). An exception to these dispersal modes (sea currents and wind) is the endemic genus *Dendropemon* (Loranthaceae) with a Central American ancestor and with seeds exclusively consumed by frugivorous birds (Kuijt, 2011). My results of a Central American ancestor disagree with those reported in the literature in two cases. The first is the endemic clade *Lasiocroton–Leucocroton* (Euphorbiaceae), for which Jestrow et al. (2012a) and Cervantes et al. (2016) suggested an ancestor in eastern Cuba and the Antilles, respectively. The second is *Penelopeia* (Cucurbitaceae), for which Schaefer et al. (2009) proposed a South American origin for the subfamily Cucurbitoideae. I attribute this disagreement to the differences in taxon sampling.

Despite floristic similarities between SA and the Caribbean at the genus level (Acevedo-Rodríguez & Strong, 2008), I only found five endemic genera with South American ancestors. Of these five, the mean stem ages of *Leptocereus* and Rhodopis (11 and 21 Ma, respectively) were too young for GAARlandia to have acted as a dispersal route, and the evolution of *Hebestigma* was too old (Figure 2.3). Only *Brya* and *Calycogonium* could have used GAARlandia to colonize the Antilles from SA as the 95% HPD age at the stem and/or crown nodes fell within the land bridge's time span. The literature only reports a biogeographic analysis for *Leptocereus,* which agrees with my SA ancestral area result (Hernández-Hernández et al., 2014).

My analyses revealed Old World ancestors for five endemic genera. For example, *Acidoton* (Euphorbiaceae) formed a clade with two North American species of *Tragia*,

and this clade is sister to southeast Asian and African species. The ancestor of this clade could have used Northern Hemisphere corridors, or a trans-Atlantic or trans-Pacific dispersal to reach the American continent (Heads, 2008; Michalak et al., 2010; Wei et al., 2015). However, Cervantes et al. (2016) proposed a South American ancestor for *Acidoton*, which I also attribute to their different taxon sampling. I found no formal biogeographic analysis for any of the other four genera in the literature.

My analysis recovered Antillean ancestors for 15 endemic genera within nine clades. For example, for *Microcycas*, the only gymnosperm endemic genus included in this study, Salas-Leiva et al. (2013) proposed an African Caribbean ancestor for the clade *Stangeria– Zamia–Microcycas*. The fact that the African genus *Stangeria* is not present in my analyses might explain the disagreement between the Antillean ancestor I recovered for *Microcycas* and their study. These two biogeographic analyses do not support the hypothesis based on fossil evidence that *Microcycas* originated in continental America, reached Cuba, and then became extinct in the continent (Hermsen et al., 2006). Antillean-distributed ancestors were recovered for all endemic orchids (seven genera within three clades). For the genus *Bonania* (Euphorbiaceae), my results showed an Antillean and North American distributed ancestor, whereas Cervantes et al. (2016), who did not include the endemic sister genus *Grimmeodendron*, recovered an ancestor distributed in Mesoamerica, SA, and the Caribbean. My finding of Antillean-distributed ancestors for the endemic legume genera *Pictetia* and *Poitea* was not in agreement with that of Lavin et al. (2001b) who using a cladistic vicariance analysis hypothesized on a boreotropical origin for these two endemic genera. I attribute this disagreement to differences in taxon sampling and biogeographic analysis method. My analysis recovered the Antilles as the

most likely ancestral range for the two palm genera *Hemithrinax* and *Zombia* and for *Cubanola* (Rubiaceae) corroborating the results of (Cano et al., 2018), and Antonelli et al. (2009), respectively.

My phylogenetic framework also identified nine instances for which Antillean taxa acted as source for continental taxa. This result is in line with Bellemain & Ricklefs (2008), which highlights the important and traditionally neglected role of islands as sources to colonize continental masses as seen for some terrestrial animals. Islands acting as reservoir of genetic diversity for the assemblage of continental floras have been reported for plants in other island systems (Andrus et al., 2004; Carine et al., 2004; Patiño et al., 2015; Condamine et al., 2017). My results showed that all the operational areas defined in this study received immigrants from the islands. The time of these recolonizations ranges from late Pleistocene to early Oligocene (between 2 and 25 Ma). By then, most geological events that led to the current formation of the islands had taken place, rendering over-water, bird, or wind dispersal the most plausible explanation.

2.6. References

Acevedo-Rodríguez P. & Strong M.T. (2008) Floristic richness and affinities in the West Indies. *Botanical Review*, **74**, 5–36.

Akaike H. (1974) A New Look at the Statistical Model Identification. *IEEE Transactions on Automatic Control*, **19**, 716–723.

Ali J.R. (2012) Colonizing the Caribbean: Is the GAARlandia land-bridge hypothesis gaining a foothold? *Journal of Biogeography*, **39**, 431–433.

Alonso R., Crawford A.J., & Bermingham E. (2012) Molecular phylogeny of an endemic radiation of Cuban toads (Bufonidae: *Peltophryne*) based on mitochondrial and nuclear genes. *Journal of Biogeography*, **39**, 434–451.

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.

Antonelli A., Hettling H., Condamine F.L., Vos K., Nilsson R.H., Sanderson M.J., Sauquet H., Scharn R., Silvestro D., Töpel M., Bacon C.D., Oxelman B., & Vos R.A. (2017) Toward a Self-Updating Platform for Estimating Rates of Speciation and Migration, Ages, and Relationships of Taxa. *Systematic Biology*, **66**, 152–166.

Antonelli A., Nylander J.A.A., Persson C., & Sanmartı I. (2009) Tracing the impact of the Andean uplift on Neotropical plant evolution. *Proceedings of the National Academy of Sciences of the United States of America PNAS* **, 106**, 9749–9754.

APG III (2009) An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. *Botanical Journal of the Linnean Society*, **161**, 105–121.

Appelhans M.S., Keßler P.J.A., Smets E., Razafimandimbison S.G., & Janssens S.B. (2012) Age and historical biogeography of the pantropically distributed Spathelioideae (Rutaceae, Sapindales). *Journal of Biogeography*, **39**, 1235–1250.

Bacon C.D., Baker W.J., & Simmons M.P. (2012) Miocene dispersal drives island radiations in the palm tribe Trachycarpeae (Arecaceae). *Systematic Biology*, **61**, 426–442.

Bacon C.D., Silvestro D., Jaramillo C., Tilston B., & Chakrabarty P. (2015) Biological evidence supports an early and complex emergence of the Isthmus of Panama. *Proceedings of the National Academy of Sciences of the United States of America PNAS*, **112**, 6110–6115.

De Baets K., Antonelli A., & Donoghue P.C.J. (2016) Tectonic blocks and molecular clocks. *Philosophical Transactions of the Royal Society of London B*, **371**, 20160098.

Baldwin B.G., Wessa B.L., & Panero J.L. (2002) Nuclear rDNA Evidence for Major Lineages of Helenioid Heliantheae (Compositae). *Systematic Botany*, **27**, 161–198.

Barraclough T.G. (2010) Evolving entities: towards a unified framework for understanding diversity at the species and higher levels. *Philosophical transactions of the Royal Society of London B*, **365**, 1801–1813.

Barraclough T.G. & Humphreys A.M. (2015) The evolutionary reality of species and

higher taxa in plants: A survey of post-modern opinion and evidence. *New Phytologist*, **207**, 291–296.
Bell C.D., Soltis D.E., & Soltis P.S. (2010) The age and diversification of the Angiosperms re-revisited. *American Journal of Botany*, **97**, 1296–1303.
Bellemain E. & Ricklefs R.E. (2008) Are islands the end of the colonization road? *Trends in Ecology and Evolution*, **23**, 1–8.
Binford G.J., Callahan M.S., Bodner M.R., Rynerson M.R., Núñez P.B., Ellison C.E., & Duncan R.P. (2008) Phylogenetic relationships of *Loxosceles* and *Sicarius* spiders are consistent with Western Gondwanan vicariance. *Molecular Phylogenetics and Evolution*, **49**, 538–553.
Bouckaert R., Heled J., Kühnert D., Vaughan T., Wu C.-H., Xie D., Suchard M.A., Rambaut A., & Drummond A.J. (2014) BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. *PLoS Computational Biology*, **10**, e1003537.
Bremer B. & Eriksson T. (2009) Time tree of Rubiaceae: phylogeny and dating the family, subfamilies, and tribes. *International Journal of Plant Sciences*, **170**, 766–793.
Burke K. (1988) Tectonic Evolution of the Caribbean. *Annual Review of Earth and Planetary Sciences*, **16**, 201–30.
Cano A., Bacon C.D., Stauffer F.W., Antonelli A., Serrano-Serrano M.L., & Perret M. (2018) The roles of dispersal and mass extinction in shaping palm diversity across the Caribbean. *Journal of Biogeography*, **45**, 1432–1443.
Carine M.A., Russell S.J., 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.
Caujapé-Castells J. (2011) Preface. *The Biology of Island Floras* (ed. by D. Bramwell and J. Caujape-Castells), pp. xiii–xvi. Cambridge University Press, Cambridge, UK.
Cervantes A., Fuentes S., Gutiérrez J., Magallón S., & Borsch T. (2016) Successive arrivals since the Miocene shaped the diversity of the Caribbean Acalyphoideae (Euphorbiaceae). *Journal of Biogeography*, **43**, 1773–1785.
Chamberland L., McHugh A., Kechejian S., Binford G.J., Bond J.E., Coddington J., Dolman G., Hamilton C.A., Harvey M.S., Kuntner M., & Agnarsson I. (2018) From Gondwana to GAARlandia: Evolutionary history and biogeography of ogre-faced spiders (*Deinopis*). *Journal of Biogeography*, **45**, 2442–2457.
Chandler M. (1962) *The Lower Tertiary Floras of Southern England II Flora of the Pipe-Clay Series of Dorset (Lower Bagshot).* London, British Museum (Natural History),
Christophel D.C., Harris W.K., & Syber a. K. (1987) The Eocene flora of the Anglesea Locality, Victoria. *Alcheringa: An Australasian Journal of Palaeontology*, **11**, 303–323.
Cohen K.M., Finney S.C., Gibbard P.L., & Fan J.. (2013) The ICS International Chronostratigraphic Chart. Episodes 36. 199–204.
Condamine F.L., Leslie A.B., & Antonelli A. (2017) Ancient islands acted as refugia and pumps for conifer diversity. *Cladistics*, **33**, 69–92.
Darriba D., Taboada G.L., Doallo R., & Posada D. (2012) jModelTest 2: more models, new heuristics and parallel computing. *Nature methods*, **9**, 772–772.

Darwin C. (1876) *Geological observations on the volcanic islands and parts of South America visited during the voyage of H.M.S. "Beagle."* London: Smith, Elder, & Co., London.

Dávalos L.. (2004) Phylogeny and biogeography of Caribbean mammals. *Biological Journal of the Linnean Society*, **81**, 373–394.

Davis C.C., Webb C.O., Wurdack K.J., Jaramillo C.A., & Donoghue M.J. (2005) Explosive radiation of Malpighiales supports a Mid-Cretaceous origin of modern Tropical Rain Forests. *The American Naturalist*, **165**, E36–E65.

Dransfield J., Uhl N., Asmussen C.B., Baker W.J., Harley M.M., & Lewis C.E. (2008) *Genera Palmarum: the evolution and classification of palms.* Kew Publishing, Richmond (VA).

Drummond A.J., Ho S.Y.W., Phillips M.J., & Rambaut A. (2006) Relaxed phylogengetic and dating with confidence. *PLoS Biology*, **4**, 0699–0710.

Van Ee B.W., Berry P.E., Riina R., & Gutiérrez Amaro J.E. (2008) Molecular phylogenetics and biogeography of the Caribbean-Centered *Croton* subgenus *Moacroton* (Euphorbiaceae s.s.). *Botanical Review*, **74**, 132–165.

Federhen S. (2012) The NCBI Taxonomy Database. *Nucleic Acids Research*, **40**, D136–D143.

Francisco-Ortega J., Santiago-Valentín E., Acevedo-Rodríguez P., Lewis C., Pipoly III J., Meerow A.W., & Maunder M. (2007) Seed plant genera endemic to the Caribbean Island Biodiversity Hotspot: A review and a molecular phylogenetic perspective. *The Botanical Review*, **73**, 183–234.

Fritsch P.W. (2003) Multiple Geographic Origins of Antillean *Styrax*. *Systematic Botany*, **28**, 421–430.

Fritsch P.W. & McDowell T.D. (2003) Biogeography and phylogeny of Caribbean Plants-Introduction. *Systematic Botany*, **28**, 376–377.

Gladenkov A.Y., Oleinik A.E., Marincovich L., & Barinov K.B. (2002) A refined age for the earliest opening of Bering Strait. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **183**, 321–328.

Graham A. (2003) Geohistory models and Cenozoic paleoenvironments of the Caribbean region. *Systematic Botany*, **28**, 378–386.

Haila Y. (1990) Toward an ecological definition of an Island - a Northwest European perspective. *Journal of Biogeography*, **17**, 561–568.

Heads M. (2008) Globally basal centres of endemism: The Tasman-Coral Sea region (south-west Pacific), Latin America and Madagascar/South Africa. *Biological Journal of the Linnean Society*, **96**, 222–245.

Hedges S.B. (1996) Historical Biogeography of West Indian Vertebrates. *Annual Review of Ecology and Systematics*, **27**, 163–196.

Hedges S.B. (2001) *Biogeography of the West Indies: An overview*. CRC Press, Boca Raton,

Hedges S.B. (2006) Paleogeography of the Antilles and origin of West Indian terrestrial vertebrates. *Annals of the Missouri Botanical Garden*, **93**, 231–244.

Herendeen P.S., Crepet W.L., & Dilcher D.L. (1992) The fossil history of the Leguminosae: phylogenetic and biogeographical implications. *Advances in legume systematics*, **4**, 303–16.

Herendeen P.S., Lewis G.P., & Bruneau A. (2003) Floral Morphology in Caesalpinioid Legumes: Testing the Monophyly of the "Umtiza Clade" *International Journal of Plant Sciences*, **164**, S393–S407.

Hermsen E.L.J., Taylor T.H.N., Taylor E.D.L., & Stevenson D.E.W.M. (2006) Cataphylls of the Middle Triassic cycad *Antarcticycas schopfii* and new insights into cycad evolution. *American Journal of Botany*, **93**, 724–738.

Hernández-Hernández T., Brown J.W., Schlumpberger B.O., Eguiarte L.E., & Magallón S. (2014) Beyond aridification: Multiple explanations for the elevated diversification of cacti in the New World Succulent Biome. *New Phytologist*, **202**, 1382–1397.

Hulsey C.D., Keck B.P., & Hollingsworth P.R. (2011) Species tree estimation and the historical biogeography of heroine cichlids. *Molecular Phylogenetics and Evolution*, **58**, 124–131.

Iturralde-Vinent M. & MacPhee R.D.. (1996) Age and Paleogeographical Origin of Dominican Amber. *Science*, **273**, 1850–1852.

Iturralde-Vinent M. & MacPhee R.D.. (1999) Paleogeography of the Caribbean Region: Implications for Cenozoic Biogeography. *Bulletin of the American Museum of Natural History*, **238**, 1–95.

Jaramillo C., Montes C., Cardona A., Silvestro D., Antonelli A., & Bacon C.D. (2017) Comment (1) on "Formation of the Isthmus of Panama" by O'Dea *et al*. *Science Advances*, **3**, e1602321.

Jestrow B., Gutiérrez Amaro J., & Francisco-Ortega J. (2012a) Islands within islands: A molecular phylogenetic study of the *Leucocroton* alliance (Euphorbiaceae) across the Caribbean Islands and within the serpentinite archipelago of Cuba. *Journal of Biogeography*, **39**, 452–464.

Jestrow B., Rodríguez F.J., & Francisco-Ortega J. (2010) Generic delimitation in the Antillean Adelieae (Euphorbiaceae) with description of the Hispaniolan endemic genus Garciadelia. *Taxon*, **59**, 1801–1814.

Jestrow B., Valdés J.J., Rodríguez F.J., Jestrow B., Valdês J.J., Rodríguez F.J., & Francisco-Ortega J. (2012b) Phylogenetic placement of the Dominican Republic endemic genus *Sarcopilea* (Urticaceae). *International Association for Plant Taxonomy*, **61**, 592–600.

Jung S.H. & Lee S.J. (2009) Fossil-winged fruits of *Fraxinus* (Oleaceae) and *Liriodendron* (Magnoliaceae) from the Duho Formation, Pohang Basin, Korea. *Acta Geologica Sinica*, **83**, 845–852.

Katoh K. & Standley D.M. (2013) MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution*, **30**, 772–780.

Kearse M., Moir R., Wilson A., Stones-Havas S., Cheung M., Sturrock S., Buxton S., Cooper A., Markowitz S., Duran C., Thierer T., Ashton B., Meintjes P., & Drummond A. (2012) Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, **28**, 1647–1649.

Kimball R.T. & Crawford D.J. (2004) Phylogeny of Coreopsideae (Asteraceae) using ITS sequences suggests lability in reproductive characters. *Molecular Phylogenetics and Evolution*, **33**, 127–139.

Kuijt J. (2011) Monograph of *Dendropemon* (Loranthaceae). *Systematic Botany Monographs*, **92**, .

Landis M.J., Matzke N.J., Moore B.R., & Huelsenbeck J.P. (2013) Bayesian Analysis of Biogeography when the Number of Areas is Large. *Systematic Biology*, **62**, 789–804.

Lavin M., Herendeen P.S., & Wojciechowski M.F. (2005) Evolutionary rates Analysis of Leguminosae Implicates a Rapid Diversification. *Systematic Biology*, **54**, 575–594.

Lavin M., Pennington R.T., Klitgaard B.B., Sprent J.I., De Lima H.C., & Gasson P.E. (2001a) The dalbergioid legumes (Fabaceae): Delimitation of a pantropical monophyletic clade. *American Journal of Botany*, **88**, 503–533.

Lavin M., Wojciechowski M.F., Gasson P.E., Hughes C.E., & Wheeler E. (2003) Phylogeny of Robinioid Legumes (Fabaceae) Revisited: *Coursetia* and *Gliricidia* Recircumscribed , and a Biogeographical Appraisal of the Caribbean Endemics. *Systematic Botany*, **28**, 387–409.

Lavin M., Wojciechowski M.F., Richman A., Rotella J., Sanderson M.J., & Beyra A. (2001b) Identifying Tertiary Radiations of Fabaceae in the Greater Antilles: Alternatives to Cladistic Vicariance Analysis. *International Journal of Plant Sciences*, **162**, S53–S76.

Linder H.P., Hardy C.R., & Rutschmann F. (2005) Taxon sampling effects in molecular clock dating: An example from the African Restionaceae. *Molecular Phylogenetics and Evolution*, **35**, 569–582.

Lomolino M.V., Riddle B.R., Whittaker R.J., & Brown J.H. (2010) *Biogeography*. Sinauer Associates, Inc., Sunderland, Massachusetts.

Magallón S., Hilu K.W., & Quandt D. (2013) Land plant evolutionary timeline: Gene effects are secondary to fossil constraints in relaxed clock estimation of age and substitution rates. *American Journal of Botany*, **100**, 556–573.

Magallón S., Sandra G., & Luna L.S. (2015) A metacalibrated time-tree documents the early rise of flowering plant phylogenetic diversity. *New Phytologist*, **207**, 437–453.

Martínez-Millán M. (2010) Fossil record and age of the Asteridae. *Botanical Review*, **76**, 83–135.

Matzke N.J. (2013) Probabilistic historical biogeography:new models for founder-event speciation, imperfect detection, and fossils allow improved accuracy and model-testing. *Frontiers of Biogeography*, **5**, 242–248.

Maunder M., Leiva A., Santiago-Valentín E., Stevenson D.W., Acevedo-Rodríguez P., Meerow A.W., Mejía M., Clubbe C., & Francisco-Ortega J. (2008) Plant conservation in the Caribbean Island biodiversity hotspot. *Botanical Review*, **74**, 197–207.

Michalak I., Zhang L., & Renner S.S. (2010) Trans-Atlantic, trans-Pacific and trans-Indian Ocean dispersal in the small Gondwanan Laurales family Hernandiaceae. *Journal of Biogeography*, **37**, 1214–1226.

Miller M.A., Pfeiffer W., & Terri Schwartz (2010) Available at: https://www.phylo.org/.

Mittermeier R.A., Gil R.R., Hoffman M., Pilgrim J., T. B., Mittermeier C.G., Lamoreux J., & da Fonseca G.A.B. (2004) *Hotspots revisited: Earth's biologically richest and most threatened terrestrial ecoregions.* CEMEX, Mexico D.F.

Monceau K., Cezilly F., Moreau J., Motreuil S., & Wattier R. (2013) Colonisation and

diversification of the Zenaida Dove (*Zenaida aurita*) in the Antilles: Phylogeography, contemporary gene flow and morphological divergence. *PLoS ONE*, **8**, e82189.

Montes C., Cardona A., Jaramillo C., Pardo A., Silva J.C., Valencia V., Ayala C., Pérez-Angel L.C., Rodriguez-Parra L.A., Ramirez V., & Niño H. (2015) Middle Miocene closure of the Central American Seaway. *Science*, **348**, 226–229.

Motley T.J., Wurdack K.J., & Delprete P.G. (2005) Molecular systematics of the Catesbaeeae-Chiococceae complex (Rubiaceae): Flower and fruit evolution and biogeographic implications. *American Journal of Botany*, **92**, 316–329.

Myers N., Mittermeier R.A., Mittermeier C.G., Fonseca G.A.B., & Kent J. (2000) Biodiversity hotspots for conservation priorities. *Nature*, **403**, 853–858.

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

Patiño J.P., Carine M., Mardulyn P., Devos N., Mateo R., González-Mancebo J.M., Shaw A.J., & Vanderpoorten A. (2015) Approximate Bayesian computation reveals the rucial role of oceanic islands for the assembly of continental biodiversity. *Systematic Biology*, **64**, 579–589.

Pindell J.L. & Kennan L. (2009) Tectonic evolution of the Gulf of Mexico, Caribbean and northern South America in the mantle reference frame: an update. *Geological Society, London, Special Publications*, **328**, 1–55.

Pirie M.D. & Doyle J.A. (2012) Dating clades with fossils and molecules: the case of Annonaceae. *Botanical Journal of the Linnean Society*, **169**, 84–116.

Poinar Jr G. (2002) Fossil palm flowers in Dominican and Mexican amber. *Botanical Journal of the Linnean Society*, **138**, 57–61.

Poole D.A.R. & Vorren T.O. (1993) Miocene to Quaternary paleoenvironments and uplift history on the mid Norwegian shelf. *Marine Geology*, **115**, 173–205.

Rambaut A., Suchard M.A., Xie D., & Drummond A.J. (2014) Available at: http://beast.bio.ed.ac.uk/Tracer.

Ree R.H., Moore B.R., Webb C.O., & Donoghue M.J. (2005) A likelihood framework for inferring the evolution of geographic range on phylogenetic trees. *Evolution*, **59**, 2299–2311.

Ree R.H. & Smith S.A. (2008) Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. *Systematic Biology*, **57**, 4–14.

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

Ricklefs R. & Bermingham E. (2008) The West Indies as a laboratory of biogeography and evolution. *Philosophical Transactions of the Royal Society B*, **363**, 2393–2413.

Roncal J., Zona S., & Lewis C.E. (2008) Molecular Phylogenetic Studies of Caribbean Palms (Arecaceae) and Their Relationships to Biogeography and Conservation. *Botanical Review*, **74**, 78–102.

Ronquist F. (1997) Dispersal-Vicariance Analysis: A new approach to the quantification of Historical Biogeography. *Systematic Biology*, **46**, 195–203.

Rosen D.E. (1975) A Vicariance Model of Caribbean Biogeography. *Systematic Zoology*, **24**, 431–464.

Rosen D.E. (1985) Geological Hierarchies and Biogeographic Congruence in the Caribbean. *Annals of the Missouri Botanical Garden*, **72**, 636–659.

Rova J.H.E., Delprete P.G., Andersson L., & Albert V.A. (2002) A *trnL-F* cpDNA sequence study of the Condamineeae-Rondeletieae-Sipaneeae Complex with implications on the phylogeny of the Rubiaceae. *American Journal of Botany*, **89**, 145–159.

Salas-Leiva D.E., Meerow A.W., Calonje M., Griffith M.P., Francisco-Ortega J., Nakamura K., Stevenson D.W., Lewis C.E., & Namoff S. (2013) Phylogeny of the cycads based on multiple single-copy nuclear genes: congruence of concatenated parsimony, likelihood and species tree inference methods. *Annals of Botany*, **112**, 1–16.

Santiago-Valentín E. & Olmstead R.G. (2004) Historical Biogeography of Caribbean Plants: Introduction to current knowledge and possibilities from a phylogenetic perspective. *Taxon*, **53**, 299–319.

Schaefer H., Heibl C., & Renner S.S. (2009) Gourds afloat: a dated phylogeny reveals an Asian origin of the gourd family (Cucurbitaceae) and numerous oversea dispersal events. *Proceedings. Biological Sciences*, **276**, 843–851.

Schnitker D. (1980) North Atlantic oceanography as possible cause of Antarctic glaciation and eutrophication. *Nature*, **284**, 615–616.

Sepulchre P., Arsouze T., Donnadieu Y., Dutay J.C., Jaramillo C., Le Bras J., Martin E., Montes C., & Waite A.J. (2014) Consequences of shoaling of the Central American Seaway determined from modeling Nd isotopes. *Paleoceanography*, **29**, 176–189.

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

Smith S.A., Beaulieu J.M., & Donoghue M.J. (2010) An uncorrelated relaxed-clock analysis suggests an earlier origin for fl owering plants. *Proceedings of the National Academy of Sciences of the United States of America*, **107**, 5897–5902.

Sosa V., Cameron K.M., Angulo D.F., & Hernández-Hernández T. (2016) Life form evolution in epidendroid orchids: Ecological consequences of the shift from epiphytism to terrestrial habit in *Hexalectris*. *TAXON*, **65**, 235–248.

Stamatakis A. (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*, **22**, 2688–2690.

Sun M., Sun X., Zhao Y., Wang O., Li Z., Hu Z., & Mei P. (1989) *Sporo-pollen.pp. 6-11*. Beijing.

Tiffney B.H. (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.

Tong Y., Binford G., Rheims C.A., & Kuntner M. (2019) Huntsmen of the Caribbean: Multiple tests of the GAARlandia hypothesis. *Molecular Phylogenetics and Evolution*, **130**, 259–268.

Töpel M., Zizka A., Calió M.F., Scharn R., Silvestro D., & Antonelli A. (2016) SpeciesGeoCoder: Fast categorization of species occurrences for analyses of biodiversity, biogeography, ecology, and evolution. *Systematic Biology*, syw064.

Wallace A.R. (1892) *Island life: or, the phenomena and causes of insular faunas and*

floras, including a revision and attempted solution of the problem of geological climates. London: Macmillan.

Wei R., Xiang Q., Schneider H., Sundue M.A., Kessler M., Kamau P.W., Hidayat A., & Zhang X. (2015) Eurasian origin, boreotropical migration and transoceanic dispersal in the pantropical fern genus *Diplazium* (Athyriaceae). *Journal of Biogeography*, **42**, 1809–1819.

Weigelt P., Steinbauer M.J., Cabral J.S., & Kreft H. (2016) Late Quaternary climate change shapes island biodiversity. *Nature*, **532**, 99–102.

Weir J.T., Bermingham E., & Schluter D. (2009) The great American biotic interchange in birds. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 21737–21742.

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

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

Wojciechowski M.F., Lavin M., & Sanderson M.J. (2004) A Phylogeny of Legumes (Leguminosae) based on analysis of the plastid matK gene resolves many well-supported subclades within the family. *American Journal of Botany*, **91**, 1846–1862.

Wurdack K.J., Hoffmann P., & Chase M.W. (2005) Molecular phylogenetic analysis of uniovulate Euphorbiaceae (Euphorbiaceae sensu stricto) using plastid rbcL and trnL-F DNA sequences. *American Journal of Botany*, **92**, 1397–1420.

Zachos J., Pagani M., Sloan L., Thomas E., & Billups K. (2001) Trends, rhythms, and aberrations in global climate 65 Ma to present. *Science*, **292**, 686–693.

Zanne A.E., Tank D.C., Cornwell W.K., Eastman J.M., Smith S.A., FitzJohn R.G., McGlinn D.J., O'Meara B.C., Moles A.T., Reich P.B., Royer D.L., Soltis D.E., Stevens P.F., Westoby M., Wright I.J., Aarssen L., Bertin R.I., Calaminus A., Govaerts R., Hemmings F., Leishman M.R., Oleksyn J., Soltis P.S., Swenson N.G., Warman L., & Beaulieu J.M. (2014) Three keys to the radiation of angiosperms into freezing environments. *Nature*, **506**, 89–92.

Table 2.1. Seed plant genera endemic to the Caribbean Islands sampled in this study. Endemic genus sampling indicates the number of species sampled in this study divided by the total number of species (based on Francisco-Ortega et al., 2007); suprageneric sampling indicates the number of genera in a suprageneric taxon sampled in this study divided by the total number of genera; suprageneric rank refers to the name of suprageneric rank and in parenthesis the number of species within this taxonomic rank included in this study. NCBI Taxonomy facility (Federhen, 2012) was used to select up to 10 species for every genus within their suprageneric rank to which the endemic genera belong to.

Endemic Genus	Endemic genus sampling	Suprageneric sampling	Suprageneric rank (number of species)	Family
Acidocroton	1/3	7/10	Tribe Crotoneae (16)	Euphorbiaceae
Acidoton	1/8	6/12	Tribe Plukenetieae (9)	Euphorbiaceae
Anacaona	1/1	12/13	Tribe Cucurbiteae (20)	Cucurbitaceae
Arcoa	1/1	46/56	Tribe Caesalpinieae (79)	Fabaceae
Bonania	1/8	20/23	Tribe Hippomaneae (30)	Euphorbiaceae
Broughtonia	4/6	50/54	Subtribe Laeliinae (117)	Orchidaceae
Brya	1/4	30/48	Tribe Dalbergieae (53)	Fabaceae
Calycogonium	1/36	3/19-23	Tribe Miconieae (13)	Melastomataceae
Chascotheca	1/2	3/3	Subtribe Astrocasiinae (5)	Phyllanthaceae
Cubanola	1/2	9/28	Tribe Chiococceae (19)	Rubiaceae
Dendropemon	1/36	4/9	Subtribe Psittacanthinae (4)	Loranthaceae
Dilomilis	1/5	50/54	Subtribe Laeliinae (117)	Orchidaceae
Ditta	1/2	7/7	Tribe Adenoclineae (9)	Euphorbiaceae
Doerpfeldia	1/1	1/1	Tribe Doerpfeldieae (1)	Rhamnaceae
Domingoa	2/3	50/54	Subtribe Laeliinae (117)	Orchidaceae
Espadaea	1/1	2/4	Subfamily Goetzeoideae (2)	Solanaceae
Fuertesia	1/1	2/4	Subfamily Gronovioideae (2)	Loasaceae
Goetzea	1/2	2/4	Subfamily Goetzeoideae (2)	Solanaceae
Grimmeodendron	1/2	20/23	Tribe Hippomaneae (30)	Euphorbiaceae
Haenianthus	1/2	15/18	Tribe Oleeae (53)	Oleaceae
Hebestigma	1/1	10/13	Tribe Robinieae (23)	Fabaceae
Hemithrinax	3/3	9/10	Tribe Cryosophileae (14)	Arecaceae
Lasiocroton	3/5	5/6	Tribe Adelieae (18)	Euphorbiaceae
Leptocereus	1/12	14/27	Tribe Echinocereeae (22)	Cactaceae
Leucocroton	3/28	5/6	Tribe Adelieae (18)	Euphorbiaceae
Microcycas	1/1	2/8	Family Zamiaceae (6)	Zamiaceae

Moacroton	1/8	7/10	Tribe Crotoneae (16)	Euphorbiaceae
Neobracea	3/8	2/2	Subtribe Pachypodiinae (7)	Apocynaceae
Neocogniauxia	1/2	50/54	Subtribe Laeliinae (117)	Orchidaceae
Penelopeia	1/1	12/13	Tribe Cucurbiteae (20)	Cucurbitaceae
Petitia	1/2	2/7	Subfamily Viticoideae (2)	Lamiaceae
Picrodendron	1/1	13/19	Family Picrodendraceae (13)	Picrodendraceae
Pictetia	1/8	30/48	Tribe Dalbergieae (53)	Fabaceae
Poitea	3/12	10/13	Tribe Robinieae (23)	Fabaceae
Psychilis	2/15	50/54	Subtribe Laeliinae (117)	Orchidaceae
Quisqueya	1/4	50/54	Subtribe Laeliinae (117)	Orchidaceae
Rhodopis	1/2	47/84	Tribe Phaseoleae (86)	Fabaceae
Stahlia	1/1	46/56	Tribe Caesalpinieae (79)	Fabaceae
Synapsis	1/1	2/3	Family Schlegeliaceae (2)	Schlegeliaceae
Tetramicra	1/13	50/54	Subtribe Laeliinae (117)	Orchidaceae
Zombia	1/1	9/10	Tribe Cryosophileae (14)	Arecaceae

Table 2.2. Calibration points used for divergence time estimation in BEAST v2.3.1. The offset values from the BEAUti settings column correspond to assigned fossil ages.

Fossil name	Clade constrained	Plant organs and synapomorphies	Primary reference	BEAUti settings
Machaerium	Stem of Tribe Dalbergieae (Leguminosae)	Fossil leaflets. Strong marginal vein, poorly organized higher order venation, numerous closely spaced craspedodromous secondary veins, and epidermal cell structure are diagnostic characters for *Machaerium* (Tribe Dalbergieae)	Herendeen et al., 1992	Offset=40, Mean=1.0, SD=0.5
Fraxinus excelsior	Crown of family Oleaceae	Fruit fossils. Winged (samara type) fruit that resembles *Fraxinus* in peduncle, vein structure and shape, and position of seed	Jung & Lee, 2009	Offset=5.33, Mean=1.0, SD=0.5 (as used in Magallón et al., 2015)
Sabalites carolinensis	Stem of Tribe Cryosophileae (Subfamily Coryphoideae, Arecaceae)	Leaf fossil. Oldest known palm fossil assignable to Subfamily Coryphoideae with costapalmate leaf.	Dransfield et al., 2008	Offset=86.7, Mean=1.7, SD=0.3 (as used in Bacon et al., 2012)
Micrantheum spinyspora	Stem of family Picrodendraceae	Pollen fossils	Christophel et al., 1987	Offset=35.55, Mean=1.0, SD=0.5
Acalypha	Stem of tribes Adelieae and Pluketenieae (Subfamily Acalyphoideae, Euphorbiaceae	Pollen fossils. Diagnostic characters of Acalyphoideae include pollen and pores of small size; sculpture punctate-reticulate; thick nexine and separate from sexine around pore, making sexine in the aperture protruding in a fastigium-like chamber.	Sun et al., 1989	Offset=61.0 , Mean=1.0, SD=0.5 (as used in Davis et al., 2005)

Solanispermum reniforme	Stem of family Solanaceae	Fossil seeds; one of the earliest fossils assigned to Solanaceae	Chandler, 1962	Offset=47.0 , Mean=1.0, SD of 0.5 (as used in Martínez-Millán, 2010)
Trithrinax dominicana,	Stem of genus *Trithrinax* (Arecaceae)	Flower fossils. Stamen filaments exerted and tips bent inwards are diagnostic characters for *Trithrinax*	Poinar Jr, 2002	Offset=24.5, Mean=1.0, SD of 0.5
Prosopis linearifolia	Stem of *Umtiza* clade (Fabaceae)	Fossil leaves. Mix of pinnate and bipinnate leaves. Leaflets linear and asymmetric. Terminal group of three pinnae in a single bipinnate leaf, rising from a sessile terminal pinna. These diagnostic characters are associated to *Arcoa* (*Umtiza* clade)	Herendeen et al., 2003	Offset=34.0 , Mean=1.0, SD of 0.5 (as used in Lavin et al., 2005)
	Stem of Angiosperms	Secondary calibration point	Silvestro et al., 2015	Laplace prior distribution, Offset=143.7, μ=1.0, scale=4.36
	Stem of Spermatophytes	Secondary calibration point	Silvestro et al., 2015	Gamma prior distribution, Offset=366.0, Mean=1.0SD=0.5

Table 2.3. Dispersal matrices used in BioGeoBEARS for complex biogeographic modeling. a) Complex model number 1 accounts for equal dispersal probabilities in both directions between areas; b) complex model number 2 is identical as complex model 1 except for the dispersal probabilities from South America to the Antilles and from South America to North America which are increased by 0.25 in model 2, thus favouring dispersal South to North for the 0-15 Ma period. Palaeographical events for the five time periods that informed the dispersal constrains are described in Appendix 1.

a) Complex model 1

0-15 Ma					**15-33 Ma**					**33-35 Ma**				
AN	CA	NA	RW	SA	AN	CA	NA	RW	SA	AN	CA	NA	RW	SA
1	0.5	0.5	0.1	0.5	1	0.5	0.5	0.1	0.5	1	0.5	0.5	0.1	1
0.5	1	1	0.1	1	0.5	1	1	0.1	1	0.5	1	1	0.1	1
0.5	1	1	0.1	0.5	0.5	1	1	0.1	0.1	0.5	1	1	0.1	0.1
0.1	0.1	0.1	1	0.1	0.1	0.1	0.1	1	0.1	0.1	0.1	0.1	1	0.1
0.5	1	0.5	0.1	1	0.5	1	0.1	0.1	1	1	1	0.1	0.1	1

35-50 Ma					**50-130 Ma**					**130-378 Ma**				
AN	CA	NA	RW	SA	AN	CA	NA	RW	SA	AN	CA	NA	RW	SA
1	0.5	0.5	0.01	0.5	1	0.01	0.1	0.01	0.1	1	0	0	0	0
0.5	1	0.1	0.01	0.1	0.01	1	0.01	0.01	0.01	0	1	0.01	0.01	0.01
0.5	0.1	1	0.1	0.1	0.1	0.01	1	0.01	0.01	0	0.01	1	0.01	0.01
0.01	0.01	0.1	1	0.1	0.01	0.01	0.01	1	0.01	0	0.01	0.01	1	0.01
0.5	0.1	0.1	0.1	1	0.1	0.01	0.01	0.01	1	0	0.01	0.01	0.01	1

b) Complex model 2

0-15 Ma					15-33 Ma					33-35 Ma				
AN	CA	NA	RW	SA	AN	CA	NA	RW	SA	AN	CA	NA	RW	SA
1	0.5	0.5	0.1	0.5	1	0.5	0.5	0.1	0.5	1	0.5	0.5	0.1	1
0.5	1	1	0.1	1	0.5	1	1	0.1	1	0.5	1	1	0.1	1
0.5	1	1	0.1	0.5	0.5	1	1	0.1	0.1	0.5	1	1	0.1	0.1
0.1	0.1	0.1	1	0.1	0.1	0.1	0.1	1	0.1	0.1	0.1	0.1	1	0.1
0.75	1	0.75	0.1	1	0.5	1	0.1	0.1	1	1	1	0.1	0.1	1

35-50 Ma					50-130 Ma					130-378 Ma				
AN	CA	NA	RW	SA	AN	CA	NA	RW	SA	AN	CA	NA	RW	SA
1	0.5	0.5	0.01	0.5	1	0.01	0.1	0.01	0.1	1	0	0	0	0
0.5	1	0.1	0.01	0.1	0.01	1	0.01	0.01	0.01	0	1	0.01	0.01	0.01
0.5	0.1	1	0.1	0.1	0.1	0.01	1	0.01	0.01	0	0.01	1	0.01	0.01
0.01	0.01	0.1	1	0.1	0.01	0.01	0.01	1	0.01	0	0.01	0.01	1	0.01
0.5	0.1	0.1	0.1	1	0.1	0.01	0.01	0.01	1	0	0.01	0.01	0.01	1

Table 2.4. Compilation of independent dated phylogenies from the literature. Crown and stem ages in millions of years (Ma).

Endemic Genus	Crown Age in Ma	Stem Age in Ma	Ancestral Area Estimation	Reference
Acidoton	1.7	2.6	SA	Cervantes et al., 2016
Anacaona	13	17	SA	Schaefer et al, 2009
Arcoa	34			Lavin et al., 2005
Bonania	41.6	46.39	MX, SA, MS	Cervantes et al., 2016
Broughtonia	15.68	20.74		Sosa et al., 2016
Brya	41.9	47.2		Lavin et al., 2005
Cubanola	27.6*	34.4*		Bremer & Eriksson, 2009
Dilomilis	16.01	46.72		Sosa et al., 2016
Ditta	95	105		van Ee et al., 2008
Domingoa	19.34	20.9		Sosa et al., 2016
Hebestigma	38.1	48.3		Lavin et al., 2003
Hemithrinax	6.99	17.54	AN	Cano et al., 2018
Leptocereus	2.8	4.8	SA	Hernández-Hernández et al., 2014
Lasiocroton	1.17	10.76	AN	Cervantes et al., 2016
Leucocroton	5.27	10.76	AN	Cervantes et al., 2016
Microcycas	36.5	60.32	AF-CA	Salas-Leiva et al., 2013
Neocogniauxia	16.01	46.72		Sosa et al., 2016
Penelopeia	13	17	SA	Schaefer et al, 2009
Pictetia	14.5	45.6		Lavin et al., 2005
Poitea	9.2	16.4		Lavin et al., 2001b
Psychilis	15.68	20.74		Sosa et al., 2016
Quisqueya	15.68	20.74		Sosa et al., 2016
Tetramicra	15.68	20.74		Sosa et al., 2016
Zombia	3.75	21.7	AN	Cano et al., 2018

**** TRIBE AGES***

Table 2.5. Divergence times resulting from a Bayesian dating analysis in BEAST v2.3.1 at crown and stem nodes and ancestral area reconstruction for each genus showing the most likely ancestral area based on the Complex 1 DEC*j* model.

Endemic Genus	Crown node number	Mean ages at crown nodes in Ma (95% HPD)	Stem node number	Mean ages at stem nodes in Ma (95% HPD)	Ancestral Reconstruction probabilities (at stem nodes)
Acidoton	880	31.74 (13.65-50.99)	879	42.49 (23.65-60.72)	RW 0.38
Anacaona	628	12.71 (4.05-22.29)	626	27.67 (17.54-36.97)	CA 0.39*
Arcoa	640	34.09 (26.56-40.05)	637	38.24 (34.68-41.72)	RW 0.80
Bonania	926	9.92 (2.81-17.60)	925	19.41 (8.32-30.44)	AN 0.19; ANNAm 0.16*
Broughtonia	1136	19.42 (10.91-27.63)	1135	26.73 (18.45-36.05)	AN 0.83*
Brya	851	13.72 (3.52-24.82)	850	44.17 (28.59-57.89)	SA 0.94
Calycogonium	957	26.81 (9.67-44.44)	956	47.27 (26.64-67.79)	SA 0.53
Chascotheca	952	32.86 (15.37-52.74)	951	56.43 (31.72-80.80)	RW0.26; ANRW 0.22
Cubanola	1004	17.65 (9.73-26.12)	1002	20.25 (11.98-28.28)	AN 0.97
Dendropemon	1083	20.45 (5.34-37.22)	1082	44.19 (20.76-70.21)	CA 0.65
Dilomilis	1199	18.74 (5.10-35.04)	1111	52.80 (36.69-69.27)	AN 0.93*
Ditta	937	15.09 (5.32-25.90)	936	28.50 (13.42-43.68)	RW 0.80
Domingoa	1144	15.009 (5.61-24.63)	1116	30.67 (24.48-34.96)	AN 0.57**
Grimmeodendron	927	6.61 (0.99-12.81)	925	19.41 (8.32-30.44)	AN 0.19; ANNAm 0.16*
Hebestigma	806	53.12 (33.13-72.98)	716	105.48 (88.57-123.62)	SA 0.29
Hemithrinax	1210	6.07 (1.71-10.28)	1208	9.35 (3.68-14.72)	AN 0.79*
Leptocereus	979	9.07 (2.43-16.32)	978	11.12 (3.35-17.42)	SA 0.89
Lasiocroton	891	22.59 (11.16-35.23)	890	28.85 (18.81-41.40)	CA 0.54
Leucocroton	891	22.59 (11.16-35.23)	890	28.85 (18.81-41.40)	CA 0.54
Microcycas	1216	28.23 (12.43-45.78)	1215	51 (26.15-79.07)	AN 0.95
Neobracea	1015	6.49 (0.81-13.42)	1014	35.45 (18.29-53.39)	ANRW 0.61

Neocogniauxia	1199	18.74 (5.10-35.04)	1111	52.80 (36.69-69.27)	AN 0.93*
Penelopeia	627	20.91 (9.58-31.80)	626	27.67 (17.54-36.97)	CA 0.39*
Picrodendron	942	19.13 (12.37-26.11)	941	23.93 (16.27-30.93)	RW 0.48
Pictetia	842	13.96 (5.15-23.03)	841	17.67 (7.47-27.37)	AN 0.45
Poitea	826	11.50 (4.63-18.89)	823	17.10 (8.50-26.01)	AN 0.25; ANCA 0.21
Psychilis	1140	15.93 (8.51-24.31)	1135	26.73 (18.45-36.05)	AN 0.83*
Quisqueya	1142	12.08 (4.73-19.92)	1135	26.73 (18.45-36.05)	AN 0.83*
Rhodopis	797	17.47 (6.42-29.48)	796	21.89 (10.33-33.92)	SA 0.14
Stahlia	666	3.40 (0.0078-8.50)	665	8.64 (1.86-15.97)	CA 0.67
Tetramicra	1142	12.08 (4.73-19.92)	1135	26.73 (18.45-36.05)	AN 0.83*
Zombia	1209	8.13 (3.03-13.74)	1208	9.35 (3.68-14.72)	AN 0.79*

Table 2.6. Biogeographic model testing in BioGeoBEARS. Null models have equal plant dispersal probabilities across all areas and through time. Complex models account for unequal dispersal probabilities considering the geological history. LnL, log likelihood; # params, number of parameters; d, dispersal rate per million years along branches; e, extinction rate per million years along branches; j, founder event speciation weighted per speciation event; AIC, Akaike Information Criterion; AIC wt, relative weight for each model. Best model according to AIC marked with *.

Null Model	**LnL**	**# params**	**d**	**e**	**j**	**AIC**	**AIC_wt**
BAYAREALIKE	-1480.99	2	0.003897926	0.019351951	0	2965.992259	3.63E-96
BAYAREALIKE*j*	-1260.30	3	0.002286047	0.000617994	0.046701184	2526.614158	0.933204519
DEC	-1318.92	2	0.004494901	1.00E-12	0	2641.850873	8.84E-26
DEC*j**	-1262.94	3	0.003566693	1.00E-12	0.028423902	2531.888136	0.066795481
DIVALIKE	-1330.90	2	0.005416508	1.00E-12	0	2665.809754	5.55E-31
DIVALIKE*j*	-1294.20	3	0.004284076	1.08E-09	0.022922813	2594.403385	1.78E-15
Complex Model 1	**LnL**	**# params**	**d**	**e**	**j**	**AIC**	**AIC_wt**
BAYAREALIKE	-1465.10	2	0.011318446	0.016440112	0	2934.215946	3.10E-106
BAYAREALIKE*j*	-1235.58	3	0.00847181	0.000807486	0.126630982	2477.169101	5.48E-07
DEC	-1288.10	2	0.015131683	0.000614288	0	2580.218724	2.30E-29
DEC*j*	-1221.16	3	0.011441675	4.87E-05	0.115873802	2448.333573	0.999999452
DIVALIKE	-1295.18	2	0.018188404	0.000809571	0	2594.36264	1.95E-32
DIVALIKE*j*	-1272.81	3	0.01468759	0.000472286	0.020573461	2551.631877	3.71E-23
Complex Model 2	**LnL**	**# params**	**d**	**e**	**j**	**AIC**	**AIC_wt**
BAYAREALIKE	-1468.83	2	0.010448962	0.016429327	0	2941.666889	8.73E-102
BAYAREALIKE*j*	-1243.37	3	0.008059871	0.000673963	0.122983955	2492.751708	0.00026405
DEC	-1296.90	2	0.014309739	0.00060304	0	2597.809953	4.06E-27
DEC*j*	-1235.13	3	0.011095838	5.33E-05	0.063058035	2476.273491	0.99973595
DIVALIKE	-1304.63	2	0.017191254	0.000805488	0	2613.262891	1.79E-30
DIVALIKE*j*	-1281.03	3	0.015055036	0.000812842	0.019658281	2568.078678	1.16E-20

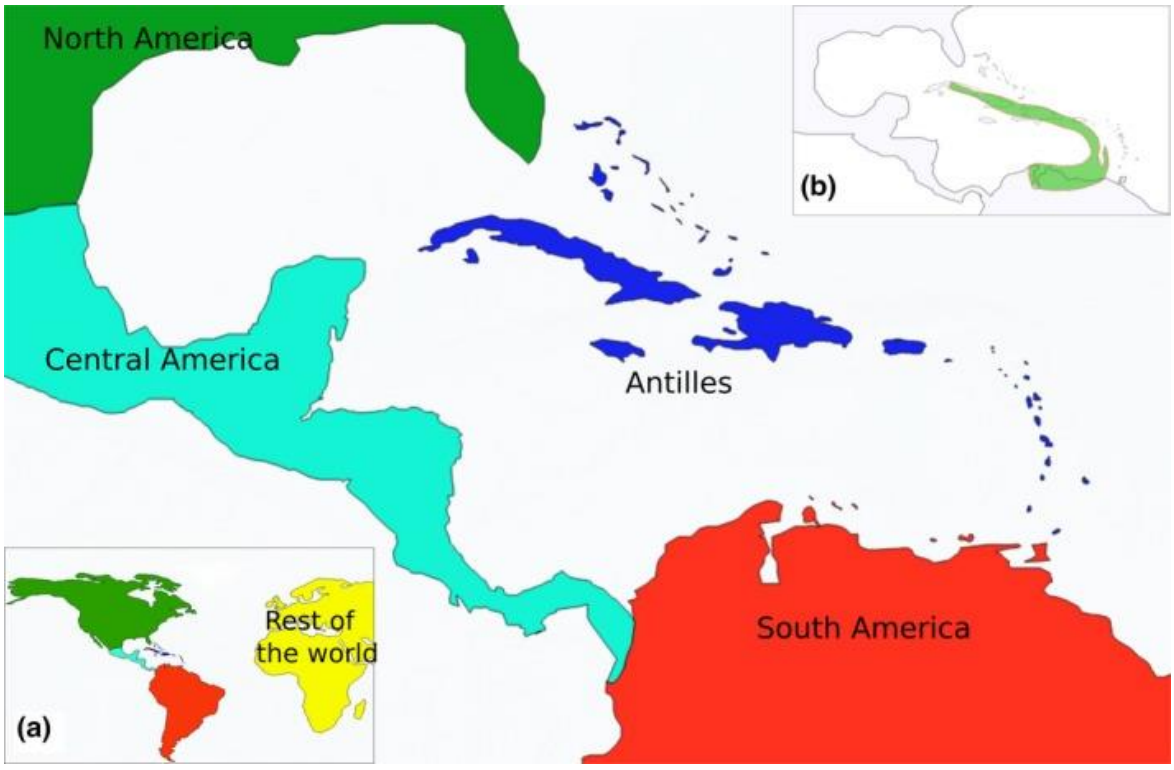


Figure 2.1. Operational biogeographic areas used in this study. Insets: (a) the five operational areas. (b) The GAARlandia land bridge hypothesized to have existed between 33 and 35 Ma (modified from Iturralde-Vinent & MacPhee, 1999).

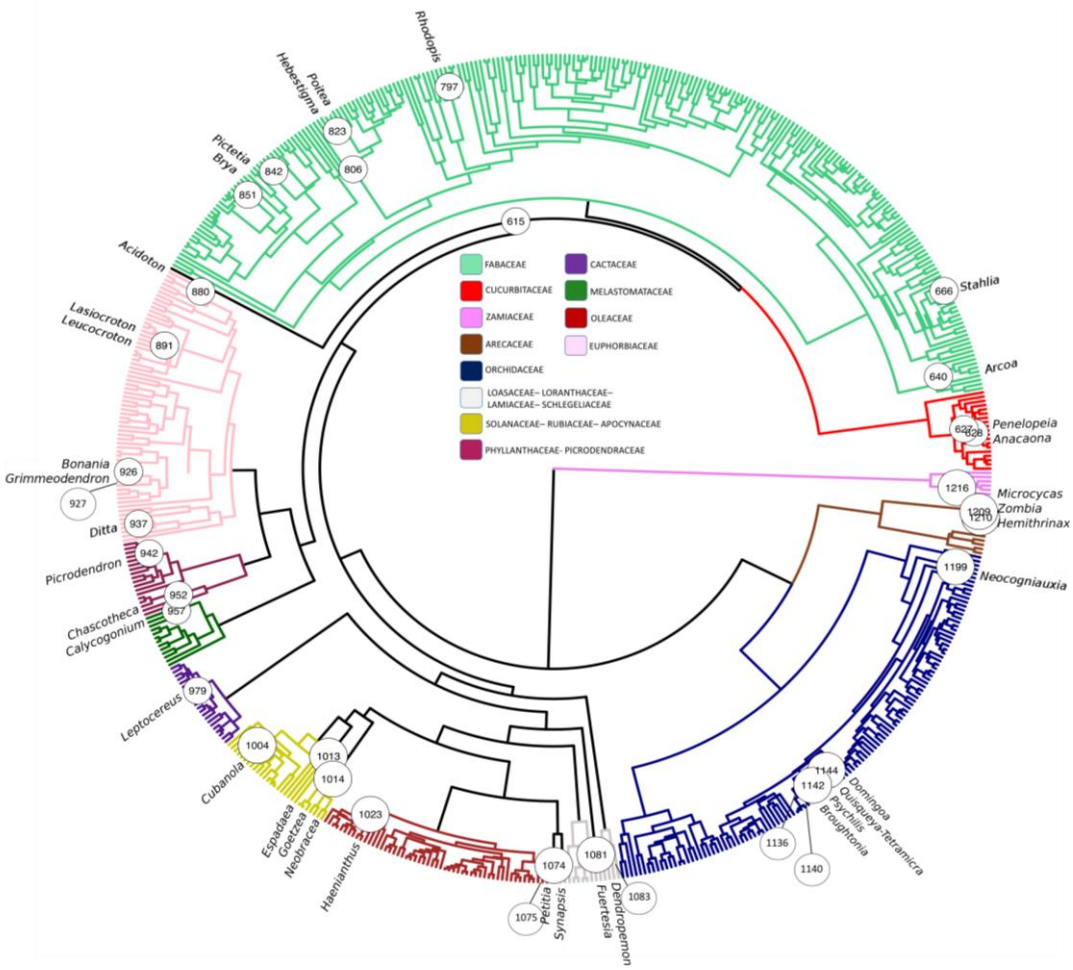


Figure 2.2. Phylogenetic relationships of plant families obtained from the Bayesian dating analysis (maximum clade credibility tree). Numbered circles indicate node number that subtends each Caribbean endemic genus or clade (same as in Table **2.5**). Families within a clade represented in this study by a small number of taxa have been lumped into one color.

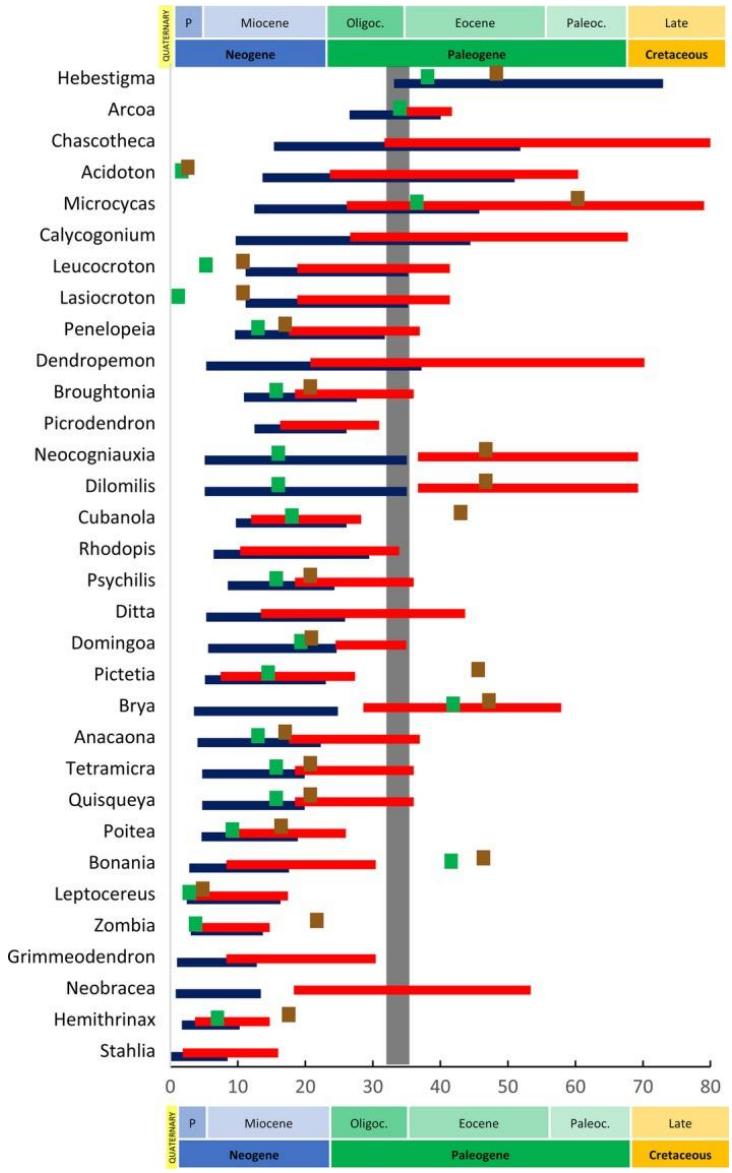


Figure 2.3. Bayesian divergence times of Caribbean endemic genera ordered by age. Blue and red bars correspond to the 95% HPD for the crown and stem node ages, respectively, obtained in my broad-scale analysis. Green and brown squares represent crown and stem ages, respectively, obtained from the literature. Gray vertical band indicates the GAARlandia time frame (33–35 Ma). Geological timescale according to the International Commission on Stratigraphy(v2016/04; Cohen et al., 2013). Pliocene is abbreviated as P, the Oligocene as Oligoc., and the Paleocene as Paleoc. Note the 95% HPD for the stem of *Hebestigma* (88.57–123.62 Ma) and the mean crown (95 Ma) and stem (105 Ma) ages of *Ditta* are not shown in the figure because they fall outside the geological scale.

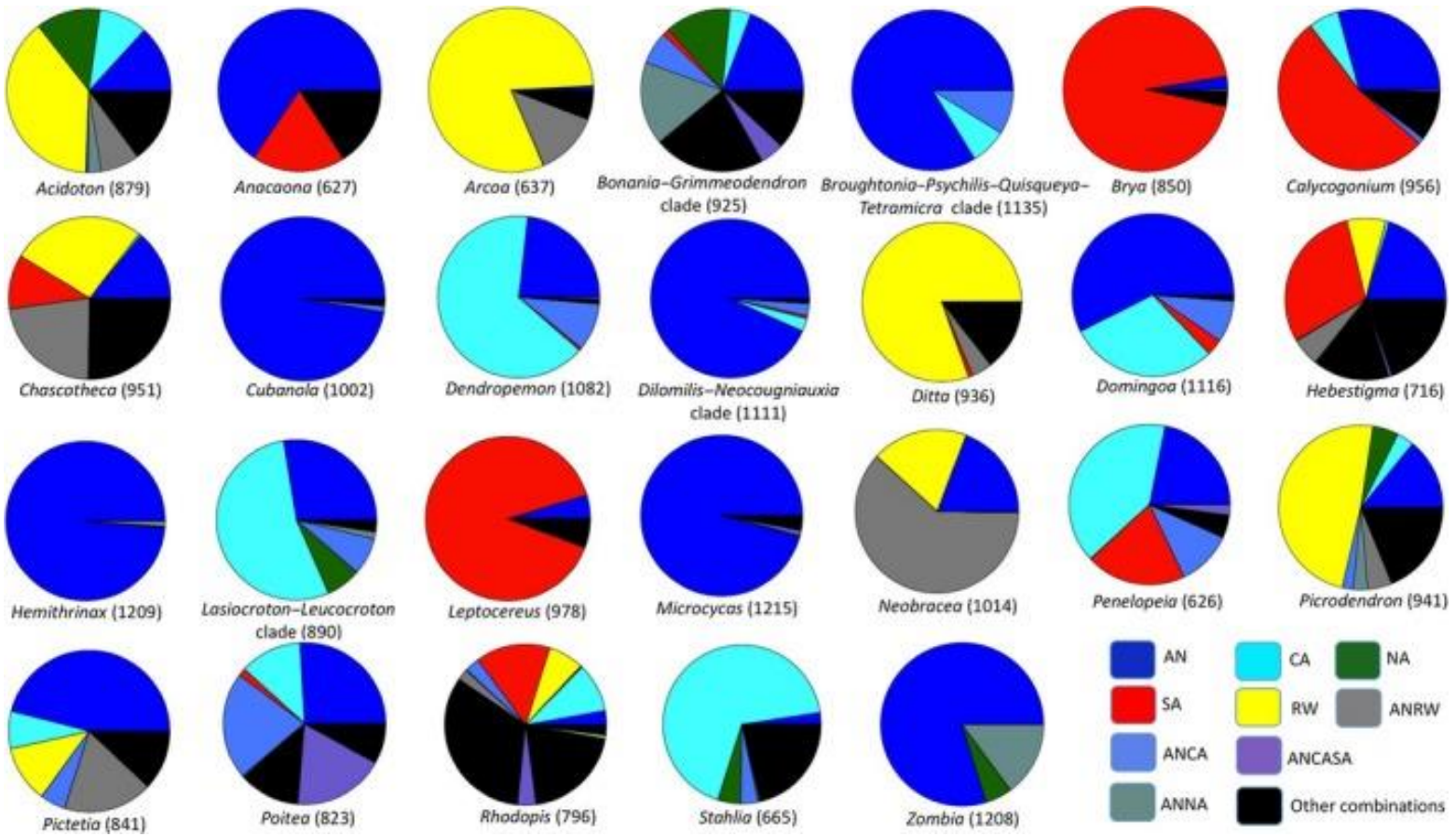


Figure 2.4. Ancestral area estimation for Caribbean endemic genera or clades based on the DEC*j* Complex model 1. Each pie chart contains the likelihood percentage for each estimated area per genus or clade. Numbers in parenthesis are selected nodes that subtend each endemic genus or clade in the tree and that were used for plotting results (same as in Table **2.5**). Ancestors distributed in Antilles and Central America are abbreviated as ANCA; ancestors distributed in Antilles and rest of the world are abbreviated as ANRW; ancestors distributed in Antilles and North America are abbreviated as ANNA; and ancestors distributed in Antilles, Central America, and South America are abbreviated as ANCASA.

CHAPTER 3

Historical biogeography of the conifer genus *Podocarpus* in the Caribbean

3.1. Abstract

The Caribbean region is a biodiversity hotspot and the most species diverse archipelago in the Neotropics. Hypotheses such as vicariance and dispersal have been long discussed as explanations for the origin of Caribbean biota. An alternative hypothesis, the Progression Rule, that older lineages inhabit older islands and colonize newer ones as they emerge, has seldom been tested due to the geological complexity of the Caribbean. I explore the evolutionary history, biogeography, and diversification rates of the conifer genus *Podocarpus* in the Caribbean. I present the most comprehensive sampling for Caribbean *Podocarpus* to date in a Bayesian dated phylogenetic tree with four fossil calibration points and a genotyping by sequencing DNA matrix of 67,589 bp. I used BioGeoBEARS with all models available for all biogeographic inferences. Caribbean *Podocarpus* is the result of a single colonization from South America in the Oligocene (ca. 30 Ma), and the Lesser Antillean species originated from the Greater Antilles in the late Oligocene (ca. 18-21 Ma). Because colonization of the Greater Antilles occurred at the time when the Lesser Antilles were at least partly exposed, the progression rule was not supported. Vicariance can explain the inter-island divergence of Cuban and Hispaniolan species, and Jamaican species originated as a result of dispersals from Cuban and Hispaniolan ancestors. Despite the availability of new resource opportunities, insular *Podocarpus* lineages do not show higher diversification rates than continental taxa.

3.2. Introduction

The Caribbean region is a biodiversity hotspot with high levels of endemism (Mittermeier et al., 2004). Among the biogeographic factors that have contributed to the assembly of the rich Caribbean flora are the complex geological history of the islands, their topographic diversity (which leads to habitat and microclimate heterogeneity), and their proximity to continental America. The Caribbean archipelago, also known as the West Indies, consists of the Greater and Lesser Antilles, the Bahama archipelago, and the islands off the northern coast of Venezuela. The time and geological processes that led to the formation of the Greater and Lesser Antilles are distinct: see Graham (2003) and Roncal et al. (in press) for a paleogeographical review of the Antilles. The former originated as a submerged to subaerially exposed chain of volcanic edifices during the lower Cretaceous, ca. 130-110 Ma (million years ago) (Pindell & Kennan, 2009). This chain, known as the proto-Antilles, drifted northeastwards as the oceanic crust that separated the two Americas was subducted. Present day Cuba and Hispaniola, along with Puerto Rico, formed a unit (Mann et al., 1991). Puerto Rico split first from the proto-Antilles in the Oligocene/early Miocene, and subsequently western and northern Hispaniola separated from Cuba in the middle Miocene (Graham, 2003). The southern portion of Hispaniola island joined northern Hispaniola during the middle Miocene (ca. 15 Ma) (Mann et al., 1991). Jamaica also originated during the Cretaceous as part of the volcanic chain. The island was submerged between 42 and 10 Ma (Lewis & Draper, 1990), and there is no evidence of land connection after the Miocene (Graham, 2003). However, some authors have suggested that north and northeastern Jamaica emerged

during the early Miocene (Robinson, 1971; Buskirk, 1985). Also, rock evidence of terrestrial origin dates to the middle Eocene to late Miocene (Lewis & Draper, 1990).

The Lesser Antilles originated as a volcanic chain product of a subduction of the South American Plate under the Caribbean Plate (Macdonald et al., 2000). The system is divided into a northern part (north of Martinique), which dates from the middle Eocene (ca. 47-38 Ma), and a southern part (from Martinique to northern South America), which dates from the Oligocene (ca. 34-23 Ma). The northern part of the Lesser Antilles is further divided in an eastern arc dating from the Eocene to Oligocene, and a western arc dating from the Miocene (Graham, 2003; Macdonald et al., 2000). One of the outstanding issues in Caribbean geology is the hypothesized existence of a land bridge known as GAARlandia (Greater Antilles+Aves Ridge; Iturralde-Vinent & MacPhee, 1999), which is hypothesized to have connected the Antilles with northern South America during the early Oligocene (35-33 Ma), so as to facilitate migration between these two regions.

In light of the different geological ages for the Greater and Lesser Antilles, a hypothesis that has been extensively explored in oceanic island systems is the progression rule (Hennig, 1966). The rule states that patterns of colonization and diversification of insular organisms are closely linked to geological history, with older taxa found in older islands and more recent taxa formed by subsequent colonization of younger islands as they emerge (Funk & Wagner, 1995). Support for the progression rule comes from Hawai'ian animals and plants (e.g. Hawaiian silverswords, Baldwin & Robichaux, 1995; *Psychotria*, Nepokroeff et al., 2003; lobeliads, Givnish et al., 2009). The progression rule has also been observed in the Canary Islands, an archipelago close to the African continent, at a similar distance as that between the Antilles and continental South

America [e.g arachnids (López-Mercader, 2005; Macías-Hernández et al., 2008; Planas & Ribera, 2014); beetles (Faria et al., 2016)]. Evidence for progression is accumulating for many other archipelagos such as the Galapagos, the Australs, and the Marquesas (Shaw & Gillespie, 2016). However, the Caribbean remains largely unexplored mainly because of its complex geological history. By means of the PACT approach (phylogenetic analysis for comparing trees), Eckstut et al. (2011) explored the progression rule in the Greater Antilles for plants and animals. They concluded that some clades showed progression rule patterns in the Caribbean, while others showed taxon pulse dynamic, a complex pattern of biotic expansions alternating with episodes of *in situ* speciation (Erwin, 1981). Their study did not include the Lesser Antilles; therefore, the progression rule could not be tested in the Caribbean system as a whole.

The formation of islands represents opportunities for new colonizers to exploit new habitats and resources (Losos & Ricklefs, 2009), which might lead to increased lineage diversification rates (Bellemain & Ricklefs, 2008). Even though there are well-documented examples of adaptive radiations in the Caribbean (e.g. *Lyonia*, Judd, 2001; *Anolis*, Losos & Thorpe, 2004; *Cocothrinax*, Cano et al., 2018), there are few empirical studies that have compared diversification rates between Caribbean and continental taxa (*Cocothrinax*, Cano et al., 2018). Very few studies have explored diversification rate shifts for Caribbean plant clades (e.g. *Spathelia*, Appelhans et al., 2012; *Coccothrinax*, Baker & Couvreur, 2013 and Cano et al., 2018).

In recent years, efforts to explain the evolution and assembly of Caribbean endemic flora (e.g. Appelhans et al., 2012; Cervantes et al., 2016; Regalado et al., 2017; Cano et al., 2018) have shown a predominance of Oligocene to Miocene colonizations and *in situ*

speciation that occurred in the last 30 Ma , but not of proto-Antillean vicariance during the late Cretaceous (Nieto-Blázquez et al., 2017; Roncal et al., in press). Proto-Antillean vicariance proposes that fragments of ancient islands, that were situated between North and South America, carried ancient biota as the Caribbean plate drifted eastward in the late Cretaceous (Rosen, 1975). These studies also showed that South and Central America were the main sources of Caribbean ancestors, which agrees with the taxonomic affinities demonstrated in floristic studies (Acevedo-Rodríguez & Strong, 2008).

Here I use the genus *Podocarpus* L'Hér. ex Pers. as a case study to investigate the evolutionary and biogeographic history of endemic plants in the Caribbean. *Podocarpus* is the most speciose and widespread genus of the plant family Podocarpaceae, which is morphologically and ecologically the most diverse family of conifers (Kelch, 1998). Podocarpaceae has a mainly tropical distribution, however, it is an important element of the southern hemisphere temperate forest. Fossil evidence has shown that Podocarpaceae has an origin in Gondwanaland dating from at least the Jurassic, and diversified through the Cretaceous and earliest Cenozoic (Morley, 2011). The fossil record shows the presence of Podocarpaceae in northern South America since the late Eocene to Oligocene (van der Hammen & Hooghiemstra, 2000).

Podocarpus comprises 31 species in the Neotropics (Figure 3.1; Dalling et al., 2011), nine of which occur in the Caribbean (Mill, 2015a). All eight Greater Antilles species are endemic to single islands, and one, *Podocarpus coriaceus* Rich & A. Richin is found in Puerto Rico and the Lesser Antilles. Mill (2015a) did not include a tenth species, *P. trinitensis*, in his review of Caribbean species, because Trinidad and Tobago properly belong to the Orinoco bioregion. I include *P. trinitensis* here because of its

closer geographic proximity to *P. coriaceus* than to other *Podocarpus* species in South America (Figure 3.2).

Previous molecular phylogenetic studies based on a few DNA regions of the chloroplast and nuclear genomes suggest that Antillean podocarps are paraphyletic, albeit with low support (Leslie et al., 2012; Little et al., 2013; Quiroga et al., 2016), such that the Greater and Lesser Antillean species occur in two different clades (Quiroga et al., 2016; Leslie et al., 2018). Mill (2003) proposed that a strong geographic barrier exists between the western Greater Antilles (Cuba, Hispaniola, Jamaica) and Puerto Rico, based on species taxonomic relationships and distributions. Stark Schilling (2004) hypothesized that colonization of the Lesser Antilles and Puerto Rico by *P. coriaceus*, and of the Greater Antilles by the remaining Caribbean species, are unrelated events, based on anatomical and molecular DNA studies.

The phylogenetic relationships amongst Antillean species, their diversification times, and the sequence of inter-island and island-continent colonization events are yet to be explored in depth. Next-generation sequencing (NGS) methods with reduced genome representation libraries such as genotyping by sequencing (GBS, Elshire et al., 2011) have yielded hundreds to thousands of DNA loci to unravel the evolutionary history of plants at genus and species levels (e.g. Wong et al., 2015; Hamon et al., 2017; Alam et al., 2018). I present here a GBS-based, dated phylogenetic tree that comprises about 65% of Neotropical *Podocarpus* species, including 90% of Caribbean species. This is the most comprehensive analysis of Caribbean species to date and the first to use high throughput sequencing technology in *Podocarpus*. I conduct biogeographic analyses to infer the ancestral ranges of Caribbean *Podocarpus* and inter-island colonization patterns, in order

to test the progression rule in the Caribbean. I also use the dated phylogenetic tree to compare diversification rates between insular and continental taxa. My specific questions are: (1) when did Caribbean *Podocarpus* colonize and diversify in the Antilles?; (2) what was the most likely distribution of Caribbean ancestors of *Podocarpus*?; (3) are Greater and Lesser Antillean *Podocarpus* the result of two independent colonization events: the first one into the Greater Antilles older than the second one into the Lesser Antilles as predicted by the progression rule?; and (4) do insular *Podocarpus* have higher diversification rates compared to continental taxa as expected from the new resource opportunities that insular colonists face, or does *Podocarpus* exhibit diversification rate shifts unrelated to its geographical distribution?

3.3. Methods

3.3.1. Sample collection and DNA isolation

Sampling included 29 *Podocarpus* accessions representing 27 species from both subgenera (*Podocarpus* de Laub. and *Foliolatus* de Laub.) and all major clades as recovered in Quiroga et al. (2016) (e.g. Asian, Austral, African and, tropical and subtropical South American clades). I obtained silica-dried leaf samples from living botanical collections at the Royal Botanic Garden Edinburgh (UK), Montgomery Botanical Center (USA), Royal Botanic Gardens Kew (DNA aliquots, UK), and field collections (Table 3.1). Each *Podocarpus* species was represented by a single accession, except for the widespread *P. oleifolius* D. Don, for which three samples were included (Colombia, Costa Rica and Bolivia). My sampling included all Antillean species except

for the eastern Cuban *P. victorinianus* Carabia. Outgroup consisted of seven species across six Podocarpaceae genera (Table 3.1).

I conducted DNA isolation following the standard protocol of the DNeasy Plant MiniKit (Qiagen, Valencia, California, USA) using 30-40 mg of plant tissue with the following modifications: 1) an increase of AP1 buffer for cell lysis from 600 to 750µl; 2) an increase of cell lysis incubation time to 60 minutes; and 3) an increase of P3 buffer used for the precipitation of polysaccharides, detergent and proteins from 195 to 225µl. I diluted DNA extractions in EB buffer to a concentration of 20 ng/µl.

3.3.2. Genotyping by sequencing library preparation and loci selection

The Institut de Biologie Intégrative et de Systèmes (IBIS) of the Université Laval in Canada conducted the GBS. Library preparation and sequencing followed the protocol of Abed et al. (2019). Genomic libraries were prepared for the 36 DNA samples using two restriction enzymes, *SbfI* (high fidelity) and *MspI* (New England BioLabs Inc., Ipswich, MA). Unique barcodes of a length 10-12 bp were added to each sample to facilitate posterior demultiplexing. Single-end sequencing reads of variable length (up to 200 bp) were obtained using 2 chips of an Ion Proton system, producing a raw data FASTQ file of 46.41 GB. For data quality assurance, I used FastQC (Banraham Bioinformatics, Cambridge, England) for high throughput sequence data, where Phred quality score and % GC content were inspected.

I demultiplexed sequencing reads using the *process_radtags* function from *Stacks* v1.47 (Catchen et al., 2013). I trimmed reads at 92 bp length, and removed uncalled reads and reads with low quality scores (phred score of 10). I processed sample reads using

ipyrad v0.7.28 (Eaton & Overcast, 2016) to conduct a *de novo* assembly of loci. *ipyrad* is suitable for phylogenetic studies that includes divergent taxa since it allows indels and lower similarity thresholds across loci (Eaton, 2014). I used the *vclust* function as implemented in VSEARCH (Edgar, 2010) to cluster reads using an 85% similarity threshold within each sample. I excluded clusters that contain less than six reads. To test the effect of the minimum number of samples (mns) that must have data for a locus to be processed, I conducted five runs using: mns = 20, mns = 12; mns = 6; mns = 4; and mns = 2. I used a custom python script to extract loci present in samples within targeted groups (e.g. outgroup, Asian clade, African species, Austral species, tropical species, sub-tropical species, and Caribbean species). The script also concatenated the selected loci in a final matrix used for phylogenetic analyses.

3.3.3. Phylogenetic and divergence time estimation analyses

To obtain a dated phylogenetic tree, I used the Bayesian method as implemented in BEAST v2.4.7. (Bouckaert et al., 2014). I specified the uncorrelated lognormal (UCLN) relaxed clock (Drummond et al., 2006), and the GTR + G + I nucleotide substitution model. To test the effect of tree prior selection on divergence time estimation, I compared the marginal log-likelihoods of a Yule versus the Birth-Death model. Since the likelihood and branch support of the Yule (likelihood = -120846.72) was higher than that of the Birth-Death model (likelihood = -120874.43), the former was chosen to reconstruct the dated phylogenetic tree. I conducted two independent Markov Chain Monte Carlo (MCMC) runs on the Westgrid's "*Cedar*" cluster (Compute Canada Service) for 200 million generations each, sampling every 10,000th generation.

Despite the abundant fossil pollen records available for *Podocarpus*, identification is limited to the generic level (Hooghiemstra et al., 2006; Morley, 2011). I selected four fossil calibration points using priors with lognormal distribution to account for fossil dating uncertainty. The first was the oldest reliable fossil attributed to *Podocarpus*, namely *P. andiniformis* Berry from Laguna del Hunco flora in Argentina, dated 52.22 ± 0.29 Ma (Wilf, 2012), which was used to constrain the crown of *Podocarpus* (node IV in Fig. 3). The other three fossils were: 1) *Dacrycarpus* sp. from Salamanca Formation, Argentina (64.48 ± 0.59, Iglesias, 2007) used to calibrate the crown of the *Dacrycarpus-Dacrydium* clade (node I in Fig. 3); 2) *Retrophyllum* sp. from Laguna del Hunco flora (52.22 ± 0.22, Wilf, 2012) used to constrain the crown of the *Retrophyllum-Afrocarpus-Nageia* clade (node II in Fig. 3); and 3) *Nageia hainanensis* from China (34-55 Ma, Jin et al., 2010) used to constrain the crown of the *Afrocarpus-Nageia* clade (node III in Fig. 3). Specific calibration parameters can be found in Table 3.2. Additionally, monophyly was imposed on five clades to aid finding the tree topology with deep-level relationships in agreement with previous studies (Knopf et al., 2012; Leslie et al., 2012; Little et al., 2013; Quiroga et al., 2016). The five clades were: 1) Asian *Podocarpus*; 2) non-Austral *Podocarpus*; 3) African-Subtropical South America *Podocarpus*; 4) *Prumnopitys* clade; and 5) non-*Prumnopitys* Podocarpaceae taxa.

I used Tracer v1.6 (Rambaut et al., 2014) for verification of convergence of the two MCMC runs, which showed that all estimated sample sizes (ESS) were above 200. I combined log and tree files with LogCombiner v2.4.8, and used TreeAnnotator v2.4.8

(Bouckaert et al., 2014) with a 20% burn-in to obtain the Maximum Clade Credibility (MCC) tree displaying node heights.

3.3.4. Biogeographical analysis

I defined eight biogeographical areas: Greater Antilles (GA), Lesser Antilles (LA), North-Central Andes (AN), Central America and Chocó (CA), southern South America (AU), Mata Atlántica (MA), Africa (AF), and Asia (AS) (inset Figure 3.4). I compiled species distributions from the Global Biodiversity Information Facility (GBIF, http://www.gbif.org/, accessed on 10th June 2018). I corroborated species occurrences using the literature (Mill, 2015a, 2015b; Farjon, 2017) and excluded records from oceans and cultivated specimens from botanical garden collections.

To infer the ancestral areas of Caribbean *Podocarpus* I used BioGeoBEARS v1.1.1 (Matzke, 2018) in R v3.3.1. BioGeoBEARS implements different models, such as dispersal, extinction, and cladogenesis (DEC; Ree et al., 2005, Ree & Smith, 2008), dispersal–vicariance (DIVA; Ronquist, 1997), and Bayesian biogeographic inference (BayArea; Landis et al., 2013) to infer the biogeographic histories of taxa. It also incorporates the *j* parameter which accounts for founder-event speciation, which is potentially relevant for island-continent systems. I used the dated MCC tree obtained from BEAST and set a maximum number of ancestral areas at nodes to two because this is the maximum number of biogeographic areas that any *Podocarpus* species currently occupies. I compared two dispersal models, a null analysis with no time stratification and equal dispersal probabilities amongst regions, and a second complex model using time periods and a dispersal probability matrix. Dispersal probabilities ranged from 1 for

contiguous areas to 0 when areas were not formed. Intermediate probabilities (0.5, 0.1 and 0.01) were assigned depending on distance between areas and the presence of dispersal barriers such as the ocean (Table 3.3).

Five time periods were defined based on geological history: (1) 0-15 Ma: I set a dispersal probability of 0.5 between the Greater and Lesser Antilles (and for time periods 2-4). Dispersal between the Antilles and North-Central Andes was set to 0.5. I gave maximum dispersal probability of 1 between Central America and North-Central Andes to indicate the connection of these two landmasses (Montes et al., 2015; Jaramillo et al., 2017). Dispersal probability between Mata Atlántica and North-Central Andes was set to 0.5, and to 0.1 with Greater Antilles, Lesser Antilles and Central America reflecting the distance to these areas. Dispersal between Africa and Asia with continental America and the Caribbean was set to the minimum 0.01; (2) 15-33 Ma: I reduced the dispersal probability between Central America and North-Central Andes to 0.1 to account for the pre-closure of the Panama Isthmus (Montes et al., 2015). I kept the remaining dispersal probabilities as in the first period; (3) 33-35 Ma: I increased the dispersal probability between North-Central Andes and the Antilles to 1 to account for the hypothesized GAARlandia land bridge (Iturralde-Vinent & MacPhee, 1999) that might have facilitated biotic interchange; (4) 35-50 Ma: I decreased the dispersal probability back to 0.5 between North-Central Andes and the Antilles to indicate the nonexistence of GAARlandia. I increased the dispersal probability to 0.1 between Africa and all three South American biogeographic areas because the distance between continents was shorter (Sanmartín, 2011). At last, (5) 50-76 Ma: The dispersal probability between any biogeographical area and the Lesser Antilles or Central America were reduced to 0 since

they were not fully formed (Graham, 2003; Pindell & Kennan, 2009). I increased the dispersal probability to 0.5 between Africa and the South American areas because continents were even closer (Sanmartín, 2011). See Table 3.3a for dispersal matrices. The 12 biogeographic models (null and complex, and each with six inference models) were compared and the best selected using the Akaike information criterion (AIC, Akaike, 1974) in BioGeoBEARS.

I conducted a second ancestral range reconstruction analysis to elucidate the colonization pattern of the Greater Antillean species. Using the *drop.tip*() function from the R package 'phytools' v0.6.44 (Revell, 2012), I pruned the MCC tree to obtain the clade containing the seven Greater Antillean species. Cuba (CU), Jamaica (JA) and Hispaniola (HI) were the operational areas, and the distribution of species in these three areas followed Mill (2015a). I compared three different biogeographical models. First, I ran a null analysis with no time stratification and equal dispersal probabilities amongst islands. Second, I ran a complex model (complex 1) with two time periods: (1) 0-10 Ma: since the three islands were above water (Graham, 2003), I set a probability of 0.5 to all pairwise dispersal events; and (2) 10-21 Ma: according to Graham (2003) Jamaica was submerged most of the period between 10-42 Ma, therefore I set dispersal constrains between Jamaica and Cuba or Hispaniola to 0.000001 and kept 0.5 for the dispersal between Cuba and Hispaniola. I used the constrain of 0.000001 as suggested by the software developer since a dispersal constrain of 0 made impossible to calculate a valid starting likelihood. Lastly, I ran a third model (complex 2) that reflected the potential emergence of north and northeastern Jamaica during the early Miocene (Buskirk, 1985). I used the same two time periods and dispersal probabilities as before but increased the

dispersal probability between Jamaica and the other Greater Antillean islands to 0.25 during the second time frame (10-21 Ma). See Table 3.3b for dispersal matrices for complex models.

3.3.5. Diversification rate analyses

To test the hypothesis of higher diversification rates on island versus continental taxa, I performed a Binary State Speciation and Extinction model (BiSSE) analysis (Maddison et al., 2007), as implemented in the R package 'diversitree' v0.9.8 (FitzJohn, 2012). The BiSSE model estimates rates of character transition (q10 and q01), speciation (λ) and extinction (μ), and assumes that λ and μ follow a birth–death process, and that rates are dependent on a certain character state. I used a reduced version of the MCC tree excluding the outgroup taxa, Asian and African podocarps using the *drop.tip*() function of the R package 'phytools' v0.6.44 (Revell, 2012). Continental species were assigned "0" and insular species "1". I adjusted the analysis for missing taxa (FitzJohn et al., 2009) by using a sampling fraction of 0.65 for continental taxa and 0.9 for insular taxa. Using a Maximum Likelihood (ML) approach, λ, μ, and q were estimated for a total of eight models of increasing complexity in which parameters were modeled to remain equal or to vary between states. Additionally, I performed an analysis without the sampling correction to test the effect of missing taxa on diversification rates estimates. I used AIC scores (Akaike, 1974) to select the best-fit model. I ran a second BiSSE analysis using a MCMC and the sampling correction model with 10,000 generations. The estimated speciation and extinction rates were plotted using a helper function from 'diversitree' (FitzJohn, 2012).

To explore diversification rate shifts across the dated phylogenetic tree, I ran a Bayesian analysis of macroevolutionary mixtures using BAMM v.2.5.0 (Rabosky, 2014). The *setBAMMPriors*() function was applied to get appropriate prior parameters for the dated phylogeny. I ran BAMM for 1 million generations sampling every 1,000th and accounted for incomplete taxon sampling. BAMM output was then analyzed using the R package BAMMtools v2.1.6 (Rabosky, 2014) discarding 20% of the trees as burn-in, and estimated ESS using the R package 'coda' v0.19.2 (Plummer et al., 2006). Net diversification rate through time plots were generated using the *plotRateThroughTime*() function in BAMMtools. Theoretical and practical concerns as to the use of BAMM have been raised (Moore et al., 2016). However, Rabosky et al. (2017) evaluations of Moore's et al. (2016) critiques show that the method is accurate and consistent.

3.4. Results

3.4.1. Genotyping by sequencing and locus selection

The Ion Proton sequencing generated 46.41 GB of raw data containing 149,187,630 reads. After demultiplexing, quality filtering and discarding reads with ambiguous barcodes a total of 121,369,653 reads were retained. The number of reads per sample varied from 192,163 to 17,264,049 with an average of 1,896,400 reads used in *ipyrad*. Following filtering steps, the number of retained loci were 16 for mns20, 405 for mns12, 1,531 for mns6, 5,633 for mns4, and 52,926 loci for mns2. For subsequent analyses, I selected the data matrix mns2 with 4,778,290 bp because exploratory phylogenetic analyses using higher mns values did not give concordant topologies at deep

Podocarpaceae nodes with earlier studies (Biffin et al., 2012; Knopf et al., 2012; Little et al., 2013; Leslie et al., 2018). The final concatenated DNA matrix resulting from the custom python filtering script contained a total of 67,589 bp.

3.4.2. Phylogenetic relationships among Caribbean **Podocarpus** *and divergence time estimation*

Collectively, Caribbean species of *Podocarpus* are not monophyletic. However, all species from the Greater Antilles do form a monophyletic group (PP = 1.0) with crown and stem ages estimated at 20 Ma (95% HPD 15-26 Ma) and 30 Ma (95% HPD 24-34 Ma), respectively. The two species endemic to Hispaniola (*P. buchii* and *P. hispaniolensis*) are a well-supported clade (PP = 0.85) whose sister is *P. urbanii* from Jamaica (PP = 1.0). The other four subclades within the Greater Antilles clade are all well supported (PP > 0.84), except for *P. angustifolius* (Cuba) and *P. purdieanus* (Jamaica) (PP = 0.48).

The two Lesser Antilles species *P. coriaceus* (Lesser Antilles and Puerto Rico) and *P. trinitensis* (Trinidad and Tobago) occur in a second clade, but not as a monophyletic pair. The other four species in this clade are all from South (*P. sellowii*; *P. ballivianensis*) or Central America (*P. guatemalensis*; *P. matudae*) (PP=0.92). Figure 3.3 and Table 3.4 present divergence times for nodes of interest.

The monophyletic Greater Antilles clade and the Lesser Antilles inclusive clade are weakly paired (PP = 0.61) with respect to a third clade of exclusively South American species (PP = 1.0) (Fig. 3.2).

3.4.3. Biogeographical analysis

Model comparison using AIC (Table 3.5) showed that the best-fit biogeographical model was the DEC*j* complex (LnL = -59.01 and AICw = 0.80%) followed by the DIVAREALIKE*j* complex (LnL = -60.41 and AICw = 0.19%). South America appeared as the most likely ancestral area for the Caribbean taxa, with a colonization event from North-Central Andes into the Greater Antilles estimated between 32 and 30 Ma during the early Oligocene (Figure 3.4). I found a single colonization event from the Greater Antilles into the Lesser Antilles between 30 and 26 Ma. My analysis recovered two recolonization events of the American continent from the Lesser Antilles. The first was between 26 and 18 Ma for the ancestor of *P. matudae* and *P. ballivianensis* into Central America first and then into North-Central Andes. The second was for the ancestor of *P. sellowii* and *P. guatemalensis* into North-Central Andes between 18 and 17 Ma. The origin of *P. sellowii* in the Mata Atlántica of Brazil was estimated at 17 Ma. Figure 3.4 and Table 3.4 present the most likely ancestral ranges for nodes of interest.

For the second ancestral range reconstruction (Greater Antillean species only) the best-fit biogeographical model was the DIVALIKE*j* null (LnL = -5.35 and AICw = 0.34%) followed by the DEC*j* null (LnL = -5.85 and AICw = 0.20%). According to the best-fit model, the ancestor of Greater Antillean taxa was most likely distributed in Cuba and Jamaica (PP=0.25), although the reconstruction showed high uncertainty (PP=0.21 for Cuba and Hispaniola, and PP=0.20 for Jamaica and Hispaniola). From this ancestral area, a divergence of two lineages occurred, one in Hispaniola and the other in Cuba. Jamaican *P. urbanii* originated from the Hispaniolan ancestor 17 Ma (95%HPD 11-23

Ma) at the earliest, while *P. purdieanus* originated from the Cuban ancestor at 16 Ma (95% HPD 9-21 Ma) at the earliest (Figure 3.5).

3.4.4. Diversification rates in continental versus *insular taxa*

The BiSSE analysis without the taxonomic correction recovered the equal lm (AIC = 64.56 and AICw = 0.27%) as the best-fit diversification model, which suggests equal speciation and extinction rates between continental and insular taxa. When the taxonomic correction was incorporated, the best-fit model was equal lmq, suggesting equal speciation, extinction and transition rates (AIC = 61.56 and AICw = 0.25%) (Table 3.6). Missing taxa thus had no effect on the estimated speciation and extinction rates associated to continental and insular taxa. The Bayesian exploration of the posterior distribution of parameters showed a different distribution for speciation rates (λ), which was higher for insular than for continental taxa. The Bayesian exploration also showed an equal distribution of extinction rates (μ) between continental and insular taxa (Figure 3.6a).

The BAMM analysis converged as indicated by the ESS values over 200. There were no significant diversification rate shifts across the dated phylogeny, with zero shifts posterior distribution of 0.79, followed by one shift posterior distribution of 0.17. There is a trend of decreasing net speciation rates through time as shown in Figure 3.6b. The slow decrease in speciation rates was from ca. 0.075 to ca. 0.06 lineages per million years (Figure 3.6c).

3.5. Discussion

My research used high-throughput sequencing technology and the most comprehensive taxon sampling to date to elucidate the evolution, historical biogeography, and diversification rates of the conifer genus *Podocarpus* in the Caribbean. This study increased the resolution and support of phylogenetic relationships amongst these taxa including their times of origin. Diversity of *Podocarpus* in the Greater Antilles arises from a single colonization event from South America during the early Oligocene, and the Lesser Antillean species, *P. coriaceous* and *P. trinitensis,* originated from a late Oligocene Greater Antillean ancestor. The two endemic Lesser Antillean species, *P. coriaceous* and *P. trinitensis,* originated in the late Oligocene from a Greater Antillean ancestor. The progression rule hypothesis is not supported Antillean *Podocarpus* in the ground of ages. Although the Greater Antilles taxa are of slightly greater age than those of the Lesser Antilles, at the time of the colonization of the former, island chain of the latter were all already at least partly exposed (Graham, 2003; Macdonald et al., 2000), rather than emerging stepwise so as to be available for successive colonization. This study also shows the role of vicariance and dispersal processes in the diversification of Greater Antillean *Podocarpus*. Diversification rates for these species did not increase after island colonization, and no major shifts in diversification rates were found among species in this clade.

3.5.1. Systematics of Caribbean Podocarpus

This study provides the most comprehensive phylogenetic analysis of Caribbean *Podocarpus* to date (90% of Caribbean species). The sister relationship of the clade

containing Caribbean species (Figure 3.2) and tropical South American species is concordant with previous work (Knopf et al., 2012; Quiroga et al., 2016; Leslie et al., 2018). All Greater Antillean species form a well-supported (PP=1) clade, contrary to Quiroga et al. (2016) in which Greater Antillean species appeared paraphyletic (PP=0.77). Quiroga et al. (2016) also found the Lesser Antillean species (*P. coriaceus* and *P. trinitensis*) to be sisters, with high support (PP=1), and *P. hispaniolensis* branched off first in a clade formed by Greater Antillean and tropical South American species. Their study did not include *P. buchii*, which is sister to *P. hispaniolensis* in the present study (PP=0.85). Contrary to the results here, Biffin et al. (2012) thought that the Greater Antillean species formed a clade except for *P. hispaniolensis*, which grouped with South American, Central American and Lesser Antillean species. The co-occurrence here of Cuban and Hispaniolan species as sister clades is concordant with the within-island speciation pattern reported in *Anolis* lizards, where species are more closely related to other species from the same island (Losos et al., 1998).

Caribbean species are spread across three of the four sections within subgenus *Podocarpus* as recognized by de Laubenfels (1985): 1) Section Pumilis (*P. ekmanii*, *P. angustifolius*, *P. victorianianus*, *P. urbanii*, *P. buchii*); 2) Section Nemoralis (*P. purdieanus*, *P. guatemalensis*, *P. hispaniolensis*, *P. trinitensis*); and 3) Section Lanceolatus (*P. coriaceus*, *P. matudae*, *P. costaricensis*). de Laubenfels (1985) proposed a sectional circumscription based on external morphological characters such as a groove or a ridge on the upper leaf surface, vegetative buds and their scales, leaf shapes and pollen cones. Results here (Figure 3.2) are concordant with previous molecular systematic studies (Biffin et al., 2011; Knopf et al., 2012) that do not support de Laubenfels (1985)

sections. Thus the molecular evidence agrees with the cuticle micromorphology of Stark Schilling & Mill (2011), who did not find any synapomorphies that support de Laubenfels' (1985) sections.

Taking this and other molecular studies into account, it is clear that Greater Antillean species are phylogenetically more closely related to each other than to *P. coriaceus*, *P. trinitensis* or Central American species, which also explains their similar morphology (Mill, 2015a). However, no synapomorphies for the Greater Antillean species have been identified. The aristate leaf apex has been proposed as a synapomorphy for species in the Greater Antillean taxa, although is poorly developed in *P. hispaniolensis* and *P. urbanii* (Mill, 2015a).

The higher number of *Podocarpus* species in the Greater than Lesser Antilles resembles other plant and animal radiations in the Caribbean (e.g *Anolis*, Thorpe & Losos, 2004; *Spathelia*, Appelhans et al., 2012; *Coccothrinax*, Cano et al., 2018; *Zamia*, Meerow et al., 2018). Time for diversification could be an explanation for this pattern, since the Lesser Antilles is a younger system than the Greater Antilles (Graham, 2003). Another potential explanation is island size (MacArthur & Wilson, 1967; Ricklefs & Bermingham, 2008; Whittaker et al., 2008), which might result in higher extinction rates on smaller islands (e.g. Palmeirim et al., 2018). Island size might also reduce opportunities for cladogenesis in small islands (Emerson & Gillespie, 2008) For example, in the fern genus *Adiantum* (Regalado et al., 2017), Lesser Antillean communities are assembled from new migrants rather than *in situ* speciation. Sympatric speciation is also reduced, at least for birds, in small islands (Coyne & Price, 2000). The Lesser Antilles

may also have fewer *Podocarpus* species due to higher extinction rates (Carson et al., 1990) during recent active volcanism (Macdonald et al., 2000).

3.5.2. Oligocene colonization and early Miocene diversification of **Podocarpus** *in the Caribbean*

The ancestor of Caribbean *Podocarpus* colonized the Greater Antilles in the Oligocene and diversified during the Miocene. In a study on *Limia* fish, Weaver et al. (2016) argued for the importance of global climate in the Eocene-Oligocene transition (30-35 Ma, Iturralde-Vinent & MacPhee, 1999; Iturralde-Vinent, 2006). According to these authors, the transition from the warmer Eocene climate to a cooler Oligocene climate, along with a drop in sea level, played a role on the evolution of organisms globally.

The divergences of *P. trinitensis* and *P. coriaceus* occurred at a time when the Lesser Antilles, but not Puerto Rico, were fully emerged (Graham, 2003). Although the phylogenetic pattern supports a progression rule scenario (i.e. Lesser Antillean species originated from a Greater Antillean ancestor), divergence times do not support this hypothesis. My results show that colonization of the Greater Antilles between 32 to 30 Ma occurred after the Lesser Antilles start forming in the north (ca. 47-38 Ma) and south (ca. 37-34 Ma) (Graham, 2003). Therefore, the Lesser Antilles (or part of them) were above water when the first *Podocarpus* ancestor arrived in the Caribbean, and the necessary circumstances for progression rule were absent. My results are partly concordant with those of Eckstut et al. (2011), whose meta-analysis of plant and animal taxa for the Greater Antilles shows that inter-island relationships have been produced in a

progression rule manner for certain clades. However, their conclusions were based on the inference of biotic expansions between islands, and number of lineages accumulated in relation to island age, rather than on testing of historical biogeographic models. The geological complexity of the Caribbean has therefore hindered evaluation of this hypothesis. More refined paleogeographic knowledge of the Lesser Antilles will allow further testing in multiple lineages.

3.5.3. South American origin of Caribbean Podocarpus

My results are concordant with recent biogeographical meta-analyses that suggest South America was an important source for Caribbean lineages (Nieto-Blázquez et al., 2017, Antonelli et al., 2018; Roncal et al., in press.). The most likely ancestral area for Caribbean *Podocarpus* was the region labelled North-Central Andes (AN) in south America. This is not surprising because of the proximity of this continental mass, and the high dispersal capacity of *Podocarpus*. Island colonization during the Oligocene likely involved over-water dispersal by birds. Studies have shown that several bird families feed on the fleshy female cones of South American *P. parlatorei* (Blendinger, 2017). It is likely that Caribbean *Podocarpus* are also bird dispersed based on the similar morphology of the female fleshy cones. Bird-mediated colonization of the Antilles from South America was also proposed for *Brunfelsia* (Solanaceae), as its Antillean taxa have fleshy bright capsules (Filipowicz & Renner, 2012). Hedges (2006) highlights the importance of water currents coming from the east across northeastern South America into the Caribbean Sea, which may transport non-flying vertebrate organisms. Additional support for a South American origin of Caribbean *Podocarpus* comes from the floristic affinities

between these two regions, and the composition of fossil biota recorded from Miocene deposits from Cuba, Hispaniola, and Puerto Rico with a distinctive South American origin (Borhidi, 1991; Iturralde-Vinent, 2006). The GAARlandia hypothesis to explain colonization of the Antilles from South America cannot be discarded for *Podocarpus* since the stem age of the Greater Antillean clade (95% HPD 23-34 Ma) falls within the hypothesized existence of this land bridge.

3.5.4. Inter-island and continental recolonization patterns

The biogeographic reconstruction for the Greater Antillean clade suggests a Cuban-Jamaican ancestor at the root of the tree (Figure 3.5). This result should be taken with caution, as Jamaica was likely submerged at this time (Lewis & Draper, 1990) and the ancestral area reconstruction shows a high degree of uncertainty. Based on the inferred divergence times and the tree topology, a vicariant explanation can be invoked for the origin of Cuban and Hispaniolan species. The most recent common ancestor (MRCA) for the clades containing Cuban and Hispaniolan species date back ca. 17.5 and 16.7 Ma, respectively. These ages match the proposed time for the split between western Hispaniola and Cuba during the middle Miocene (Graham, 2003) that opened the Windward Passage between both islands. The next most likely ancestral area for Greater Antillean species is Cuba and Hispaniola, assuming a vicariance scenario. After an origin by vicariance, within-island speciation may have given rise to the rest of species in Cuba and Hispaniola. Examples of lineages that also fit the vicariance hypothesis for these two islands include *Calisto* butterflies (Matos-Maraví et al., 2014) and *Limia* fishes (Weaver et al., 2016).

The crown age of Hispaniola species coincides with the estimated time of the north / south palaeo-island collision during the middle Miocene (ca. 15 Ma; Graham, 2003), and this event might have triggered within-island diversification as suggested for Hispaniolan animal taxa (e.g. crickets, Oneal et al., 2010; birds, Sly et al., 2011). The Puerto Rican species *P. coriaceus* should appear sister to the Cuban-Hispaniolan species, as Puerto Rico split from the proto-Antilles in the Oligocene-early Miocene, prior to the Cuba-Hispaniola split. If so, *P. coriaceus* should have diverged early: the results did not support this prediction. Examples within the Caribbean where vicariance amongst islands might have played a role in diversification include the spider genus *Deinops* (Chamberland et al., 2018), and the cycad genus *Zamia* (Meerow et al., 2018), both for the Hispaniola and Puerto Rico break up. The *Zamia* study by Meerow et al. (2018) does not explicitly imply vicariance for the divergence of Hispaniolan-Puerto Rican lineages, but show evidence of early admixture which would indicate common ancestry of populations.

The ancestral reconstruction confidently showed that the Jamaican species *P. urbanii* and *P. purdieanus* originated from Hispaniolan and Cuban ancestors, respectively. However, the time of colonization might be too old if we consider the earliest inferred ages of 16-17 Ma, because Jamaica was likely submerged from the middle Eocene to late Miocene (42-10 Ma) (Graham, 2003). However, there is some evidence that the north and northeastern parts of Jamaica were emergent during the early Miocene (Robinson, 1971; Buskirk, 1985), making two independent colonization events from Cuba and Hispaniola possible.

Colonization of the Lesser Antilles occurred from a Greater Antillean ancestor during the mid-late Oligocene. Since the Lesser Antilles have never been connected to the continent or to the Greater Antilles, this colonization is necessarily the result of over-water dispersal. The pattern of larger islands (i.e. Greater Antilles in this case) acting as source for the colonization of smaller islands (Ricklefs & Bermingham, 2008) has been documented in the spider genus *Selenops* in the Caribbean (Crews et al., 2010). Recolonization of the South American continent from the Lesser Antilles has been documented in Caribbean angiosperms (Cano et al., 2018; Nieto-Blázquez et al., 2017). Continental recolonization has been shown for *Podocarpus* in the Australasian region (Condamine et al., 2017), which reinforces the idea of islands not just as sinks but also sources of biodiversity.

3.5.5. Equal diversification rates for continental and insular taxa

I found no association between island colonization and higher diversification rates in the maximum likelihood BiSSE analysis. The wider diversity of resources and habitat heterogeneity available to continental species might lead to higher diversification rates and potentially mask signals of increased diversification rates for insular *Podocarpus*. However, the Bayesian posterior distribution of parameters from BiSSE showed higher speciation rates for insular *Podocarpus*, and equal extinction rates on the continent and in the islands. This contrasts with comparisons of conifers from New Zealand and New Caledonia, where higher speciation and extinction rates in continental *versus* insular taxa were found with the BiSSE model (Condamine et al., 2017).

In a geologically dynamic archipelago such as the Caribbean, new ecological opportunities for successful island colonizers might be linked to shifts in diversification rates. Global sea level drops during the Oligocene glaciation (Houben et al., 2012) and the climatic changes of the middle Miocene climatic optimum (MMCO) have also been invoked as triggers for shifts in diversification rates. However, the BAMM analysis did not identify any such shifts either in Caribbean *Podocarpus*, or elsewhere across the phylogenetic tree. In this, *Podocarpus* is concordant with other Caribbean plant and fungal taxa where no shifts in diversification rates associated with island colonization have been found (e.g. Myrtaceae, Vasconcelos et al., 2017; *Coccothrinax*, Cano et al., 2018; *Sticta*, Widhelm et al., 2018; *Amphilophium*, Thode et al., 2019).

The absence of shifts in diversification rate may be explained by diversity-dependent processes (Rabosky, 2009). Although niche diversity and ecological opportunities on islands are predicted to increase cladogenesis (Losos, 2010), *Podocarpus* may have arrived in the Caribbean when niche spaces were already filled, preventing further diversification. As well, although extinction rates are generally expected to be high on islands (e.g. Warren et al., 2015), in *Podocarpus* the evidence instead suggests long-term persistence of endemic species, where populations divergence could be a long-term strategy in contrast with species diversification. This tendency of long-term persistence has been shown for the subtropical *P. parlatorei* in South America (Quiroga & Premoli, 2007). The diversification rate slowdown might also be due to incomplete phylogenetic sampling. Cusimano & Renner (2010) suggest that at least 80% of extant taxa for a particular clade should be sampled for an accurate estimation of diversification rates. While this holds true for the Caribbean clade, the 80% sampling is not achieved in

the rest of the phylogeny. Moen & Morlon (2014) proposed additional explanations for the diversification rate slowdown related to time dependent processes, and protracted speciation (Etienne & Rosindell, 2012)(i.e. underestimation of branching events near the tips of a phylogeny due to the gradual, rather than instantaneously, process of speciation).

3.6. References

Abed A., Légaré G., Pomerleau S., St-Cyr J., Boyle B., & Belzile F. (2019) Genotyping-by-Sequencing on the Ion Torrent Platform in Barley: Methods and Protocols. *Methods in Molecular Biology* (ed. by W. Hardwood), pp. 233–252. Humana Press, New York, NY.

Acevedo-Rodríguez P. & Strong M.T. (2008) Floristic richness and affinities in the West Indies. *Botanical Review*, **74**, 5–36.

Akaike H. (1974) A New Look at the Statistical Model Identification. *IEEE Transactions on Automatic Control*, **19**, 716–723.

Alam Z., Roncal J., & Peña-Castillo L. (2018) Genetic variation associated with healthy traits and environmental conditions in *Vaccinium vitis-idaea*. *BMC Genomics*, **19**, 4.

Antonelli A., Zizka A., Carvalho F.A., Scharn R., Bacon C.D., Silvestro D., & Condamine F.L. (2018) Amazonia is the primary source of Neotropical biodiversity. *Proceedings of the National Academy of Sciences*, **115**, 6034 LP-6039.

Appelhans M.S., Keßler P.J.A., Smets E., Razafimandimbison S.G., & Janssens S.B. (2012) Age and historical biogeography of the pantropically distributed Spathelioideae (Rutaceae, Sapindales). *Journal of Biogeography*, **39**, 1235–1250.

Baker W.J. & Couvreur T.L.P. (2013) Global biogeography and diversification of palms sheds light on the evolution of tropical lineages. II. Diversification history and origin of regional assemblages. *Journal of Biogeography*, **40**, 286–298.

Baldwin B.G. & Robichaux R.H. (1995) Historical biogeography and ecology of the Hawaiian silversword alliance (Asteraceae): New molecular phylogenetic perspectives. *Hawaiian Biogeography: Evolution on a hot spot archipelago* (ed. by W.L. Wagner and V.A. Funk), pp. 259–287. Smithsonian Institution Press, Washington D.C.

Bellemain E. & Ricklefs R.E. (2008) Are islands the end of the colonization road? *Trends in Ecology and Evolution*, **23**, 1–8.

Biffin E., Brodribb T.J., Hill R.S., Thomas P., & Lowe A.J. (2012) Leaf evolution in Southern Hemisphere conifers tracks the angiosperm ecological radiation Leaf evolution in Southern Hemisphere conifers tracks the angiosperm ecological radiation. *Proceedings of the Royal Society B: Biological Sciences*, **279**, 341–348.

Biffin E., Conran J.G., & Lowe A.J. (2011) Podocarp Evolution: A Molecular Phylogenetic Perspective. *Ecology of the Podocarpaceae in Tropical Forests* (ed. by B.L. Turner and L.A. Cernusak), pp. 1–20. Smithsonian Institution Scholarly Press, Washington D.C.

Blendinger P.G. (2017) Functional Equivalence in Seed Dispersal Effectiveness of *Podocarpus parlatorei* in Andean Fruit-Eating Bird Assemblages. *Frontiers in Ecology and Evolution*, **5**, 57.

Borhidi A. (1991) *Phytogeography and vegetation ecology of Cuba.* Akadémiai Kiadó, Budapest.

Bouckaert R., Heled J., Kühnert D., Vaughan T., Wu C.-H., Xie D., Suchard M.A., Rambaut A., & Drummond A.J. (2014) BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. *PLoS Computational Biology*, **10**, e1003537.

Buskirk R.E. (1985) Zoogeographic Patterns and Tectonic History of Jamaica and the

Northern Caribbean. *Journal of Biogeography*, **12**, 445–461.
Cano A., Bacon C.D., Stauffer F.W., Antonelli A., Serrano-Serrano M.L., & Perret M. (2018) The roles of dispersal and mass extinction in shaping palm diversity across the Caribbean. *Journal of Biogeography*, **45**, 1432–1443.
Carson H.L., Lockwood J.P., & Craddock E.M. (1990) Extinction and recolonization of local populations on a growing shield volcano. *Proceedings of the National Academy of Sciences*, **87**, 7055 LP-7057.
Catchen J., Hohenlohe P.A., Bassham S., Amores A., & Cresko W.A. (2013) Stacks: an analysis tool set for population genomics. *Molecular Ecology*, **22**, 3124–3140.
Cervantes A., Fuentes S., Gutiérrez J., Magallón S., & Borsch T. (2016) Successive arrivals since the Miocene shaped the diversity of the Caribbean Acalyphoideae (Euphorbiaceae). *Journal of Biogeography*, **43**, 1773–1785.
Chamberland L., McHugh A., Kechejian S., Binford G.J., Bond J.E., Coddington J., Dolman G., Hamilton C.A., Harvey M.S., Kuntner M., & Agnarsson I. (2018) From Gondwana to GAARlandia: Evolutionary history and biogeography of ogre-faced spiders (*Deinopis*). *Journal of Biogeography*, **45**, 2442–2457.
Cohen K.M., Finney S.C., Gibbard P.L., & Fan J.. (2013) The ICS International Chronostratigraphic Chart. Episodes 36. 199–204.
Condamine F.L., Leslie A.B., & Antonelli A. (2017) Ancient islands acted as refugia and pumps for conifer diversity. *Cladistics*, **33**, 69–92.
Coyne J.A. & Price T.D. (2000) Little evidence for sympatric speciation in island birds. *Evolution*, **54**, 2166–2171.
Crews S.C., Puente-Rolón A.R., Rutstein E., & Gillespie R.G. (2010) Molecular Phylogenetics and Evolution A comparison of populations of island and adjacent mainland species of Caribbean *Selenops* (Araneae: Selenopidae) spiders. *Molecular Phylogenetics and Evolution*, **54**, 970–983.
Cusimano N. & Renner S.S. (2010) Slowdowns in Diversification Rates from Real Phylogenies May Not be Real. *Systematic Biology*, **59**, 458–464.
Dalling J.W., Barkan P., Bellingham J., Healey J.R., & Tanner E.V.J. (2011) Ecology and Distribution of Neotropical Podocarpaceae. *Ecology of the Podocarpaceae in Tropical Forests* (ed. by B.L. Turner and L.A. Cernusak), pp. 43–56. Smithsonian Institution Scholarly Press, Washington D.C.
Drummond A.J., Ho S.Y.W., Phillips M.J., & Rambaut A. (2006) Relaxed phylogengetic and dating with confidence. *PLoS Biology*, **4**, 0699–0710.
Eaton D.A.R. (2014) PyRAD: assembly of de novo RADseq loci for phylogenetic analyses. *Bioinformatics*, **30**, 1844–1849.
Eaton D.A.R. & Overcast E. (2016) Available at: https://ipyrad.readthedocs.io/.
Eckstut M.E., McMahan C.D., Crother B.I., Ancheta J.M., McLennan D.A., & Brooks D.R. (2011) PACT in practice: comparative historical biogeographic patterns and species-area relationships of the Greater Antillean and Hawaiian Island terrestrial biotas. *Global Ecology and Biogeography*, **20**, 545–557.
Edgar R.C. (2010) Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*, **26**, 2460–2461.
Elshire R.J., Glaubitz J.C., Sun Q., Poland J.A., Kawamoto K., Buckler E.S., & Mitchell S.E. (2011) A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High

Diversity Species. *PLoS ONE*, **6**, e19379.
Emerson B.C. & Gillespie R.G. (2008) Phylogenetic analysis of community assembly and structure over space and time. *Trends in Ecology and Evolution*, **23**, 619–630.
Erwin T.C. (1981) Taxon pulses, vicariance, and dispersal: an evolutionary synthesis illustrated by carabid beetles. *Vicariance biogeography: a critique* (ed. by G. Nelson and D.E. Rosen), pp. 159–196. Columbia University Press, New York.
Etienne R.S. & Rosindell J. (2012) Prolonging the Past Counteracts the Pull of the Present: Protracted Speciation Can Explain Observed Slowdowns in Diversification. *Systematic Biology*, **61**, 204.
Faria C.M.A., Machado A., Amorim I.R., Gage M.J.G., Borges P.A. V, & Emerson B.C. (2016) Evidence for multiple founding lineages and genetic admixture in the evolution of species within an oceanic island weevil (Coleoptera, Curculionidae) super-radiation. *Journal of Biogeography*, **43**, 178–191.
Farjon A. (2017) *A handbook of the world's conifers.* Brill, Leiden.
Filipowicz N. & Renner S.S. (2012) *Brunfelsia* (Solanaceae): A genus evenly divided between South America and radiations on Cuba and other Antillean islands. *Molecular Phylogenetics and Evolution*, **64**, 1–11.
FitzJohn R.G. (2012) Diversitree: comparative phylogenetic analyses of diversification in R. *Methods in Ecology and Evolution*, **3**, 1084–1092.
FitzJohn R.G., Maddison W.P., & Otto S.P. (2009) Estimating Trait-Dependent Speciation and Extinction Rates from Incompletely Resolved Phylogenies. *Systematic Biology*, **58**, 595–611.
Funk V.A. & Wagner W.L. (1995) Biogeography Patterns in the Hawaiian Islands. *Hawaiian Biogeography: Evolution on a hot spot archipelago* (ed. by W.L. Wagner and V.A. Funk), pp. 379–419. Smithsonian Institution Institution Press.
Givnish T.J., Millam K.C., Mast A.R., Paterson T.B., Theim T.J., Hipp A.L., Henss J.M., Smith J.F., Wood K.R., & Sytsma K.J. (2009) Origin, adaptive radiation and diversification of the Hawaiian lobeliads (Asterales: Campanulaceae). *Proceedings of the Royal Society B*, **276**, 407–416.
Graham A. (2003) Geohistory models and Cenozoic paleoenvironments of the Caribbean region. *Systematic Botany*, **28**, 378–386.
van der Hammen T. & Hooghiemstra H. (2000) Neogene and Quaternary history of vegetation, climate, and plant diversity in Amazonia. *Quaternary Science Review*, **19**, 725–742.
Hamon P., Grover C.E., Davis A.P., Rakotomalala J., Raharimalala N.E., Albert V.A., Sreenath H.L., Stoffelen P., Mitchell S.E., Couturon E., Hamon S., De Kochko A., Crouzillat D., Rigoreau M., Sumirat U., Akaffou S., & Guyot R. (2017) Genotyping-by-sequencing provides the first well-resolved phylogeny for coffee (*Coffea*) and insights into the evolution of caffeine content in its species GBS coffee phylogeny and the evolution of caffeine content. *Molecular Phylogenetics and Evolution*, **109**, 351–361.
Hedges S.B. (2006) Paleogeography of the Antilles and origin of West Indian terrestrial vertebrates. *Annals of the Missouri Botanical Garden*, **93**, 231–244.
Hennig W. (1966) *Phylogenetic Systematics.* University of Illinois Press,
Hooghiemstra H., Wijninga V.M., & Cleef A.M. (2006) The paleobotanical record of

Colombia: implications for Biogeography and Biodiversity. *Annals of the Missouri Botanical Garden*, **93**, 297–324.

Houben A.J.P., van Mourik C.A., Montanari A., Coccioni R., & Brinkhuis H. (2012) The Eocene – Oligocene transition: Changes in sea level, temperature or both? *Palaeogeography, Palaeoclimatology, Palaeoecology*, **335**–**336**, 75–83.

Iglesias A. (2007) *Estudio paleobotanico, paleoecologico y paleoambiental en secuencias de la Formacion Salamanca, Paleoceno Inferior en el sur de la provincia de Chubut, Patagonia, Argentina.* Universidad Nacional de La Plata,

Iturralde-Vinent M.A. (2006) Meso-Cenozoic Caribbean Paleogeography: Implications for the Historical Biogeography of the Region. *International Geology Review*, **48**, 791–827.

Iturralde-Vinent M. & MacPhee R.D.. (1999) Paleogeography of the Caribbean Region: Implications for Cenozoic Biogeography. *Bulletin of the American Museum of Natural History*, **238**, 1–95.

Jaramillo C., Montes C., Cardona A., Silvestro D., Antonelli A., & Bacon C.D. (2017) Comment (1) on "Formation of the Isthmus of Panama" by O'Dea *et al*. *Science Advances*, **3**, e1602321.

Jin J., Qiu J., Zhu Y., & Kodrul T.M. (2010) First fossil record of the genus *Nageia* (Podocarpaceae) in South China and its phytogeographic implications. *Plant Systematics and Evolution*, **285**, 159–163.

Judd W.S. (2001) Phylogeny and biogeography of *Lyonia* sect. *Lyonia* (Ericaceae). *Biogeography of the West Indies: Patterns and perspectives Patterns and perspectives* (ed. by C.A. Woods and F.E. Sergile), pp. 63–75. CRC Press, Boca Raton.

Kelch D.G. (1998) Phylogeny of Podocarpaceae: Comparison of evidence from morphology and 18s rDNA. *American Journal of Botany*, **85**, 986–996.

Knopf P., Schulz C., Little D.P., Stutzel T., & Stevenson D.W. (2012) Relationships within Podocarpaceae based on DNA sequence, anatomical, morphological, and biogeographical data. *Cladistics*, **28**, 271–299.

Landis M.J., Matzke N.J., Moore B.R., & Huelsenbeck J.P. (2013) Bayesian Analysis of Biogeography when the Number of Areas is Large. *Systematic Biology*, **62**, 789–804.

de Laubenfels D.J. (1985) A taxonomic revision of the genus *Podocarpus*. *Blumea*, **30**, 251–278.

Leslie A.B., Beaulieu J., Holman G., Campbell C.S., Mei W., Raubeson L.R., & Mathews S. (2018) An overview of extant conifer evolution from the perspective of the fossil record -. *American Journal of Botany*, **105**, 1–14.

Leslie A.B., Beaulieu J.M., Rai H.S., Crane P.R., & Donoghue M.J. (2012) Hemisphere-scale differences in conifer evolutionary dynamics. *Proceedings of the National Academy of Sciences*, **109**, 16217–16221.

Lewis J. & Draper G. (1990) Geology and tectonic evolution of the northern Caribbean margin. *The geology of North America, volume H, the Caribbean region* (ed. by G. Dengo and J. Case), pp. 77–140. The Geological Society of America, Boulder.

Little D.P., Knopf P., & Schulz C. (2013) DNA Barcode Identification of Podocarpaceae — The Second Largest Conifer Family. *PLoS ONE*, **8**, e81008.

López-Mercader N. (2005) *Evolutionary processes of the genus Spermophorides (Araneae, Pholcidae) in the Canary Islands.* Universitat de Barcelona, Spain.
Losos J.B. (2010) Adaptive Radiation, Ecological Opportunity, and Evolutionary Determinism. *The American Naturalist*, **175**, 623–639.
Losos J.B., Jackman T.R., Larson A., Queiroz K. de, & Rodríguez-Schettino L. (1998) Contingency and Determinism in Replicated Adaptive Radiations of Island Lizards. *Science*, **279**, 2115–2118.
Losos J.B. & Ricklefs R.E. (2009) Adaptation and diversification on islands. *Nature*, **457**, 830–836.
Losos J.B. & Thorpe R.S. (2004) Evolutionary Diversification of Caribbean Anolis Lizards. *Adaptive Speciation* (ed. by U. Dieckmann, M. Doebeli, J.A.J. Metz, and D. Tautz), pp. 322–324. Cambridge University Press, Cambridge.
MacArthur R.H. & Wilson E.O. (1967) *The theory of island biogeography.* Princeton University Press, Princeton, N.J.
Macdonald R., Hawkesworth C.J., & Heath E. (2000) The Lesser Antilles volcanic chain: a study in arc magmatism. *Earth-Science Reviews*, **49**, 1–76.
Macías-Hernández N., Oromí P., & Arnedo M.A. (2008) Patterns of diversification on old volcanic islands as revealed by the woodlouse-hunter spider genus *Dysdera* (Araneae, Dysderidae) in the eastern Canary Islands. *Biological Journal of the Linnean Society*, **94**, 589–615.
Maddison W.P., Midford P.E., & Otto S.P. (2007) Estimating a Binary Character's Effect on Speciation and Extinction. *Systematic Biology*, **56**, 701–710.
Mann P., Draper G., & Lewis J. (1991) An overview of the geologic and tectonic development of Hispaniola. *Geologic and Tectonic Development of the North American–Caribbean Plate Boundary in Hispaniola* (ed. by P. Mann, G. Draper, and J. Lewis), pp. 1–18. The Geological Society of America, Inc., Boulder, CO.
Matos-Maraví P., Águila R.N., Peña C., Miller J.Y., Sourakov A., & Wahlberg N. (2014) Causes of endemic radiation in the Caribbean: evidence from the historical biogeography and diversification of the butterfly genus *Calisto* (Nymphalidae: Satyrinae: Satyrini). *BMC Evolutionary Biology*, **14**, 1–18.
Matzke N.J. (2018) Available at: https://github.com/nmatzke/BioGeoBEARS.
Meerow A.W., Salas-Leiva D.E., Calonje M., Francisco-Ortega J., Griffith M.P., Nakamura K., Jiménez-Rodríguez F., Lawrus J., & Oberli A. (2018) Contrasting Demographic History and Population Structure of *Zamia* (Cycadales: Zamiaceae) on Six Islands of the Greater Antilles Suggests a Model for Population Diversification in the Caribbean Clade of the Genus. *International Journal of Plant Sciences*, **179**, 730–757.
Mill R.R. (2003) Towards a Biogeography of the Podocarpaceae. *Acta Hortic.*, **65**, 137–147.
Mill R.R. (2015a) A Monographic Revision of the Genus *Podocarpus* (Podocarpaceae): II . The Species of the Caribbean Bioregion*. *Edinburgh Journal of Botany*, **72**, 61–185.
Mill R.R. (2015b) A Monographic Revision of the genus Podocarpus (Podocarpaceae): III. The Species of the Central America and Northern Mexico Bioregions*. *Edinburgh Journal of Botany*, **72**, 243–341.

Mittermeier R.A., Gil R.R., Hoffman M., Pilgrim J., T. B., Mittermeier C.G., Lamoreux J., & da Fonseca G.A.B. (2004) *Hotspots revisited: Earth's biologically richest and most threatened terrestrial ecoregions.* CEMEX, Mexico D.F.

Moen D. & Morlon H. (2014) Why does diversification slow down? *Trends in Ecology and Evolution*, **29**, 190–197.

Montes C., Cardona A., Jaramillo C., Pardo A., Silva J.C., Valencia V., Ayala C., Pérez-Angel L.C., Rodriguez-Parra L.A., Ramirez V., & Niño H. (2015) Middle Miocene closure of the Central American Seaway. *Science*, **348**, 226–229.

Moore B.R., Höhna S., May M.R., Rannala B., & Huelsenbeck J.P. (2016) Critically evaluating the theory and performance of Bayesian analysis of macroevolutionary mixtures. *Proceedings of the National Academy of Sciences*, **113**, 9569 LP-9574.

Morley R.J. (2011) Dispersal and Paleoecology of Tropical Podocarps. *Ecology of the Podocarpaceae in Tropical Forests* (ed. by B.L. Turner and L.A. Cernusak), pp. 21–41. Smithsonian Institution Scholarly Press, Washington D.C.

Nepokroeff M.N., Sytsma K.J., Wagner W.L., & Zimmer E.A. (2003) Reconstructing Ancestral Patterns of Colonization and Dispersal in the Hawaiian Understory Tree Genus *Psychotria* (Rubiaceae): A Comparison of Parsimony and Likelihood Approaches. *Systematic Biology*, **52**, 820–838.

Nieto-Blázquez M.E., Antonelli A., & Roncal J. (2017) Historical Biogeography of endemic seed plant genera in the Caribbean: Did GAARlandia play a role? *Ecology and Evolution*, **7**, 10158–10174.

Oneal E., Otte D., & Knowles L.L. (2010) Testing for biogeographic mechanisms promoting divergence in Caribbean crickets (genus *Amphiacusta*). *Journal of Biogeography*, **37**, 530–540.

Palmeirim A.F., Benchimol M., Vieira M.V., & Peres C.A. (2018) Small mammal responses to Amazonian forest islands are modulated by their forest dependence. *Oecologia*, **178**, 191–204.

Pindell J.L. & Kennan L. (2009) Tectonic evolution of the Gulf of Mexico, Caribbean and northern South America in the mantle reference frame: an update. *Geological Society, London, Special Publications*, **328**, 1–55.

Planas E. & Ribera C. (2014) Uncovering overlooked island diversity: colonization and diversification of the medically important spider genus *Loxosceles* (Arachnida: Sicariidae) on the Canary Islands. *Journal of Biogeography*, **41**, 1255–1266.

Plummer M., Best N., Cowles K., & Vines K. (2006) CODA: convergence diagnosis and output analysis for MCMC. *R News*, **6**, 7–11.

Quiroga M.P., Mathiasen P., Iglesias A., Mill R.R., & Premoli A.C. (2016) Molecular and fossil evidence disentangle the biogeographical history of *Podocarpus*, a key genus in plant geography. *Journal of Biogeography*, **43**, 372–383.

Quiroga M.P. & Premoli A.C. (2007) Genetic patterns in *Podocarpus parlatorei* reveal the long-term persistence of cold-tolerant elements in the southern Yungas. *Journal of Biogeography*, **34**, 447–455.

Rabosky D.L. (2009) Ecological limits and diversification rate: alternative paradigms to explain the variation in species richness among clades and regions. *Ecology Letters*, **12**, 735–743.

Rabosky D.L. (2014) Automatic Detection of Key Innovations, Rate Shifts, and

Diversity-Dependence on Phylogenetic Trees. *PLOS ONE*, **9**, e89543.
Rabosky D.L., Mitchell J.S., & Chang J. (2017) Is BAMM Flawed? Theoretical and Practical Concerns in the Analysis of Multi-Rate Diversification Models. *Systematic Biology*, **66**, 477–498.
Rambaut A., Suchard M.A., Xie D., & Drummond A.J. (2014) Available at: http://beast.bio.ed.ac.uk/Tracer.
Ree R.H., Moore B.R., Webb C.O., & Donoghue M.J. (2005) A likelihood framework for inferring the evolution of geographic range on phylogenetic trees. *Evolution*, **59**, 2299–2311.
Ree R.H. & Smith S.A. (2008) Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. *Systematic Biology*, **57**, 4–14.
Regalado L., Lóriga J., Bechteler J., Beck A., Schneider H., & Heinrichs J. (2017) Phylogenetic biogeography reveals the timing and source areas of the *Adiantum* species (Pteridaceae) in the West Indies, with a special focus on Cuba. *Journal of Biogeography*, **45**, 541–551.
Revell L.J. (2012) phytools: An R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution*, **3**, 217–223.
Ricklefs R. & Bermingham E. (2008) The West Indies as a laboratory of biogeography and evolution. *Philosophical Transactions of the Royal Society B*, **363**, 2393–2413.
Robinson E. (1971) Observation on the Geology of Jamaican Bauxite. 3–9.
Ronquist F. (1997) Dispersal-Vicariance Analysis: A New Approach to the Quantification of Historical Biogeography. *Systematic Biology*, **46**, 195–203.
Rosen D.E. (1975) A Vicariance Model of Caribbean Biogeography. *Systematic Zoology*, **24**, 431–464.
Sanmartín I. (2011) A paleogeographic history of the Southern hemisphere. http://hdl.handle.net/10261/34831. .
Shaw K.L. & Gillespie R.G. (2016) Comparative phylogeography of oceanic archipelagos: Hotspots for inferences of evolutionary process. *Proceedings of the National Academy of Sciences*, **113**, 7986–7993.
Sly N.D., Townsend A.K., Rimmer C.C., Jason M., Latta S.C., & Lovette I.J. (2011) Ancient islands and modern invasions: disparate phylogeographic histories among Hispaniola's endemic birds. *Molecular Ecology*, **20**, 5012–5024.
Stark Schilling D.M. (2004) *Taxonomic studies of Caribbean and Central American species of Podocarpus subgenus Podocarpus: a multi-disciplinary approach.* Royal Botanic Garden Edinburgh/University of Edinburgh, UK.,
Stark Schilling D.M. & Mill R.R. (2011) Cuticle Micromorphology of Caribbean and Central American Species of *Podocarpus*. *International Journal of Plant Sciences*, **172**, 601–631.
Thode V.A., Sanmartín I., & Lohmann L.G. (2019) Molecular Phylogenetics and Evolution Contrasting patterns of diversification between Amazonian and Atlantic forest clades of Neotropical lianas (*Amphilophium*, Bignonieae) inferred from plastid genomic data. *Molecular Phylogenetics and Evolution*, **133**, 92–106.
Thorpe R.S. & Losos J.B. (2004) Concluding Comments. *Adaptive Speciation* (ed. by U. Dieckamnn, M. Doebeli, J. Metz, and D. Tautz), pp. 343–344. Cambridge University

Press, Cambridge.
Vasconcelos T.N.C., Proença C.E.B., Ahmad B., Aguilar D.S., Aguilar R., Amorim B.S., Campbell K., Costa I.R., De-Carvalho P.S., Faria J.E.Q., Giaretta A., Kooij P.W., Lima D.F., Mazine F.F., Peguero B., Prenner G., Santos M.F., Soewarto J., Wingler A., & Lucas E.J. (2017) Myrteae phylogeny, calibration, biogeography and diversification patterns: Increased understanding in the most species rich tribe of Myrtaceae. *Molecular Phylogenetics and Evolution*, **109**, 113–137.
Warren B.H., Simberloff D., Ricklefs R.E., Aguilée R., Condamine F.L., Gravel D., Morlon H., Mouquet N., Rosindell J., Casquet J., Conti E., Cornuault J., Fernández-Palacios J.M., Hengl T., Norder S.J., Rijsdijk K.F., Sanmartín I., Strasberg D., Triantis K.A., Valente L.M., Whittaker R.J., Gillespie R.G., Emerson B.C., & Thébaud C. (2015) Islands as model systems in ecology and evolution: prospects fifty years after MacArthur-Wilson. *Ecology Letters*, **18**, 200–217.
Weaver P.F., Cruz A., Johnson S., Dupin J., & Weaver K.F. (2016) Colonizing the Caribbean: biogeography and evolution of livebearing fishes of the genus *Limia* (Poeciliidae). *Journal of Biogeography*, **43**, 1808–1819.
Whittaker R.J., Triantis K.A., & Ladle R.J. (2008) A general dynamic theory of oceanic island biogeography. *Journal of Biogeography*, **35**, 977–994.
Widhelm T.J., Bertoletti F.R., Asztalos M.J., Mercado-Díaz J.A., Huang J.-P., Moncada B., Lücking R., Magain N., Sérusiaux E., Goffinet B., Crouch N., Mason-Gamer R., & Lumbsch H.T. (2018) Oligocene origin and drivers of diversification in the genus *Sticta* (Lobariaceae, Ascomycota). *Molecular Phylogenetics and Evolution*, **126**, 58–73.
Wilf P. (2012) Rainforest conifers of Eocene Patagonia: attached cones and foliage of the extant southeast Asian and Australasian genus *Dacrycarpus* (Podocarpaceae). *American Journal of Botany*, **99**, 562–584.
Wong M.M.L., Gujaria-Verma N., Ramsay L., Yuan H.Y., Caron C., Diapari M., Vandenberg A., & Bett K.E. (2015) Classification and Characterization of Species within the Genus *Lens* Using Genotyping-by- Sequencing (GBS). *PLoS ONE*, **10**, e0122025.

Table 3.1. List of sampled *Podocarpus* species and related genera, their distribution, voucher information and collection information. RBGE = Royal Botanic Gardens Edinburgh; HAC = Herbario de la Academia de Ciencias, La Habana; JBSD = Jardín Botánico de Santo Domingo; MCB = Montgomery Botanical Center; and KEW = Royal Botanic Gardens Kew.

Species name	Distribution	Voucher information	Collection Information
Podocarpus angustifolius	Cuba	20031689H (RBGE)	Cultivated from material wild collected in Cuba, Sancti Spiritus
Podocarpus ekmanii	Cuba	43222 (HAC)	Wild collected, Cuba, Guantanamo, Yateras
Podocarpus aristulatus	Cuba	43223 (HAC)	Wild collected, Cuba, Guantanamo, Yateras
Podocarpus purdieanus	Jamaica	20011344A (RBGE)	Cultivated from material wild collected in Jamaica, Middlesex
Podocarpus urbanii	Jamaica	20011361A (RBGE)	Cultivated from material wild collected in Jamaica, Surrey
Podocarpus hispaniolensis	Hispaniola	128153 (JBSD)	Wild collected, Dominican Republic, Puerto Plata
Podocarpus buchii	Hispaniola	128145 (JBSD)	Wild collected, Dominican Republic, Independencia
Podocarpus coriaceus	Puerto Rico, Lesser Antilles	19901428A3 (RBGE)	Cultivated from material wild collected in Dominica
Podocarpus trinitensis	Trinidad and Tobago	20030492A1 (RBGE)	Cultivated from material wild collected in Trinidad and Tobago, St Andrew
Podocarpus guatemalensis	Mexico, Guatemala, Belize, Honduras, Nicaragua and Colombia	20140203 (MBC)	Cultivated from material wild collected in Belize
Podocarpus matudae	Mexico, Guatemala, El Salvador, Honduras, Nicaragua, Costa Rica, Panama	19972326A (RBGE)	Cultivated from material wild collected in Mexico, Oaxaca
Podocarpus oleifolius	Bolivia, Ecuador, El Salvador, Guatemala, Honduras, Mexico, Costa Rica, Panama, Colombia, Peru, Venezuela	E00617101 (RBGE)	Wild collected, La Paz, Franz Tamayo
		19923163A3 (RBGE)	Cultivated from material wild collected in Costa Rica
		20001722A2 (RBGE)	Cultivated from material wild collected in Colombia
Podocarpus glomeratus	Ecuador, Peru, Bolivia	E00593344 (RBGE)	Wild collected, Santa Cruz, Manuel María Caballero
Podocarpus parlatorei	Peru, Bolivia, NW Argentina, Chile	Vargas, I. 3036 (LPB)	Wild collected Bolivia, Dept. Chiquisaca, Vallegrande province, El Palmar Reserve
Podocarpus ballivianensis	At least Bolivia	E00593365 (RBGE)	Wild collected, La Paz, Franz Tamayo

Podocarpus ingensis	Bolivia, Peru	E00617086 (RBGE)	Wild collected, La Paz, Sud Yungas
Podocarpus salignus	Chile	Quiroga, M.P. (Individual 4)	Reserva Nacional Nonguén, Concepción, Chile
Podocarpus nubigenus	S. Chile to S. Argentina	Quiroga, M.P. 1 BCRU	Rio Frio, Chile
Podocarpus lambertii	SE. & S. Brazil to Argentina	20091135A (RBGE)	Cultivated from material wild collected in Brazil
Podocarpus sellowii	Brazil	20071743A (RBGE)	Cultivated from material wild collected in Brazil
Podocarpus costalis	Philippines, Taiwan	19763954A3 (RBGE)	Cultivated from material wild collected in Hong Kong, Lantau Island
Podocarpus rubens	Taiwan, Philippines	20000597A (RBGE)	Cultivated from material wild collected in Indonesia, Sulawesi
Podocarpus neriifolius	Nepal to W. & C. Malesia	19950517A1 (RBGE)	Cultivated from material wild collected in Vietnam, Lào Cai
Podocarpus brasii	New Guinea	19661928B (RBGE)	Cultivated from material wild collected in Papua New Guinea, Eastern Highlands
Podocarpus nakaii	Taiwan	19763844A6 (RBGE)	Cultivated from material wild collected in Taiwan, Taipei
Podocarpus milanjianus	Tropical Africa	19340272A (RBGE)	Cultivated from material wild collected. Country not specified
Podocarpus henkelii	Tanzania to Zimbabwe, Cape Prov. to KwaZulu-Natal	19790678 (RBGE)	Cultivated from material wild collected in South Africa, Natal
Retrophyllum rospigliosii	NW. Venezuela to W. South America	19951955 (RBGE)	Cultivated from material wild collected in Venezuela
Prumnopitys harmsiana	W. South America to Venezuela	E00593362 (RBGE)	Wild collected, La Paz, Franz Tamayo
Prumnopitys exigua	Bolivia	E00593349 (RBGE)	Wild collected, Santa Cruz, Manuel María Caballero
Nageia fleuryi	S. China to Indo-China	20020806A (RBGE)	Cultivated from material wild collected in Vietnam
Afrocarpus manii	São Tomé	19960586 (RBGE)	Cultivated from material wild collected in Sao Tomé and Principe, St Thomas Is
Dacrycarpus compactus	New Guinea	13219 (KEW)	DNA aliquot from material wild collected in Indonesia, Papua, Mimika Regency
Dacrydium beccarii	Malesia to Solomon Is.	37524 (KEW)	DNA aliquot from material wild collected in Malaysia, Sarawak

Table 3.2. Four fossil calibration points used in the Bayesian divergence time analysis of *Podocarpus* as implemented in BEAUti and BEAST v2.4.7.

Fossils	Fig. 4 Node #	Age (epoch)	Location	Reference	BEAUti settings
Dacrycarpus sp.	I	Early Paleocene	Salamanca Formantion, Argentina	Iglesias, 2007	Offset = 64.48, Mean = 1.0, *SD* = 0.59
Retrophyllum sp.	III	Early Eocene	Laguna del Hunco, Argentina	Wilf, 2012	Offset = 52.22, Mean = 1.0, *SD* = 0.22
Nageia hainanensis	III	Eocene	China	Jin et al. 2010	Offset = 43.0, Mean = 1.0, *SD* = 6.0
Podocarpus andiniformis	IV	Early Eocene	Laguna del Hunco, Argentina	Wilf, 2012	Offset = 52.22, Mean = 1.0, *SD* = 0.29

Table 3.3. Dispersal matrices used in historical biogeographical analyses as implemented in BioGeoBEARS. a) For the first analysis (full taxon sampling) time periods correspond to: 1) 0-15 Ma; 2) 15-33 Ma; 3) 33-35 Ma; 4) 35-50 Ma; and 5) 50-76 Ma. Biogeographical areas as follows: Greater Antilles (GA), Lesser Antilles (LA), North and Central Andes (AN), Central America and Chocó (CA), Southern South America (AU), Mata Atlántica (MA), Africa (AF) and Asia (AS). b) For the second analysis (greater Antillean clade) time periods correspond to: 1) 0-10 Ma; and 2) 10-21 Ma.

a) First Ancestral Range Reconstruction

0-15 Ma	**GA**	**LA**	**CA**	**SA**	**AU**	**MA**	**AF**	**AS**
GA	-							
LA	0.5	-						
CA	0.5	0.5	-					
SA	0.5	0.5	1	-				
AU	0.1	0.1	0.1	0.5	-			
MA	0.1	0.1	0.1	0.5	0.5	-		
AF	0.01	0.01	0.01	0.01	0.01	0.01	-	
AS	0.01	0.01	0.01	0.01	0.01	0.01	0.01	-

15-33 Ma	**GA**	**LA**	**CA**	**SA**	**AU**	**MA**	**AF**	**AS**
GA	-							
LA	0.5	-						
CA	0.5	0.5	-					
SA	0.5	0.5	0.1	-				
AU	0.1	0.1	0.1	0.5	-			
MA	0.1	0.1	0.1	0.5	0.5	-		
AF	0.01	0.01	0.01	0.01	0.01	0.01	-	
AS	0.01	0.01	0.01	0.01	0.01	0.01	0.01	-

33-35 Ma	**GA**	**LA**	**CA**	**SA**	**AU**	**MA**	**AF**	**AS**
GA	-							
LA	0.5	-						
CA	0.5	0.5	-					
SA	1	1	0.1	-				
AU	0.1	0.1	0.1	0.5	-			
MA	0.1	0.1	0.1	0.5	0.5	-		
AF	0.01	0.01	0.01	0.01	0.01	0.01	-	
AS	0.01	0.01	0.01	0.01	0.01	0.01	0.01	-

35-50 Ma	**GA**	**LA**	**CA**	**SA**	**AU**	**MA**	**AF**	**AS**
GA	-							
LA	0.5	-						
CA	0.5	0.5	-					
SA	0.5	0.5	0.1	-				
AU	0.1	0.1	0.1	0.5	-			
MA	0.1	0.1	0.1	0.5	0.5	-		
AF	0.01	0.01	0.01	0.1	0.1	0.1	-	
AS	0.01	0.01	0.01	0.01	0.01	0.01	0.01	-

50-76 Ma	GA	LA	CA	SA	AU	MA	AF	AS
GA	-							
LA	0	-						
CA	0	0	-					
SA	0.5	0	0	-				
AU	0.1	0	0	0.5	-			
MA	0.1	0	0	0.5	0.5	-		
AF	0.01	0	0	0.5	0.5	0.5	-	
AS	0.01	0	0	0.01	0.01	0.01	0.01	-

b) Second Ancestral Range Reconstruction

Complex 1				Complex 2			
0-10 Ma	**CUBA**	**JAMAICA**	**HISPANIOLA**	**0-10 Ma**	**CUBA**	**JAMAICA**	**HISPANIOLA**
CUBA	-			**CUBA**	-		
JAMAICA	0.5	-		**JAMAICA**	0.5	-	
HISPANIOLA	0.5	0.5	-	**HISPANIOLA**	0.5	0.5	-
10-21 Ma	**CUBA**	**JAMAICA**	**HISPANIOLA**	**10-21 Ma**	**CUBA**	**JAMAICA**	**HISPANIOLA**
CUBA	-			**CUBA**	-		
JAMAICA	0.000001	-		**JAMAICA**	0.25	-	
HISPANIOLA	0.5	0.000001	-	**HISPANIOLA**	0.5	0.25	-

Table 3.4. Divergence times resulting from a Bayesian dating analysis in BEAST v2.4.7 and the most likely ancestral area for each *Podocarpus* ancestor based on the complex DEC*j* model. Node numbers are the same as in Figure 3.3. Nodes 6-11 also show the most likely ancestral area based on the DIVALIKE*j* null model for the Greater Antillean clade. North and Central Andes = AN; Greater Antilles = GA; Lesser Antilles = LA; Central America = CA; Mata Atlantica = MA; Africa = AF; Southern South America = AU; CU = Cuba; HI = Hispaniola; and JA = JA.

Node number	Mean Age (in Million years)	95% HPD	Ancestral reconstruction probabilities (%)
1	50.01	48.39-51.57	AN (0.5)
2	49.62	47.78-51.34	AN (0.79)
3	42.15	37.45-46.43	AN (0.74)
4	31.52	26.12-36.90	AN (0.79)
5	29.96	23.49-34.23	GA (0.66)
6	20.43	15.01-26.01	GA (1) / CUJA (0.24)
7	17.48	11.47-22.52	GA (1) / CU (0.85)
8	16.03	8.95-20.52	GA (1) / CU (0.65)
9	11.81	5.48-17.98	GA (1) / CU (1)
10	16.77	11.12-22.51	GA (1) / HI (0.43)
11	14.56	8.37-20.17	GA (1) / HI (1)
12	25.83	20.10-31.30	LA (0.49)
13	18.04	8.52-25.78	CA (0.56)
14	21.15	15.10-26.93	LA (0.97)
15	18.02	11.88-23.45	LA (0.48)
16	17.28	9.69-21.38	AN (0.64)
17	23.23	15.57-31.21	AN (1)
18	16.1	7.68-24.13	AN (1)
19	40.86	35.67-45.53	MA (0.63)
20	21.32	14.13-28.91	MA (0.70)
21	18.2	11.04-25.58	AN (1)
22	26.4	16.19-36.68	AF (1)
23	25.23	7.21-44.21	AU (1)

Table 3.5. Biogeographic model testing of *Podocarpus* conducted in BioGeoBEARS. The first and second ancestral range reconstructions were conducted for the full taxonomic sampling and the Greater Antillean clade only, respectively. Null models have equal plant dispersal probabilities across all areas and through time. Complex models account for unequal dispersal probabilities considering the geological history. LnL, log likelihood; # params, number of parameters; d, dispersal rate per million years along branches; e, extinction rate per million years along branches; j, founder event speciation weighted per speciation event; AIC, Akaike Information Criterion; AIC wt, relative weight for each model. Best model according to AIC marked with *.

FIRST ANCESTRAL RANGE RECONSTRUCTION

Models	**LnL**	**# params**	**d**	**e**	**j**	**AIC**	**AIC wt**
DEC null	-111.6	2	0.01	1.0E-02	0	227.20	3.15E-23
DEC*j* null	-81.53	3	0.001	1.0E-12	0.0429	169.07	1.32E-10
DIVAREALIKE null	-100.03	2	0.003	5.8E-03	0	204.05	3.36E-18
DIVAREALIKE*j* null	-81.98	3	0.001	1.0E-12	0.0383	169.95	8.51E-11
BAYAREALIKE null	-105.34	2	0.004	1.9E-02	0	214.68	1.65E-20
BAYAREALIKE*j* null	-82.99	3	0	2.5E-03	0.0409	171.98	3.10E-11
DEC complex	-81.3	2	0.009	6.5E-03	0	166.60	4.54E-10
DEC*j* complex*	-59.01	3	0.001	1.0E-12	0.2296	124.02	0.8001
DIVALIKE complex	-84.73	2	0.011	5.8E-03	0	173.46	1.47E-11
DIVALIKE*j* complex	-60.41	3	0.001	1.0E-12	0.2892	126.83	0.1969
BAYAREALIKE complex	-84.93	2	0.011	1.6E-02	0	173.87	1.20E-11
BAYAREALIKE*j* complex	-64.65	3	0.001	1.0E-07	0.2838	135.30	0.0028

SECOND ANCESTRAL RANGE RECONSTRUCTION

Models	LnL	# params	d	e	j	AIC	AIC wt
DEC null	-11.57	2	6.4E-03	4.8E-03	0	27.14	1.84E-03
DEC*j* null	-5.85	3	1.0E-12	1.0E-12	0.232655536	17.70	0.2056
DIVALIKE null	-10.80	2	9.9E-03	7.7E-03	0	25.61	3.95E-03
DIVALIKE*j* null *	-5.36	3	1.0E-12	1.0E-12	0.200449753	16.72	0.3363
BAYAREALIKE null	-13.86	2	1.9E-02	4.4E-02	0	31.73	1.85E-04
BAYAREALIKE*j* null	-6.43	3	1.0E-07	1.0E-07	0.229628376	18.85	0.1156
DEC complex 1	-14.26	2	1.7E-02	2.1E-02	0	32.53	1.24E-04
DEC*j* complex 1	-14.06	3	1.4E-02	1.7E-02	0.218592574	34.12	5.60424E-05
DIVALIKE complex 1	-14.05	2	2.1E-02	2.0E-02	0	32.11	1.53E-04
DIVALIKE*j* complex 1	-13.43	3	8.5E-03	9.6E-03	0.274553841	32.86	1.05E-04
BAYAREALIKE complex 1	-14.78	2	1.9E-02	4.7E-02	0	33.57	7.36482E-05
BAYAREALIKE*j* complex 1	-11.52	3	1.0E-07	2.7E-02	0.28487924	29.04	7.09E-04
DEC complex 2	-12.18	2	1.1E-02	1.0E-12	0	28.36	9.99E-04
DEC*j* complex 2	-6.52	3	1.0E-12	1.0E-12	0.579073898	19.04	1.05E-01
DIVALIKE complex 2	-11.66	2	2.3E-02	9.6E-03	0	27.32	1.68E-03
DIVALIKE*j* complex 2	-6.04	3	1.0E-12	1.0E-12	0.47933328	18.08	0.1702
BAYAREALIKE complex 2	-14.39	2	4.2E-02	4.5E-02	0	32.78	1.09E-04
BAYAREALIKE*j* complex 2	-7.14	3	1.0E-07	1.0E-07	0.462016307	20.27	5.69E-02

Table 3.6. Model comparison for the BiSSE analyses of geography-correlated diversification (0=continental, 1=insular), with parameter estimates for each model with and without taxonomic sampling correction. The best-fitting model as determined by the lowest Akaike information criterion (AIC) and highest AIC weight is highlighted in bold. λ=speciation rate; μ=extinction rate; q=transition rate; Df=degrees of freedom; lnLik=log likelihood.

Without taxonomic correction	**Speciation**		**Extinction**		**Transition**						
Model	**λ0**	**λ1**	**μ0**	**μ1**	**q01**	**q10**	**Df**	**lnLik**	**AIC**	**Delta**	**AICw**
full (λ0 ≠ λ 1; μ0 ≠ μ1; q01 ≠ q10)	1.192	0.915	0	0	0.259	0	6	-28.137	68.274	3.712	0.042
equal l (λ0 = λ 1; μ0 ≠ μ1; q01 ≠ q10)	1.094	1.094	0	0	0.257	0	5	-28.281	66.563	2.001	0.099
equal m (λ0 ≠ λ 1; μ0 = μ1; q01 ≠ q10)	1.192	0.915	0	0	0.259	0	5	-28.137	66.273	1.711	0.115
equal q (λ0 ≠ λ 1; μ0 ≠ μ1; q01 = q10)	1.188	0.922	0	0	0.191	0.191	5	-29.26	68.52	3.958	0.037
equal lm (λ0 = λ 1; μ0 = μ1; q01 ≠ q10)	**1.094**	**1.094**	**0**	**0**	**0.257**	**0**	**4**	**-28.281**	**64.562**	**0**	**0.271**
equal lq (λ0 = λ 1; μ0 ≠ μ1; q01 = q10)	1.094	1.094	0	0	0.19	0.19	4	-29.393	66.785	2.223	0.089
equal mq (λ0 ≠ λ 1; μ0 = μ1; q01 = q10)	1.189	0.922	0	0	0.191	0.191	4	-29.26	66.52	1.958	0.101
equal lmq (λ0 = λ 1; μ0 = μ1; q01 = q10)	1.094	1.094	0	0	0.19	0.19	3	-29.393	64.785	0.223	0.242

With taxonomic correction	**Speciation**		**Extinction**		**Transition**						
Model	**λ0**	**λ1**	**μ0**	**μ1**	**q01**	**q10**	**Df**	**lnLik**	**AIC**	**Delta**	**AICw**
full (λ0 ≠ λ 1; μ0 ≠ μ1; q01 ≠ q10)	1.533	1.008	0	0	0.22	0	6	-26.548	65.096	3.528	0.044
equal l (λ0 = λ 1; μ0 ≠ μ1; q01 ≠ q10)	1.349	1.349	0	0	0.216	0	5	-26.974	63.949	2.381	0.078
equal m (λ0 ≠ λ 1; μ0 = μ1; q01 ≠ q10)	1.533	1.008	0	0	0.22	0	5	-26.548	63.096	1.528	0.120
equal q (λ0 ≠ λ 1; μ0 ≠ μ1; q01 = q10)	1.527	1.033	0	0	0.174	0.174	5	-27.412	64.824	3.256	0.050
equal lm (λ0 = λ 1; μ0 = μ1; q01 ≠ q10)	1.349	1.349	0	0	0.216	0	4	-26.974	61.948	0.38	0.213
equal lq (λ0 = λ 1; μ0 ≠ μ1; q01 = q10)	1.355	1.355	0	0	0.174	0.174	4	-27.784	63.568	2	0.095
equal mq (λ0 ≠ λ 1; μ0 = μ1; q01 = q10)	1.527	1.033	0	0	0.174	0.174	4	-27.412	62.824	1.256	0.138
equal lmq (λ0 = λ 1; μ0 = μ1; q01 = q10)	**1.355**	**1.355**	**0**	**0**	**0.174**	**0.174**	**3**	**-27.784**	**61.568**	**0**	**0.258**

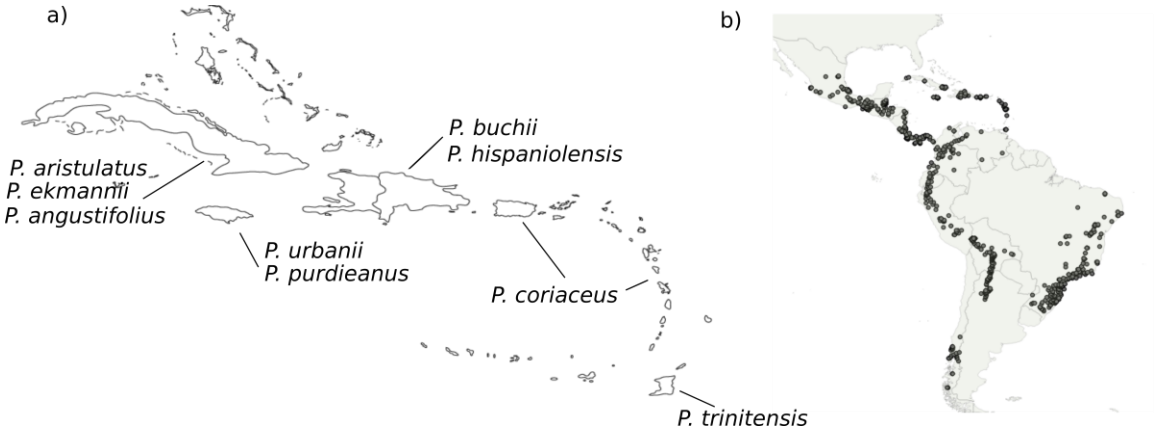


Figure 3.1. a) Map showing distribution of Caribbean *Podocarpus* species per island and, b) Neotropical *Podocarpus* species distribution used for biogeographical analysis (occurrence data from GBIF).

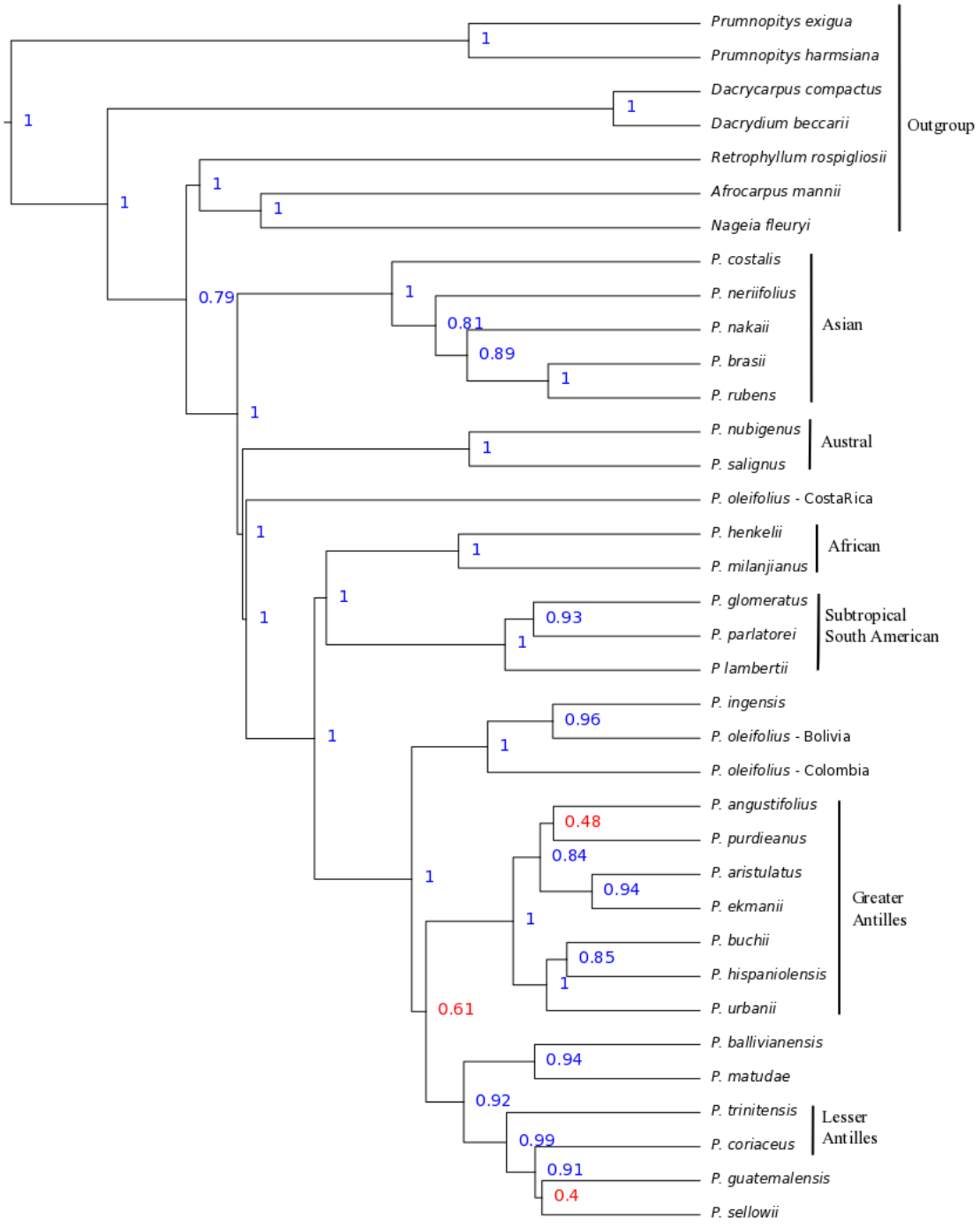


Figure 3.2. Maximum clade credibility tree obtained from a Bayesian analysis in BEAST v2.4.7. showing posterior probabilities (PP) as branch support values. PPs < 0.75 shown in red.

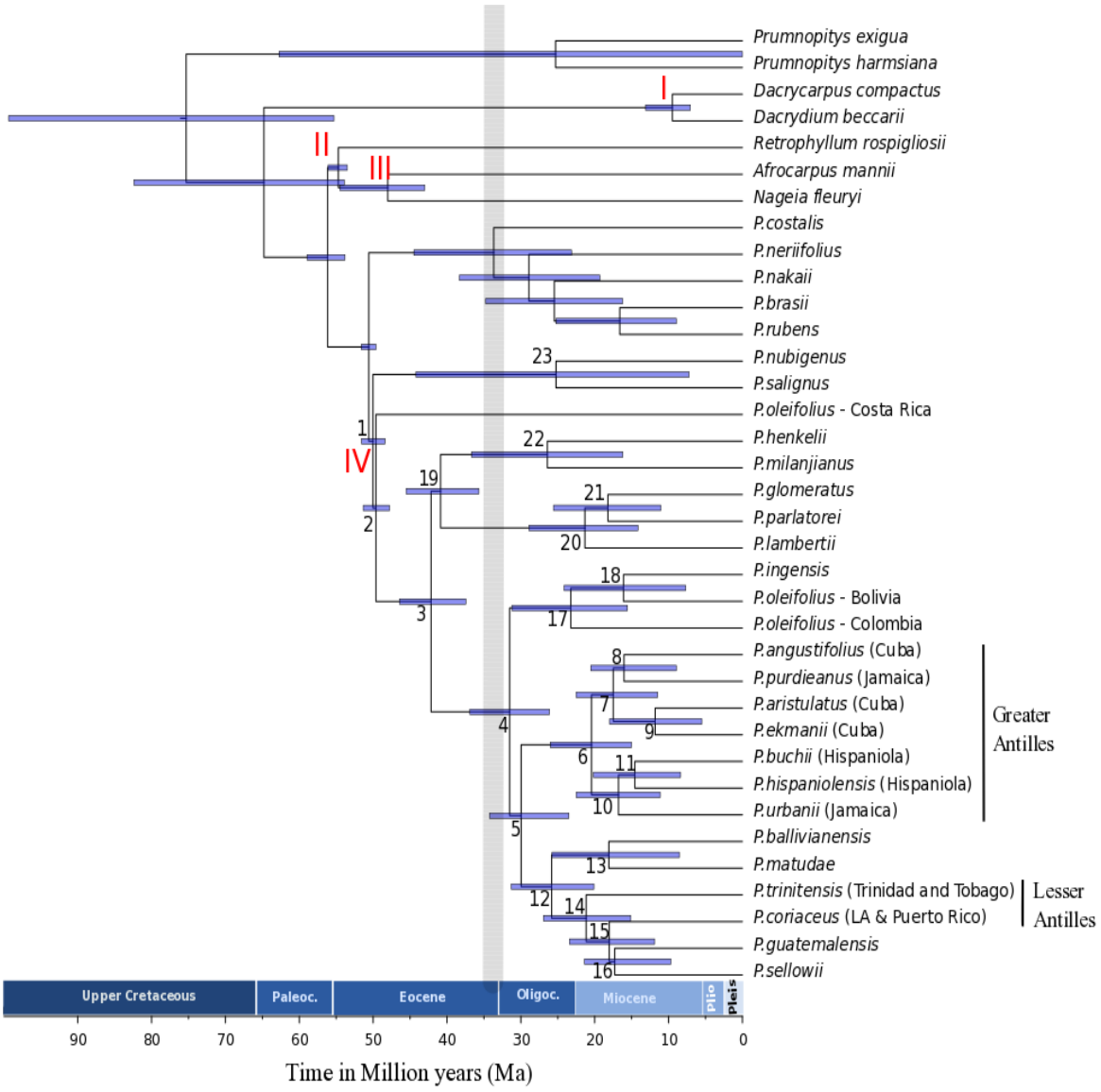


Figure 3.3. Chronogram based on the maximum clade credibility tree from BEAST v2.4.7. Blue bars represent the 95% highest posterior densities. Red roman numbers indicate fossil calibration points (Table 3.2). Divergence times for nodes numbered in black appear in Table 3.4. Grey vertical bar represents the hypothesized presence of the GAARlandia land bridge. Geological time scale at the bottom from the International Commission on Stratigraphy (v2018/04; Cohen et al., 2013). Plio = Pliocene, and Pleis = Pleistocene.

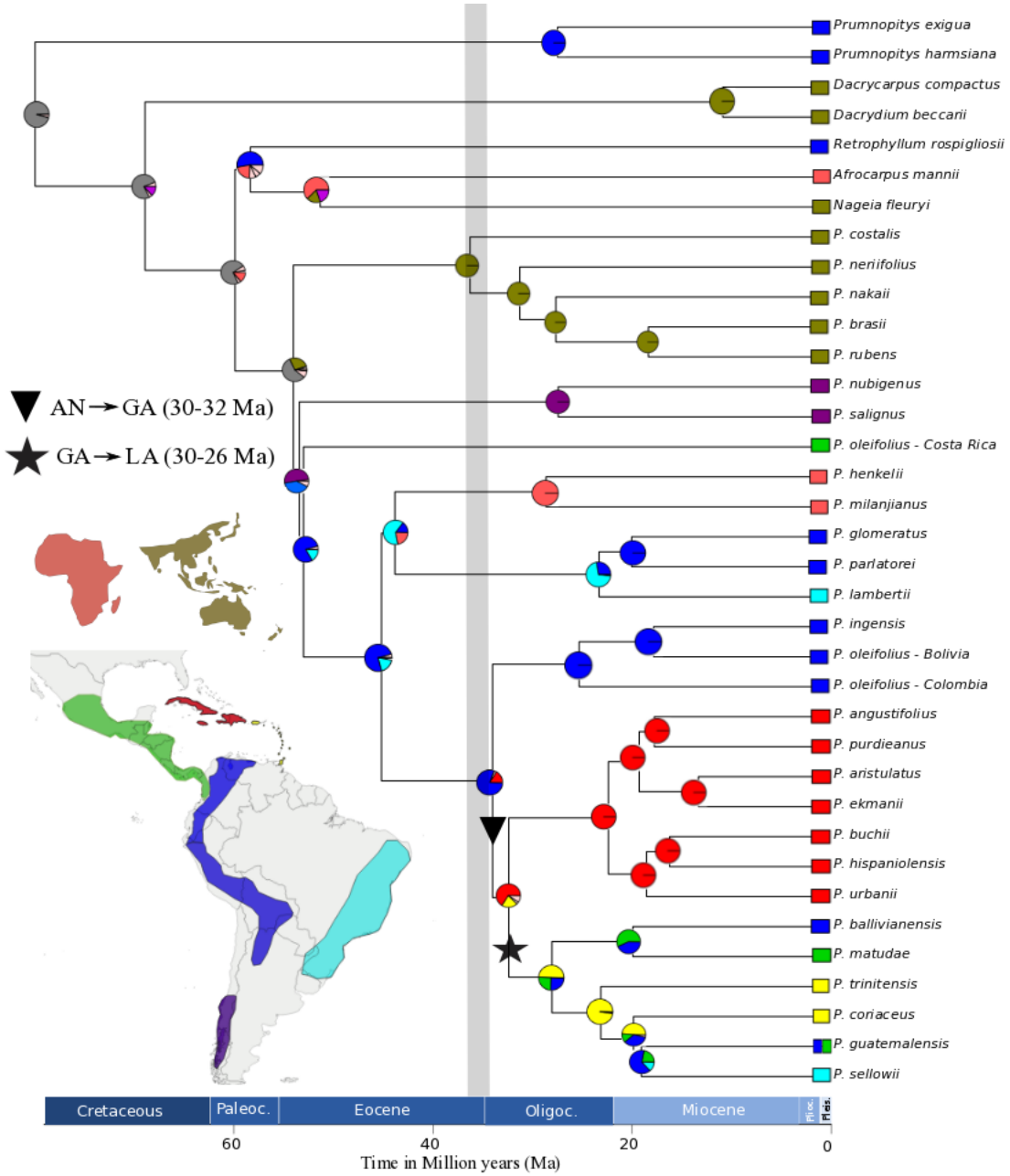


Figure 3.4. Ancestral area estimation using BioGeoBEARS based on the DEC*j* complex model. Color boxes at the tips of the tree indicate extant species geographical distribution. Grey vertical bar represents the hypothesized presence of the GAARlandia land bridge. Color code for each geographical area is indicated in inset maps. Gray color in pie charts indicate the area combination of North-Central Andes and Southern South America. Light pink in pie charts indicates combination of areas with low probabilities (<10%). Black triangle and star represent dispersal events. AN = North and Central Andes; GA = Greater Antilles; LA = Lesser Antilles; Plio = Pliocene; and Pleis = Pleistocene.

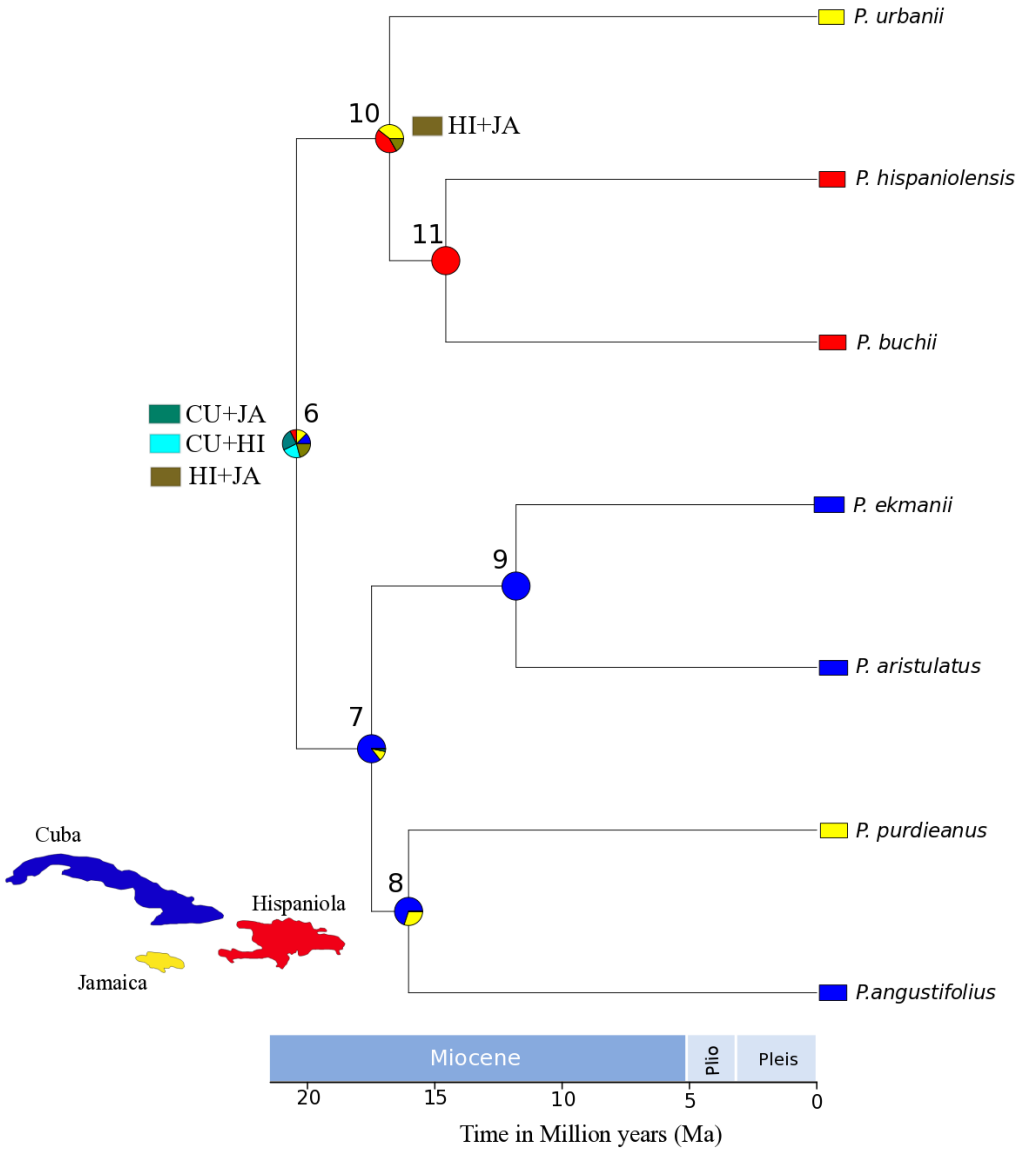


Figure 3.5. Ancestral area estimation using BioGeoBEARS for the Greater Antillean clade based on a DIVALIKE*j* model. Color boxes at the tips of the tree indicate extant species geographical distribution. Color code for each geographical area is indicated in the legend and inset map. Node numbers as in Table 3.4 and Figure 3.3. CU = Cuba; HI = Hispaniola; JA = Jamaica; Plio = Pliocene; and Pleis = Pleistocene.

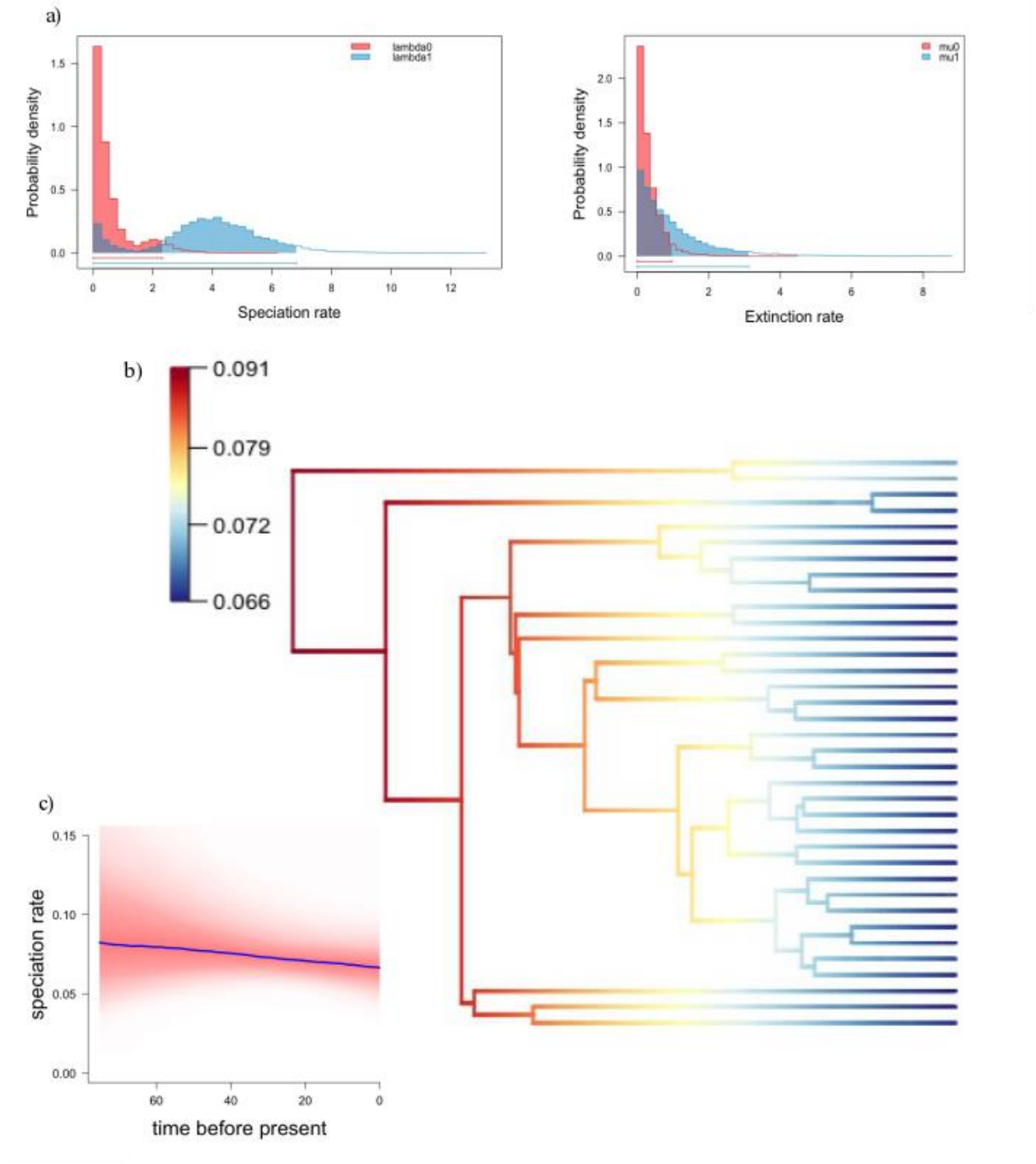


Figure 3.6. a) Bayesian posterior distribution of speciation and extinction parameters from BiSSE (0=continental, 1=insular); b) phylorate plot from BAMM analysis of posterior mean diversification rate showing the lack of significant diversification rate shifts along the phylogenetic tree; and c) net speciation rate through time plot where the blue curve is the mean diversification rate, and the red shade indicates 95% of the rates.

Chapter 4

Evolutionary history of *Podocarpus* in Hispaniola

4.1. Abstract

Hispaniola is the second largest island in the Caribbean and a hotspot of biodiversity. The island was formed by the fusion of two northern and southern palaeo-islands during the mid-Miocene (15 Ma), which are now separated by the Neiba Valley-Cul de Sac Plain. Repeated marine incursions during the Pleistocene are known to have influenced lineage divergence and genetic structure in a variety of animal taxa, but the effect on vascular plants is less understood. The tropical-subtropical conifer genus *Podocarpus* comprises two species, *P. hispaniolensis* and *P. buchii*, that are endemic to the mountainous regions of Hispaniola. The former occurs in the Cordillera Septentrional in the north, and the latter in the Sierra Bahoruco and the Sierra de Neiba in the south. They occur in sympatry in the Central Cordillera, the oldest mountain range in Hispaniola. Here I evaluate the fusion of the two palaeo-islands, and repeated marine incursions as dispersal barriers to the geographical distribution of genetic diversity, genetic structure, divergence patterns, and the historical demography of the two species. I used Genotyping by Sequencing (GBS) to identify single nucleotide polymorphisms (SNPs). The results show a population genetic structure that corresponds to the geographic distribution of the species in mountainous areas. *Podocarpus* in Hispaniola followed a stepping-stone colonization pattern from the south towards north of the island, with bottlenecks at each mountain colonization event and a progenitor-derivative speciation event in the Cordillera Central. The historical events tested do not seem to have influenced the genetic structure, diversity, or demography of *Podocarpus*, instead the current geographic barriers imposed by lowland xeric valleys did.

4.2. Introduction

The study of the geographic distribution of lineages within species and/or amongst conspecific populations and related species is known as phylogeography (Avise, 2000). Phylogeography includes the study of speciation, and the biogeographic and demographic histories of populations. It is an important tool to identify barriers or thresholds for gene flow in the biogeographic history of species and populations (Gifford et al., 2004).

The island of Hispaniola originated during the Cretaceous and the Eocene epoch of the Paleogene period as part of a volcanic arc chain that extended from Cuba to the north coast of South America. Part of this volcanic arc once formed a land unit constituted by Cuba, Hispaniola and Puerto Rico (Mann et al., 1991). As the Caribbean Plate continued to move eastward, Hispaniola and Puerto Rico became separated in the Oligocene-early Miocene, and western and northern Hispaniola separated from Cuba in the middle Miocene (ca. 20-25 Ma) (Graham, 2010).

Hispaniola, as it exists today, is the result of the juncture of two palaeo-islands that collided tectonically during the middle Miocene, ca. 15 Ma (Mann et al., 1991; Graham, 2003). The southern palaeo-island was fully emergent by the Plio-Pleistocene. These northern and southern blocks remained separated by the Neiba Valley-Cul de Sac Plain (Figure 4.1), an arid and deep rift that continued to be inundated by marine incursions repeatedly over the late Pleistocene, a period when sea levels were high (Maurrasse et al., 1980; McLaughlin et al., 1991). These events would have left the north and south palaeo-islands disconnected during the incursion periods. The southern peninsula is further divided into a west and east portion separated by the Jacmel-Fauche´

depression (known as Bond's line), which was also inundated during the Plio-Pleistocene through a sea channel (Maurrasse et al., 1980).

Hispaniola has a tropical climate that is influenced by the Atlantic wind currents and the topography of the island (Cano-Carmona et al., 2010). Its current topography is very complex and consists of mountain ranges (cordilleras) that run in parallel from NW-to-SE (Latta et al., 2006) separated by lowland xeric valleys (Heubeck & Mann, 1991; Townsend et al., 2007). The largest mountain system is the Cordillera Central, which was uplifted during the middle-to-late Eocene and is dominated by igneous and volcanic materials. The Sierra Neiba, a karstic range in the south, and the Sierra Bahoruco, a limestone mountain in the south-central part of the island, are of similar age to the Cordillera Central. The Cordillera Septentrional, located in the north of the island, is of later, Oligocene-Miocene origin and presents mostly sedimentary rocks. (Figure 4.1) (Cano-Carmona et al., 2010; Cano-Ortiz et al., 2016).

Hispaniola's biodiversity is the result of a combination of factors such as elevation, pluviometric (rainfall) gradients, and the diversity of substrates. Hispaniola has the greatest altitudinal gradient in the whole Caribbean archipelago, ranging from Lago Enriquillo (46 m) to Pico Duarte (3087 m), the highest peak in the Caribbean (Atlas de Biodiversidad y Recursos Naturales de la República Dominicana, 2012). Annual rainfall ranges from 1000-2000 mm (Cano et al., 2012): northern parts of the island receive higher rainfall due to the influence of the Atlantic Ocean, while southern parts remain drier (Cano-Carmona et al., 2010). This combination of factors generates a wide variety of habitats, which include mountain valleys that support broadleaf and pine forests, lowland grasslands, dry forest, thorny scrub habitats and agricultural land (Fahey et al.,

2012). There are 1,284 named genera of vascular plants that comprise collectively approximately 6,000 species, about a third of which (2,050 species) are endemic (Mejía, 2006). There is high representation of the families Gramineae, Orchidaceae, Bromeliaceae, Arecaceae, and Araceae (Cano-Ortiz et al., 2016). Even though Cuba is the largest island of the Caribbean and has a greater number of plant species (Borhidi, 1991), species density is slightly higher in Hispaniola with 0.064 species per km^2 in contrast to 0.050 species per km^2 in Cuba (Liogier, 2000). Floristic analyses show significant similarities between the Hispaniola flora and those from tropical Central and South America (Acevedo-Rodríguez & Strong, 2008; Cano-Ortiz et al., 2016), most likely influenced by the migratory routes into the Caribbean islands (Cano et al., 2009).

Studies on birds, rodents, reptiles, and insects have shown the role that current geography and historical marine incursions have played on the distribution of genetic structure and differentiation, as a result of isolation between populations (Glor et al., 2003; Gifford et al., 2004; Townsend et al., 2007; Sly et al., 2010, 2011; Brace et al., 2012; Matos-Maraví et al., 2014; Turvey et al., 2016). For example, mountain ranges in Hispaniola have been invoked as the explanation for the high levels of genetic differentiation amongst populations of the dry lowland lizard species *Ameiva chrysolaema*, which has led to different evolutionary lineages isolated by the mountain barriers (Gifford et al., 2004). The endemic mammal species *Solenodon paradoxus* (Turvey et al., 2016) exhibits three distinct allopatric evolutionary lineages that correspond to populations found on different island blocks, which were separated during periods of marine incursions over the late Pleistocene and the Plio-Pleistocene.

Although birds typically have greater dispersal capabilities that can overcome the barriers described above, several studies have shown patterns of geographic differentiation among bird lineages due to the configuration of north and south palaeo-islands and marine incursions (Townsend et al., 2007; Sly et al., 2010, 2011). These studies have also shown in some cases evidence of gene flow among allopatric populations. For example, studies of at least four genera of birds show that historical gene flow was not restricted, and suggest that even low levels of gene flow can homogenize populations (Sly et al., 2011). Although Townsend et al. (2007) emphasized the importance of historical geology on the differentiation of two lineages of the bird genus *Calyptophilus*, they also showed recent colonization from montane regions of the north palaeo-island into the south palaeo-island overcoming the habitat barriers between mountain ranges.

The conifer genus *Podocarpus* is represented by two species in Hispaniola, *P. buchii* Urb. and *P. hispaniolensis* de Laub. that differ in morphology, ecological attributes and elevational range. *Podocarpus hispaniolensis* is found alone in northern mountain ranges of the island, particularly in the Cordillera Septentrional of the Dominican Republic, and in the Massif du Nord in Haiti. *Podocarpus buchii* is found in the southern regions of the island, in the Sierra de Bahoruco and Neiba of the Dominican Republic, and also in the Massif de la Hotte and Chaine de la Selle in Haiti. Both species occur in the Cordillera Central, and although the two species do not typically occur in sympatry, collections that include both species have been made in relatively small areas such as Rancho Arriba (San Jose de Ocoa province) and Jarabacoa (La Vega province).

Podocarpus buchii tends to be found at higher altitude (1100-2500 m) than *P. hispaniolensis* (800-1200 m), although their altitudinal ranges slightly overlap.

The complex geological history of Hispaniola, and the peculiar distribution of the two *Podocarpus* species, offers an ideal setting to study within-island diversification and phylogeographic patterns. Here, I evaluate the effects of dispersal barriers (e.g. lowland dry valleys, marine incursions, and fusion of the two palaeo-islands) on the geographical distribution of genetic diversity, genetic structure, divergence patterns, and the historical demography of Hispaniola's *Podocarpus*. My research (1) estimates genetic diversity of populations, (2) identifies phylogeographic structure within and between species, (3) infers the phylogenetic relationship among populations, (4) elucidates the demographic history of the genus in the island, and (5) provides insight on the speciation of *Podocarpus* in Hispaniola. To address these objectives, I used the Next Generation Sequencing (NGS) technique of Genotyping by Sequencing (GBS) for SNP discovery. Caribbean (pers. observ.) and South American *Podocarpus* species do not reveal much interspecific variation when examined by traditional Sanger sequencing (Quiroga & Premoli, 2010). Thus, for the population level study of this thesis it was necessary to use many more genetic markers in order to obtain the genetic variation needed.

4.3. Material and Methods

4.3.1. Sample collection and DNA extraction

I collected silica-dried leaf samples from 20 *P. buchii* and 16 *P. hispaniolensis*, across 11 collection sites in the Dominican Republic (Table 4.1 and Figure 4.2). The low sample

size reflects their rarity in the field, as recognized by their status as IUCN Endangered species. Duplicate herbarium vouchers from each individual were deposited at the National Herbarium from Jardín Botánico Nacional Dr. Rafael Ma. Moscoso (JBSD) and the Royal Botanic Garden of Edinburgh (E). I selected adult individuals with healthy-looking leaves for both silica dried leaf samples and herbarium vouchers. I selected as outgroups single specimens from 11 *Podocarpus* species, including all species from the tropical American clade to which Hispaniola's *Podocarpus* belong as shown in Quiroga et al. (2016) (*Podocarpus urbanii, P. purdieanus, P. aristulatus, P. ekmanii, P. angustifolius, P. coriaceus, P. trinitensis, P. guatemalensis, P. matudae, P. oleifolius* and *P. sellowii*). A twelfth species, *P. rusbyi*, was unavailable. I obtained the outgroup leaf material from the Royal Botanic Garden Edinburgh and Montgomery Botanical Garden (see Table B1 in Appendix B).

I isolated DNA from 35-40 mg of dry leaf tissue with the DNeasy Plant MiniKit (Qiagen, Valencia, California, USA). I modified the manufacturer's standard protocol to improve DNA recovery: I increased AP1 buffer (for cell lysis) from 600 to 750µl, the time of incubation for cell lysis up to 60 minutes, and P3 buffer (for precipitation of polysaccharides, detergent, and proteins) from 195 to 225µl. DNA extractions were diluted in EB buffer to 20 ng/µl of DNA per sample, as required for the Genotyping by Sequencing protocol.

4.3.2. Genotyping by Sequencing and SNP discovery

The Institut de Biologie Intégrative et de Systèmes (IBIS) of the University of Laval in Canada conducted the GBS. GBS is an inexpensive NGS technique which reduces the

complexity of large genomes by use of restriction site enzymes, and identifies a large number of genetic markers. Genomic libraries were prepared for the 47 DNA samples with two restriction enzymes, *SbfI* (High Fidelity) and *MspI* (New England BioLabs Inc., Ipswich, MA). Unique barcodes of a length between 10-12 bp were added to each sample to facilitate the posterior demultiplex process. Library preparation and sequencing followed the protocol of Abed et al. (2019), except that the enzyme *SbfI* was substituted for *PstI* enzyme (High Fidelity). I obtained single-end sequencing reads of variable length (up to 200 bp) from an IonProton system, which produced a raw data FASTQ file of 32.36 GB. I inspected data quality with FastQC (Banraham Bioinformatics, Cambridge, England) for high-throughput sequence data, which reports Phred quality scores and GC content.

I carried out the SNPs discovery with *Stacks* 1.47 (Catchen et al., 2013), a pipeline that assembles large number of reads from multiple taxa by means of a statistical Maximum Likelihood approach to detect SNPs that is designed to work with restriction enzyme-based data. The workflow diagram of the *Stacks* pipeline is shown in Figure B1 in Appendix B. *Stacks* works through a series of modules. I first demultiplexed and filtered the raw GBS reads using the '*process_radtags'* module. Since IonProton produces reads of different lengths, I trimmed reads to 64 bp length. I used the option –c, which removes any read with an uncalled base and –q, which discards reads with low quality scores (below 90% probability of being correct, phred score of 10). I followed the recommendations from Paris et al. (2017) on the selection of several parameters through the pipeline. Since I did not have a reference genome, the loci were built '*de novo*' with the module '*ustacks*'. In this step, reads were aligned into matching blocks, or stacks per

sample. Then the stacks were compared, a set of loci was produced and SNPs were detected at each locus. I used: -m (minimum depth coverage to create a stack) = 5; -M (maximum distance allowed to create a stack) = 3; -p (parallel execution of several threads) = 15; and the –gapped option which allowed gapped alignments between stacks. Subsequently the module '*cstacks*' built a catalog from the loci produced in '*ustacks*' by merging stacks with at most 3 (-n = 3) mismatches between loci, and allowed parallel execution of several threads (-p = 15). The module '*sstacks*' matches the loci produced by '*ustacks*' with the catalog produced by '*cstacks*', and again parallel execution of several threads (-p = 15) was allowed.

Finally, I used the '*populations'* module of *Stacks* to obtain population level information including summary statistics, different genetic diversity indexes, F_{ST} values, and a series of SNP outputs in different formats (e.g. .vcf, .str for STRUCTURE) that were used for subsequent analyses. I conducted a first '*populations*' module where I grouped individuals in five populations as follows: SB) *P. buchii* from the Sierra Bahoruco; SN) *P. buchii* from the Sierra Neiba; CC-b) *P. buchii* from the Cordillera Central; CC-h) *P. hispaniolensis* from the Cordillera Central; and CS) *P. hispaniolensis* from the Cordillera Septentrional (Figure 4.2). Each of these 5 populations include individuals from 2 to 3 collection sites, as shown in Figure 4.2. I used: -p (minimum number of populations containing each locus) = 1; -r (minimum percentage of individuals present in a population to process a locus for that population) = 0.4; and –min_maf (minimum allele frequency) = 0.1. I used the output from this first *Stacks* run on the genetic structure (.str file), phylogenetic reconstruction, and demographic history (.vcf file) analyses described below. I run a second '*populations*' *Stacks* module, where I

grouped individuals by collection sites (11, Figure 4.2) in order to perform a phylogenetic analysis to show how collection sites are related to each other, with the same parameters as in the first '*populations*' module run. I also performed a Mantel test, a non-parametric statistical method to measure the correlation between two distance matrices to test for isolation by distance (IBD). I used the R package 'ade4' (Chessel et al., 2004) in order to measure the correlation between the F_{ST} values between pairs of the 11 collection sites and their geographic distances. I used the 11 collection sites and not the 5 populations to test for IBD in order to be more accurate with the geographic distances, since the estimation of the geographic distance among populations would not be clean and would not reflect the exact geographic location of individuals. I obtained the pairwise distance between collection sites using the Geographic Distance Matrix Generator (Ersts, 2011). Mantel test uses the Pearson coefficient and a default alpha value of 0.05.

4.3.3. Population structure analyses

I investigated the genetic structure amongst individuals with a model-based Bayesian clustering method as implemented in STRUCTURE v2.3.4 (Pritchard et al., 2000). I ran STRUCTURE independently for the two species separately and combined. I removed loci for which there was missing data for all individuals using a custom R script, keeping 15,524 loci for the 20 individuals of *P. buchii* and 5,687 loci for the 16 individuals of *P. hispaniolensis*. I determined the number of genetic clusters (*K*) with default parameter settings, the admixture model as recommended by Pritchard et al. (2000), and correlated allele frequencies. I ran analyses with and without *a priori* population information (PopData) for comparison. A total of five iterations of K = 1–7 were run for 500,000

Markov chain Monte Carlo (MCMC) generations, following 100,000 burn-in generations. For the STRUCTURE analysis using both species combined, I used the same parameters described above but tested for K = 1 to 10 for 16,381 loci. I used Structure Harvester v0.6.94, a web-based program (Earl & vonHoldt, 2012; available at http://taylor0.biology.ucla.edu/structureHarvester/) to visualize STRUCTURE results, and to explore different likelihood values across different K of all iterations in the analysis. Structure Harvester provides the lnP(D|K) and uses the Evanno method for the estimation of ΔK (Evanno et al., 2005). Finally, I used *distruct* OSX10 (Rosenberg, 2004) to graphically display the STRUCTURE results.

4.3.4. Phylogenetic relationships

To determine the phylogenetic relationship amongst collection sites, I ran Maximum Likelihood (ML) and Bayesian Inference (BI) analyses using the CIPRES Science Getaway (Miller et al., 2010). I ran the ML analysis on RAxML-HPC2 v8.2.10 (Stamatakis, 2006) where I used the nucleotide substitution model GTR+ Γ and a rapid bootstrap algorithm with 500 replicates. I selected 11 tropical American *Podocarpus* species as the outgroup. I ran the BI analysis on MrBayes v3.2.6 (Ronquist et al., 2012) for 50 million MCMC generations sampling every 10,000 generations with a burn-in of 25% of generations to be discarded. For the ML and BI analyses the concatenated SNPs matrix contained 74,260 SNPs. I also ran ML and BI phylogenetic analyses for the five populations used in the demographic history analysis using the same parameters as described above.

4.3.5. Demographic history analysis

I conducted a model selection analysis to infer the demographic history of *Podocarpus* species in Hispaniola using DIYABC v 2.1.0 (Cornuet et al., 2014). I defined five populations based on the STRUCTURE analysis (see results below) and the geographic distribution of both species which included: SB- *P. buchii* from the Sierra Bahoruco, SN- *P. buchii* from the Sierra Neiba, CC-b- *P. buchii* from the Cordillera Central, CC-h- *P. hispaniolensis* from the Cordillera Central, and CS- *P. hispaniolensis* from the Cordillera Septentrional (Figure 4.2). I discarded loci obtained from the *Stacks* '*populations*' module when all individuals in a population had missing data for any particular locus, or when loci were monomorphic, as required by DIYABC, leaving 244 polymorphic loci for the analysis. The approximate Bayesian computation (ABC) method uses a linear discriminant on summary statistics to estimate sample sizes and coalescent times. Summary statistics used in this analysis are genetic diversity, F_{ST} distances and Nei's distances. Each of the populations defined above have a population size designated by the parameter "N" (e.g. NSB corresponds to the Sierra Bahoruco, NSN corresponds to the Sierra Neiba, etc). Population sizes during bottleneck events for founder events are designated as Nf, and population size of an ancestor is designated by NA. ABC provides coalescent times (t) in number of generations, which were then multiplied by the generation time of *Podocarpus* (25 years; Blendinger, 2017).

I designed eight competing demographic scenarios described as follows.

Scenario 1) an ancestral population from the Cordillera Central gave rise to both *Podocarpus* species (CC-b and CC-h) from where subsequently one species migrated south (SN and SB) and the other north (CS).

Scenario 2) has identical topology as Scenario 1, but included bottleneck events in each migration step.

Scenario 3) proposes the Cordillera Central as a region of secondary contact of both *Podocarpus* species in the island. Migrations from the south to the Cordillera Central for *P. buchii* as indicated by arrows (SB → SN → CC – b) and from the north to the Cordillera Central for *P. hispaniolensis* (CS → CC-h).

Scenario 4) has identical topology to Scenario 3, but includes bottleneck events in each migration step.

Scenario 5) proposes a stepping-stone migration pattern from the southernmost population of *P. buchii* to the north (SB → SN → CC-b → CC-h → CS).

Scenario 6) has identical topology to Scenario 5, but includes bottleneck events in each migration step.

Scenario 7) proposes a stepping-stone migration pattern from the northernmost population of *P. hispaniolensis* to the south (CS → CC-h → CC-b → SN → SB).

Scenario 8) has identical topology to Scenario 7, but includes bottleneck events in each migration step (Figures 4.3a and, B2 in Appendix B).

I used the Log-uniform prior distributions for all parameters and set the minimum and maximum values for all population sample sizes (N, Nf and NA) to 10 and 100,000 respectively. I set temporal constrains as follow: t4>=t3>=t2>=t1. For the summary statistics of each population, I selected the mean of complete distributions for the genetic diversities, and for the two samples summary the Nei's and F_{ST} mean of complete distribution. I ran eight million simulated datasets, one million per scenario. I compared

scenarios by calculating their relative posterior probability (PP) by logistic regression on 0.1% of the simulated data sets closest to the observed datasets. For the best fit scenario, I calculated the posterior and prior predictive global errors using 250 pseudo-observed test datasets (PODs) for the logistic regression approach and assessed the confidence in the scenario calculating type I and type II errors. Posterior distributions of parameters of the best model were determined applying the logit transformation of parameters to 1% of the closest simulated datasets. In order to assess the performance of the preferred scenario parameters, I computed the mean relative bias and the square root of the relative mean integrated square error (RRMISE) across the 500 PODs using the prior parameters from 1% of the closest simulated datasets. I performed model checking on the preferred scenario to verify the goodness of the model by comparing the summary statistics between the observed and simulated datasets.

I ran a second analysis based on the results of the first DIYABC and the phylogenetic analyses to investigate the colonization patterns of *P. buchii* populations in southern Hispaniola prior to the arrival to the Cordillera Central (Figures 4.3b and B3 in Appendix B). The purpose of this second analysis was to complement the low support that deeper nodes received in the phylogenetic analysis. The competing demographic scenarios were:

1) SB $\rightarrow$ SN $\rightarrow$ CC-b $\rightarrow$ CC-h $\rightarrow$ CS;

2) SN $\rightarrow$ SB $\rightarrow$ CC-b $\rightarrow$ CC-h $\rightarrow$ CS;

3) CC -b $\rightarrow$ SN $\rightarrow$ SB $\rightarrow$ CC-h $\rightarrow$ CS;

4) SN $\rightarrow$ CC -b $\rightarrow$ SB $\rightarrow$ CC-h $\rightarrow$ CS;

5) CC -b → SB → SN → CC-h → CS;

6) SB → CC- b → SN → CC-h → CS.

All scenarios included bottleneck events in each migration step, and all parameters were kept the same as the first DIYABC analysis.

4.4. Results

4.4.1. GBS and SNP discovery

The IonProton platform generated a total of 102.6M reads from which the filtering of '*process_radtags*' eliminated 3.7M reads with low quality and 5.2M with ambiguous barcodes. A total of 93.7 million reads (approximately 91.3% of initial number of reads) of 64 bp length were kept for further analyses. The number of reads per sample ranged from ca. 258K to ca. 10.3M, with an average number of ca. 2.7M reads per sample. The inspection of the quality analysis in FastQC reported a Phred score average of 24, and GC content of 57%.

The catalog built from *Stacks* produced 788,666 loci and 276,718 SNPs. After filtering in the *populations* module, 11,615 loci (of which 8,335 were polymorphic) and 22,657 SNPs remained for further analyses. The second *population*s run recovered 116,547 loci (of which 32,456 were polymorphic) and 74,260 SNPs.

4.4.2. Patterns of population structure and genetic diversity

STRUCTURE recovered genetic clustering and admixture for *P. buchii* and *P. hispaniolensis* independently, and both species together. The analyses with and without population information gave very similar results. The following results correspond to the analysis with population information (PopData).

For *P. buchii* the highest supported *K* was 4 (Δ*K*=799.427 and lnl=- 264,556.7) followed by *k* of 2 (Δ*K*=770.002 and lnl=- 327,770.3). Both *K*=2 and *K*=4 showed the Sierra Bahoruco as a distinct genetic cluster. Relative to *K*=2, *K*=4 separates the Cordillera Central from the Sierra de Neiba and places one individual from the Cordillera Central in the Sierra Bahoruco (bar plots for *K* =2 and *K*=4 are shown in Figure 4.4a). Two individuals from the Sierra Bahoruco and the four from the Cordillera Central showed admixture for *K*=2, while individuals from the Sierra Neiba did not show any admixture. For *K*=4 four individuals from the southern mountains of the Sierra Bahoruco and two from the Sierra Neiba showed admixture, while the Cordillera Central had one individual with admixture.

For *P. hispaniolensis*, the highest supported number of clusters was *K*=2 (Δ*K*=481.48 and lnl=-105,210.78), followed by *K*=5 (Δ*K* =425.54 and lnl=- 83,844.8). *K*=2 differentiated the Cordillera Central from the Cordillera Septentrional, and *K*=5 showed a clear separation of individuals from the Cordillera Central, and collection sites 10 and 11 within the Cordillera Septentrional (bar plots for *K* =2 and *K*=5 are shown in Figure 4.4b). Individuals from the Cordillera Central for *K*=2 did not show any admixture but individuals from the Cordillera Septentrional did, with one individual having ca. 90% associated with the Cordillera Central cluster. For *K*=5, two individuals from the Cordillera Central and five from the Cordillera Septentrional showed admixture.

When species were analyzed together, *K* =7 received the highest Δ*K* and mean lnl (Δ*K*=374.46 and lnl=-365,772.56), followed by *K* =5 (Δ*K*=81.03 and lnl=- 373,275.66). In this combined analysis for both *K* =5 and *K*=7, the Sierra Bahoruco, the Sierra Neiba

and the Cordillera Central-*P. buchii* were differentiated but the Cordillera Central-*P. hispaniolensis* and the Cordillera Septentrional were not. This latter group (Cordillera Central-*P. hispaniolensis* + Cordillera Septentrional) shares genetic similarity with the Sierra Bahoruco, the Sierra Neiba and the Cordillera Central- *P. buchii* (bar plots for K =5 and K=7 are shown in Figure 4.4c). The individuals that showed the highest number of clusters corresponded to *P. buchii* from the Sierra Bahoruco, while *P. hispaniolensis* individuals did not show any admixture. Only for K = 9 and K = 10 the Cordillera Septentrional individuals showed admixture (Figure B4 in Appendix B), but these two Ks did not receive the highest support. The genetic clusters revealed in the separate and combined STRUCTURE analyses showed a geographic genetic structure by mountain chain and species identity, which supported the assignment of five populations for the historical demographic analysis.

Population genetic measures of heterozygosity and nucleotide diversity showed that *P. buchii* has higher genetic diversity than *P. hispaniolensis* (Table 4.2). This result is consistent with the high number of genetic clusters found for *P. buchii* compared to those of *P. hispaniolensis* as recovered in the STRUCTURE analysis of both species together (Figure 4.4c).

When individuals were grouped in 5 populations, that with the highest expected heterozygosity (H_e) was the population from the Sierra Neiba with H_e=0.3434, followed by the population from the Sierra Bahoruco with H_e=0.3343. Individuals with the lowest H_e belong to the Cordillera Septentrional population. *P. buchii* individuals from the Cordillera Central showed the highest nucleotide diversity (π), followed by the Sierra

Bahoruco population, and the population with the lowest π was from the Cordillera Septentrional. When H_e and π were estimated for individuals grouped by collection sites, the two sites from the Sierra Neiba had the highest H_e and π, while the lowest values were observed on collection site 6 for *P. hispaniolensis* from the Cordillera Central (Table 4.2).

Pairwise F_{ST} values amongst the 11 collection sites ranged from 0.000 in several instances (Site 2 with sites 6, 7, 8, 9 and site 9 with 2, 6, 7, 8) to 0.160 for sites 4 and 10 (see Table B2 in Appendix B for all pairwise F_{ST} values). The Mantel test showed a significant correlation between the F_{ST} of the 11 collection sites and their geographic location (R = 0.442, p-value = 0.003), which suggests the occurrence of isolation by geographic distance.

4.4.3. Phylogenetic relationship among collecting sites

ML and BI analyses for the 11 collection sites recovered a mostly congruent topology and showed that *P. hispaniolensis* was nested within *P. buchii* rendering this latter species non-monophyletic (Figures 4.5a and 4.5b). Both analyses recovered with high support (PP=1.0 and BS (bootstrap values) =0.99) a clade that included all *P. hispaniolensis* sites. The two southernmost *P. hispaniolensis* sites were sister to a clade formed by the three northernmost sites of *P. hispaniolensis*. Both analyses also recovered *P. buchii* paraphyletic with the southernmost site of *P. buchii* from the Sierra Bahoruco as sister to the *P. hispaniolensis* clade with high support (PP=1.0 and BS=0.99). The relationships of *P. buchii* sites closer to the root showed some discrepancies between the two analyses but recovered both the Cordillera Central sites as sister to each other with high support (PP=1.0 and BS=0.95). The ML identified site 2 from the Sierra Bahoruco Occidental as

sister to the rest of the sites included in the analysis, while the BI analysis identified one of the sites from the Sierra Neiba. Support in both cases was low (Figures 4.5a and 4.5b). The ML and BI analyses for the five populations, provided only poor resolution due to extremely short branches (results not shown).

4.4.4. Demographic history analyses

DIYABC identified Scenario 6 (a stepping-stone model from South to North with bottlenecks in each mountain colonization) as the best-fit model for the data under the logistic regression approach (PP = 0.98 [95% HPD = 0.9828 - 0.9897]). Under the direct approach, Scenario 8 (a stepping-stone model from North to South) was the best-fit model for the data, but with substantially less support (PP=0.52 [95% HPD = 0.0842 - 0.9598) followed closely by Scenario 6 (PP=0.47 [95% HPD = 0.0402 - 0.9158). The PP in both approaches increased for Scenario 6 and decreased for Scenario 8 with the 8×10^6 simulations. Scenarios that hypothesized the Cordillera Central as either an area of secondary contact or origin of *Podocarpus* in the island received no support (Table B3 and Figure B5 in Appendix B). The posterior predictive global error for Scenario 6 was 0.140 under the direct approach and 0.120 under the logistic approach, while the prior predictive global error was 0.436 under the direct approach and 0.412 under the logistic approach. Type error I and II were 0.46 and 0.004, respectively. With respect to support for Scenario 6, the model checking shows a good fit between the model and the data (Figure B6 in Appendix B). However, some of the observed summary statistics matched poorly (proportion of simulated<observed datasets was > 95%) with the simulated data set (Table B4 in Appendix B). Mean population size ranged from 2,860 to 13,500 (Table

B5 Appendix B) and mean coalescent times ranged from 1,672.5 years at t4 in the first split followed by 1,497.5 years at t3, 927.5 at t2 and 587.5 at t1. These coalescent times were estimated based on a *Podocarpus* generation time of 25 years (Blendinger 2017). Priors were not biased (Table B6 in Appendix B).

For the second DIYABC analysis scenario 2 (a stepping-stone model from the Sierra Neiba into the Sierra Bahoruco and finally the Cordillera Central) was the best-fit model for the data under the logistic regression approach (PP = 0.8959 [95% HPD = 0.8890 - 0.9027]; Table B7 in Appendix B). The posterior predictive global error for scenario 2 was 0.532 under the direct approach and 0.480 under the logistic approach, whereas the prior predictive global error was 0.260 under the direct approach and 0.204 under the logistic approach. Observed summary statistics are shown in Table B8 in Appendix B. Type I and II errors were 0.734 and 0.074, respectively. Mean population sizes and coalescent times are shown in Table B9 in Appendix B and bias of prior in Table B10 in Appendix B.

4.5. Discussion

In Hispaniola, species and/or population divergence has been influenced by current dispersal barriers. Similar to previous studies and according to my expectations, the distribution of *Podocarpus* in Hispaniola was geographically structured. I found genetic structure among individuals from different mountain ranges within each *Podocarpus* species. This genetic structure, however, could not be attributed to Pleistocene marine incursions or the fusion of palaeo-islands given the very recent population divergence

times as estimated in the historical demographic analyses and the unexpected topologies for the relationships among populations. The structure instead might have been influenced by contemporary topographic barriers of the dry valleys separating the different mountain ranges where *Podocarpus* grow. The history of the genus in the island seems to be the result of a single colonization in the south with subsequent northward dispersal in a stepping stone manner, and a speciation event in Cordillera Central. The phylogenetic reconstruction supports a progenitor-descendent speciation scenario where the monophyletic *P. hispaniolensis* emerged from the older and paraphyletic *P. buchii*.

4.5.1. Geographic structure and genetic diversity pattern

The distribution of *Podocarpus* in Hispaniola is geographically structured according to the main mountainous systems. The STRUCTURE analysis separates individuals in clusters that corresponded to the different mountain systems. The temporal factors that might have led to this genetic structure in *Podocarpus* will be discussed in the next section.

Contrasting spatial patterns have been observed in Hispaniola fauna. Spatial differentiation between lineages has been found in Hispaniola mammals from the genus *Plagiodontia* (Rodentia) (Brace et al., 2012) and *Solenodon paradoxus* (Soricomorpha) (Turvey et al., 2016), where the distribution of genetically distinct allopatric lineages coincides with the south and north palaeo-islands. Matos-Maraví et al. (2014) showed in *Calisto* butterflies on Hispaniola that two vicariant events prompted divergence between species, one related to the uplift of the Cordillera Central, and a second one related to the repeated marine incursions in Neiba Valley-Cul de Sac Plain. It is likely that for

Podocarpus populations, the dry valleys that lay in between the mountain chains acted as dispersal barriers that prompted divergence. The geographical structure found in this present study is concordant with Gifford et al. (2004) study on the lizard species *Ameiva chrysolaema*, where the complex topography of Hispaniola led to three different evolutionary lineages. Results here contrast with those of Sly et al. (2011), who suggest that lineage divergence in birds is not due to current topographic or ecological factors, but instead to the geological fragmentation of the palaeo-islands.

Mountain ranges in Hispaniola act as barriers to gene flow, for example in lizards (Gifford et al., 2004). However, individuals' genetic admixture observed in the STRUCTURE analyses (Figure 4.4) suggest the presence of gene flow across the different mountain ranges for *Podocarpus*. Some studies have shown gene flow to be present amongst bird populations occupying different mountain ranges, suggesting over-land dispersal between ranges from south and north palaeo-islands (Sly et al., 2011). Despite dry valleys that might act as barriers for *Podocarpus* and thus led to genetic structure, gene flow has homogenized populations (Figure 4.4c), where a genetic cluster (blue color in K = 5 and K = 7) is present in all populations. The analysis of both *Podocarpus* species combined did not show any admixture in *P. hispaniolensis* for the two best-supported numbers of genetic clusters (K=5, 7), which could be attributed to the different sets of loci included in the combined versus individual analyses.

The higher genetic diversity in *P. buchii* is likely due to an older origin with respect to the derived species *P. hispaniolensis* as discussed in the next section. The study by Brace et al. (2012) on hutias in Hispaniola shows that southern populations had a higher nucleotide diversity (π) compared to northern populations. Rodríguez-Peña et al.

(2014) did not find significant differences in genetic diversity (H_e) for northern and southern populations of the palm species *Pseudophoenix vinifera* in Hispaniola.

Overall genetic differentiation amongst populations (pairwise F_{ST} values) was low. Collections sites from the Cordillera Septentrional in the north (sites 10 and 11) showed the highest degree of genetic differentiation with the rest of sites. The highest F_{ST} was for site 4 from the Sierra Neiba with respect to sites 10 and 11 from the Cordillera Septentrional, which is also one of the greatest geographic distance among sites. Geographic distance is a major cause of genetic isolation. This might be the case for *Podocarpus* populations in Hispaniola as indicated by the Mantel test (R = 0.442, p-value = 0.003). Even in small islands, isolation by distance may play a role in lineage divergence in plants. This is the case for several plant genera from Lord Howe Island (Australia) where isolation by distance was an important factor determining lineage divergence (Papadopulos et al., 2014). Genetic distances are not always linearly correlated with geographic distances and other factors such as habitat configuration and maximum migration distance might influence isolation amongst populations (van Strien et al., 2015). This might explain why the most geographically distant collection sites (sites 1 & 2 vs 10 & 11) did not show the greatest genetic differentiation in this study.

4.5.2. Demographic history of Podocarpus in Hispaniola

The Bayesian inference analysis for the first DIYABC strongly supports scenario 6, which depicts the hypothesis of an initial colonization of ancestors of *P. buchii* into the south of the island (Sierra Bahoruco) and further stepping stone migration with bottleneck events to the north into Sierra Neiba and Cordillera Central. This contrast with the results

of the second DIYABC analysis for which scenario 2 (initial colonization into Sierra Neiba and stepping stone migration into Sierra Bahoruco and Cordillera Central) was the best fit scenario. However, based on the posterior predictive error for the second DIYABC analysis this result should be taken with caution. A speciation event occurred within Cordillera Central giving rise to *P. hispaniolensis*, which later migrated to the Cordillera Septentrional in the north. An opposite colonization direction within Hispaniola was shown for chat-tanagers where a colonization event of *Calyptophilus frugivorus* mainly distributed in the north palaeo-island into Sierra Bahoruco in the south palaeo-island, indicating the movement of taxa between mountainous ranges (Townsend et al., 2007). This north to south colonization direction was the second best-fit demographic scenario for *Podocarpus* in Hispaniola. This stepping stone migration with bottlenecks could also explain the overall lower genetic diversity of *P. hispaniolensis*, since founder events are followed by bottlenecks with an impact in genetic diversity (Barrett, 1996).

The estimated population divergence times from DIYABC (Table B5 in Appendix B) are rather young in comparison to lineage divergence for example in the hutia species *Plagiodontia aedium*, where south and north population diverged 0.594 Ma, during the time when marine incursions in the Neiba Valley-Cul de Sac Plain were still occurring (Brace et al., 2012). Given the inferred divergence dates, *Podocarpus* most likely arrived into Hispaniola after the island achieved its current configuration, the north and south palaeo-islands were already connected, and all marine incursions no longer occurred. As shown in Figure B2 and Table B5 in Appendix B the first population split followed by a bottleneck event (t4) occurred 1,672.5 years ago, much later than the Pleistocene marine

incursions and uplift of the main mountain ranges. Therefore, these marine incursions and geological reconfigurations a priori did not leave a signature in the population divergence history of *Podocarpus*. I acknowledge that the small sample size in the present study could have rendered the rather young divergence times, thus these times should be corroborated with a larger *Podocarpus* sampling in Hispaniola including Haiti.

4.5.3. Speciation event in Cordillera Central

The phylogenetic relationships amongst collection sites of both *Podocarpus* species suggest a scenario of progenitor-descendent speciation where the progenitor species (*P. buchii*) appears to be paraphyletic and the derived (*P. hispaniolensis*) monophyletic. In this progenitor-descendent speciation pattern the descendent species originates from a genetic subset of the progenitor species (Schlüter et al., 2011). This is supported by the STRUCTURE analysis of species combined (Figure 4.4c), where the derived *P. hispaniolensis* is mainly composed by one genetic cluster (blue cluster), also present in most individuals of the progenitor *P. buchii*. One of the premises of this speciation mode is that the progenitor species is the widespread taxon while the descendent has a more restricted area. Progenitor-descendent type of speciation is commonly seen in continent-descended island sister taxa, where the continental progenitor is widespread and has been seen in island systems such as Macaronesia (Valtueña et al., 2017). This speciation pattern has been seen in other tropical plant taxa such as *Andira* (Pennington, 2003), *Protium* (Fine et al., 2013), *Inga* (Dexter et al., 2010) and *Dussia* (Winterton et al., 2014), and it is more common in rain forest taxa than dry forest ones (Pennington & Lavin,

2016). This is not the case for *P. buchii*, which has a distribution as restricted as *P. hispaniolensis*.

The demographic and phylogenetic analyses suggest that a speciation event occurred in Cordillera Central. A peripheral population of *P. buchii* might have become isolated, followed by limited gene flow, and thus prompting the origin of *P. hispaniolensis*. Some studies have shown that the result of peripatric speciation is that the progenitor becomes paraphyletic (Rieseberg & Brouillet, 1994), which is the case for *P. buchii*. For *Podocarpus* in Cordillera Central, ecological specialization through the different elevation ranges each species occupy rather than geographic isolation could have led the speciation event. Glor et al. (2003) showed that lineage divergence and genetic structure observed in a group of *Anolis* lizards in Hispaniola were more likely a result of microhabitat specialization rather than due to geological or topographical causes. Another example of ecological driven diversification in Hispaniola occurred in a clade of the butterfly genus *Calisto*, where the uplift of Cordillera Central potentially provided new ecological habitats that increased diversification rates (Matos-Maraví et al., 2014). The study of Sly et al. (2011) presents evidence for ecological driven divergence between two ecologically divergent sister bird genera, *Xenoligea* (montane taxon) and *Microligea* (generalist taxon) in Hispaniola. Authors here argued that the divergence of these two taxa does not correspond to any geological barriers. This type of population isolation by environment (IBE) has been a major force of ecological speciation in insular plant genera (Papadopulos et al., 2014), and it has been shown that the altitudinal gradient can drive speciation by ecological specialization in islands in the Mediterranean region (e.g.

Corsica and Crete, Steinbauer et al., 2013). The prevalence of the progenitor-derivative speciation within and among Caribbean islands is a subject of future research.

4.6. References

Abed A., Légaré G., Pomerleau S., St-Cyr J., Boyle B., & Belzile F. (2019) Genotyping-by-Sequencing on the Ion Torrent Platform in Barley: Methods and Protocols. *Methods in Molecular Biology* (ed. by W. Hardwood), pp. 233–252. Humana Press, New York, NY.

A.B.R.N. (2012) *Atlas de Biodiversidad y Recursos Naturales de la Republica Dominicana*. MINISTERIO DE MEDIO AMBIENTE Y RECURSOS NATURALES, Santo Domingo, Republica Dominicana.

Acevedo-Rodríguez P. & Strong M.T. (2008) Floristic richness and affinities in the West Indies. *Botanical Review*, **74**, 5–36.

Avise J.C. (2000) *Phylogeography: The History and Formation of Species*. Harvard University Press, Cambridge, MA.

Barrett S.C.H. (1996) The reproductive biology and genetics of island plants. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, **351**, 725–733.

Blendinger P.G. (2017) Functional Equivalence in Seed Dispersal Effectiveness of *Podocarpus parlatorei* in Andean Fruit-Eating Bird Assemblages. *Frontiers in Ecology and Evolution*, **5**, 57.

Borhidi A. (1991) *Phytogeography and vegetation ecology of Cuba*. Akadémiai Kiadó, Budapest.

Brace S., Barnes I., Powell A., Pearson R., Woolaver L.C., Thomas M.G., & Turvey S.T. (2012) Population history of the Hispaniolan hutia *Plagiodontia aedium* (Rodentia: Capromyidae): testing the model of ancient differentiation on a geotectonically complex Caribbean island. *Molecular Ecology*, **21**, 2239–2253.

Cano-Carmona E., Veloz-Ramírez A., & Cano-Ortiz A. (2010) Contribution to the biogeography of the Hispaniola (Dominican Republic, Haiti). *Acta Botanica Gallica*, **157**, 581–598.

Cano-Ortiz A., Musarella C.M., Fuentes J.C.P., Gomes C.J.P., & Cano E. (2016) Distribution patterns of endemic flora to define hotspots on Hispaniola. *Systematics and Biodiversity*, **14**, 261–275.

Cano E., Cano-Ortiz A., Del Rio González S., Alatorre Cobos J., & Veloz A. (2012) Bioclimatic map of the Dominican Republic. *Plant Sociology*, **49**, 81–90.

Cano E., Veloz Ramirez A., Cano-Ortiz A., & Esteban Ruiz F.J. (2009) Distribution of Central American Melastomataceae: biogeographical analysis of the Caribbean islands Dis. *Acta Botanica Gallica*, **156**, 527–557.

Catchen J., Hohenlohe P.A., Bassham S., Amores A., & Cresko W.A. (2013) Stacks: an analysis tool set for population genomics. *Molecular Ecology*, **22**, 3124–3140.

Chessel D., Dufour A.B., & Thioulousse J. (2004) The ade4 package - I: One-Table Methods. *R News*, **4**, 5–10.

Cornuet J.-M., Pudlo P., Veyssier J., Dehne-Garcia A., Leblois R., Marin J.-M., & Estoup A. (2014) DIYABC v2.0: a software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. *Bioinformatics*, **30**, 1187–1189.

Dexter K.G., Pennington T.D., & Cunningham C.W. (2010) Using DNA to assess errors in tropical tree identifications: How often are ecologists wrong and when does it matter? *Ecological Monographs*, **80**, 267–286.

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

Ersts P.J. (2011) Geographic Distance Matrix Generator v1.2.3. American Museum of Natural History, Center for Biodiversity and Conservation. Available at http://biodiversityinformatics.amnh.org/open_source/gdmg. Data accesed September 2018. .

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.

Fahey A.L., Ricklefs R.E., Latta S.C., & Dewoody J.A. (2012) Comparative Historical Demography of Migratory and Nonmigratory Birds from the Caribbean Island of Hispaniola. *Evolutionary Biology*, **39**, 400–414.

Fine P.V.A., Metz M.R., Lokvam J., Mesones I., Ayarza Zuñiga J.M., Lamarre G.P.A., Vásquez Pilco M., & Baraloto C. (2013) Insect herbivores, chemical innovation, and the evolution of habitat specialization in Amazonian trees. *Ecology*, **94**, 1764–1775.

Gifford M.E., Powell R., Larson A., & Gutberlet Jr. R.L. (2004) Population structure and history of a phenotypically variable teiid lizard (*Ameiva chrysolaema*) from Hispaniola: the influence of a geologically complex island. *Molecular Phylogenetics and Evolution*, **32**, 735–748.

Glor R.E., Kolbe J.J., Powell R., Larson A., & Losos J.B. (2003) Phylogenetic analysis of ecological and morphological diversification in Hispaniolan trunk-ground anoles (*Anolis cybotes* groups). *Evolution*, **57**, 2383–2397.

Graham A. (2003) Geohistory models and Cenozoic paleoenvironments of the Caribbean region. *Systematic Botany*, **28**, 378–386.

Graham A. (2010) *Late Cretaceous and Cenozoic History of Latin American Vegetation and Terrestrial Environments.* Missouri Botanical Garden Press, St . Louis, MO.

Heubeck C. & Mann P. (1991) Structural geology and Cenozoic tectonic history of the southeastern termination of the Cordillera Central, Dominican Republic. *Geologic and Tectonic Development of the North American–Caribbean Plate Boundary in Hispaniola* (ed. by P. Mann, G. Draper, and J. Lewis), pp. 315–336. Boulder, CO.

Latta S., Rimmer C., Keith A., Wiley J., Raffaele H., McFarland K., Fernandez E., MacKay B.K., Pedersen T., Williams K., Fisher C., & Rulon B. (2006) *Birds of the Dominican Republic and Haiti.* Princeton University Press, Princeton, NJ.

Liogier A.H. (2000) *Diccionario botánico de nombres vulgares de la Española*. Santo Domingo: Jardín Botánico Nacional Dr. Rafael Ma. Moscoso. 598p.

Mann P., Draper G., & Lewis J. (1991) An overview of the geologic and tectonic development of Hispaniola. *Geologic and Tectonic Development of the North American–Caribbean Plate Boundary in Hispaniola* (ed. by P. Mann, G. Draper, and J. Lewis), pp. 1–18. The Geological Society of America, Inc., Boulder, CO.

Matos-Maraví P., Águila R.N., Peña C., Miller J.Y., Sourakov A., & Wahlberg N. (2014) Causes of endemic radiation in the Caribbean: evidence from the historical

biogeography and diversification of the butterfly genus *Calisto* (Nymphalidae: Satyrinae: Satyrini). *BMC Evolutionary Biology*, **14**, 1–18.
Maurrasse F.J.-M.R., Pierre-Louis F., & Rigaud J.-G. (1980) Cenozoic facies distribution in the southern peninsula of Haiti and the Barahona Peninsula, Dominican Republic. *9th Caribbean Geological Conference, Santo Domingo, Dominican Republic. Caribbean Geology, Collected Contributions*, **1**, 161–174.
McLaughlin P.P., van den Bold W.A., & Mann P. (1991) Geology of the Azua and Enriquillo basins, Dominican Republic; 1, Neogene lithofacies, biostratigraphy, biofacies, and paleogeography. *Geologic and Tectonic Development of the North American–Caribbean Plate Boundary in Hispaniola* (ed. by P. Mann, G. Draper, and J. Lewis), pp. 337–366. The Geological Society of America, Inc., Boulder, CO.
Mejía M. (2006) Flora de la Española: conocimiento actual y estado de conservación. Actas IX Congreso Latinoamericano de Botánica; el 18-25 de Junio, Santo Domingo, Ed. Jardín Botánico Nacional Dr. Rafael Ma. Moscoso,. 11–12.
Miller M.A., Pfeiffer W., & Terri Schwartz (2010) Available at: https://www.phylo.org/.
Papadopulos A.S.T., Kaye M., Devaux C., Hipperson H., Lighten J., Dunning L.T., Hutton I., Baker W.J., Butlin R.K., & Savolainen V. (2014) Evaluation of genetic isolation within an island flora reveals unusually widespread local adaptation and supports sympatric speciation. *Philosophical Transactions of the Royal Society B*, **369**, 20130342.
Paris J.R., Stevens J.R., & Catchen J.M. (2017) Lost in parameter space: A road map for Stacks. *Methods in Ecology and Evolution*, **8**, 1360–1373.
Pennington R.T. & Lavin M. (2016) The contrasting nature of woody plant species in different neotropical forest biomes reflects differences in ecological stability. *New Phytologist*, **210**, 25–37.
Pennington T. (2003) Monograph of *Andira* (Leguminosae-Papilionoideae). *Systematic Botany Monographs*, **64**, 1–145.
Pritchard J.K., Stephens M., & Donnelly P. (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.
Quiroga M.P., Mathiasen P., Iglesias A., Mill R.R., & Premoli A.C. (2016) Molecular and fossil evidence disentangle the biogeographical history of *Podocarpus*, a key genus in plant geography. *Journal of Biogeography*, **43**, 372–383.
Quiroga M.P. & Premoli A.C. (2010) Genetic structure of *Podocarpus nubigena* (Podocarpaceae) provides evidence of Quaternary and ancient historical events. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **285**, 186–193.
Rieseberg L.H.. & Brouillet L. (1994) Are Many Plant Species Paraphyletic? *Taxon*, **43**, 21–32.
Rodríguez-Peña R.A., Jestrow B., Meerow A.W., Clase T., Jiménez-Rodríguez F., Griffith M.P., Santiago-Valentín E., Sustache-Sustache J.A., & Francisco-Ortega J. (2014) Genetic diversity and differentiation of *Pseudophoenix* (Arecaceae) in Hispaniola. *Botanical Journal of the Linnean Society*, **176**, 469–485.
Ronquist F., Teslenko M., van der Mark P., Ayres D.L., Darling A., Höhna S., Larget B., Liu L., Suchard M.A., & Huelsenbeck J.P. (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology*, **61**, 539–542.

Rosenberg N.A. (2004) DISTRUCT: a program for the graphical display of population structure. *Molecular Ecology Notes*, **4**, 137–138.

Schlüter P.M., Ruas P.M., Kohl G., Ruas C.F., Stuessy T.F., & Paulus H.F. (2011) Evidence for progenitor – derivative speciation in sexually deceptive orchids. *Annals of Botany*, **108**, 895–906.

Sly N.D., Townsend A.K., Rimmer C.C., Jason M., Latta S.C., & Lovette I.J. (2011) Ancient islands and modern invasions: disparate phylogeographic histories among Hispaniola's endemic birds. *Molecular Ecology*, **20**, 5012–5024.

Sly N.D., Townsend A.K., Rimmer C.C., Townsend J.M., Latta S.C., & Lovette I.J. (2010) Phylogeography and conservation of the endemic Hispaniolan Palm-Tanagers (Aves: *Phaenicophilus*). *Conservation Genetics*, **11**, 2121–2129.

Stamatakis A. (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*, **22**, 2688–2690.

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

van Strien M.J., Holderegger R., & van Heck H.J. (2015) Isolation-by-distance in landscapes: considerations for landscape genetics. *Heredity*, **114**, 27–37.

Townsend A.K., Rimmer C.C., Latta S.C., & Lovette I.J. (2007) Ancient differentiation in the single-island avian radiation of endemic Hispaniolan chat-tanagers (Aves: *Calyptophilus*). *Molecular Ecology*, **16**, 3634–3642.

Turvey S.T., Peters S., Brace S., Young R.P., Crumpton N., Hansford J., Nuñez-Miño J.M., King G., Tsalikidis K., Ottenwalder J.A., Timpson A., Funk S.M., Brocca J.L., Thomas M.G., & Barnes I. (2016) Independent evolutionary histories in allopatric populations of a threatened Caribbean land mammal. *Diversity and Distributions*, **22**, 589–602.

Valtueña F.J., Rodríguez-Riaño T., López J., Mayo C., & Ortega-Olivencia A. (2017) Peripatric speciation in an endemic Macaronesian plant after recent divergence from a widespread relative. *PLoS ONE*, **12**, e0178459.

Winterton C., Richardson J.E., Hollingsworth M., Clark A., Zamora N., & Pennington T. (2014) Historical biogeography of the neotropical legume genus *Dussia*: the Andes, the Panama Isthmus and the Chocó.

Table 4.1. Sample collection information for *Podocarpus* in Hispaniola. Coll. Site = collection site (see Figure 4.2), with acronyms in parenthesis indicating the designated population used for DIYABC. # indiv = number of individuals collected per site.

Species	Coll. Site	Locality	Latitude	Longitude	# indiv
Podocarpus buchii	1 (SB)	Barahona Province, Sierra Bahoruco Oriental	18º 0.915´ N	71º 17.017´ W	5
	2 (SB)	Pedernales Province, Sierra Bahoruco Occidental	18º 12.495´ N	71º 33.353´ W	1
	3 (SN)	Loma del Quince, Independencia Province, Sierra de Neiba	18º 40.515´ N	71º 41.182´ W	5
	4 (SN)	Loma del Quince, Independencia Province, Sierra de Neiba	18º 41.478 ´ N	71º 47.258´ W	5
	7 (CC-b)	Constanza area, La Vega Province, Cordillera Central	18º 58.040´ N	70º 47.770´ W	2
	8 (CC-b)	Constanza area, La Vega Province, Cordillera Central	18º 43.565´ N	70º 52.255´ W	2
Podocarpus hispaniolensis	5 (CC-h)	San Jose de Ocoa Province, Rancho Arriba, Cordillera Central	18º 40.193´ N	70º 21.792´ W	4
	6 (CC-h)	San Jose de Ocoa Province, Loma del Firme, Cordillera Central	18º 43.687´ N	70º 22.530´ W	2
	9 (CC-h)	Ebano Verde Scientific Reserve, Casabito, La Vega Province, Cordillera Central	19º 3.3´ N	70º 34.291´ W	1
	10 (CS)	Yaroa, Puerto Plata Province, Cordillera Septentrional	19º 35.283´ N	70º 36.047´ W	5
	11 (CS)	Tenares, La Jibara, Salcedo Province, Cordillera Septentrional	19º 29.605´ N	70º 20.005´ W	4

Table 4.2. Genetic diversity statistics for the five *Podocarpus* populations and the 11 collection sites. H_o = Observed Heterozygosity, H_e = Expected Heterozygosity, π = nucleotide polymorphism diversity. Numbers in bold indicate $H_o > H_e$. SB, SN and CC – b in Population ID column, and 1, 2, 3, 4, 7 and 8 in collection site column correspond to *P. buchii*. CC- h and CS, and 5, 6, 9, 10 and 11 correspond to *P. hispaniolensis*.

5 populations				**11 collection sites**			
Population ID	H_o	H_e	π	**Collection site**	H_o	H_e	π
SB	0.2039	0.3343	0.394	1	0.1904	0.303	0.3894
SN	0.1432	0.3434	0.3863	2	**0.2532**	**0.1266**	0.2532
CC- b	0.173	0.3116	0.4031	3	0.158	0.3515	0.4559
CC - h	0.1404	0.3197	0.3755	4	0.141	0.3051	0.3943
CS	0.1524	0.2999	0.3382	5	0.1024	0.2529	0.3305
				6	**0.1716**	**0.0984**	0.1799
				7	**0.2404**	**0.1394**	0.2611
				8	**0.1665**	**0.1057**	0.1888
				9	**0.23**	**0.115**	0.23
				10	0.1822	0.3005	0.3874
				11	**0.1301**	**0.2319**	0.2999

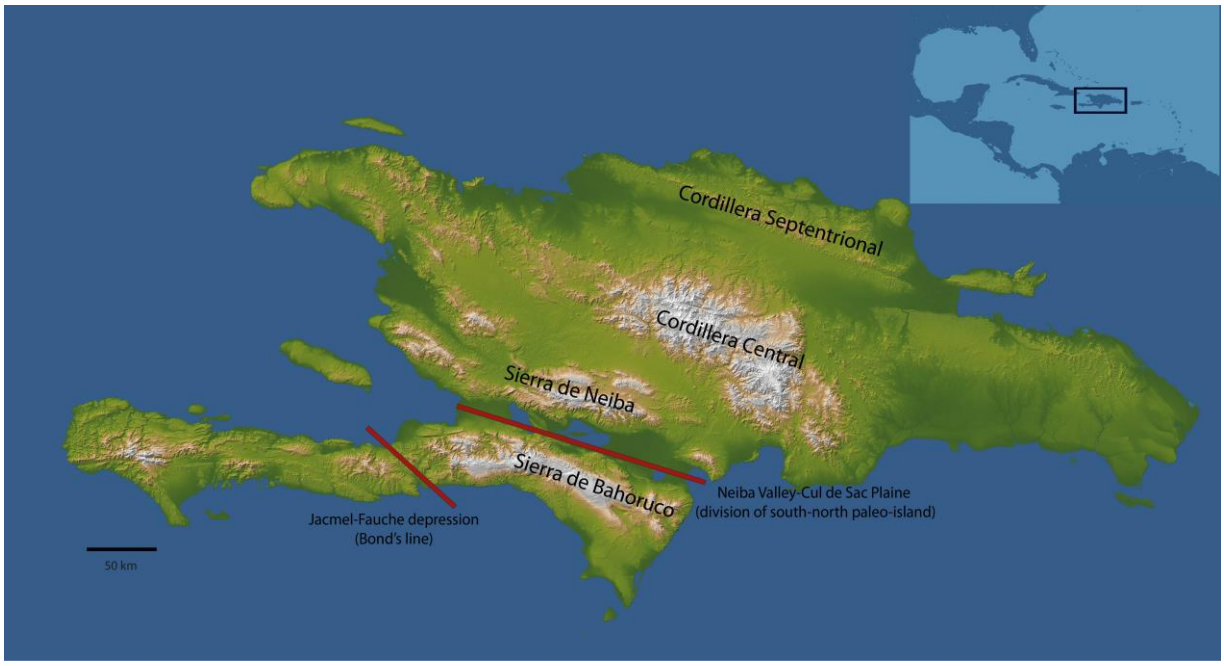


Figure 4.1. Topographical map of Hispaniola indicating the main mountain ranges studied in this work. Red lines indicate the location of the Neiba Valley-Cul de Sac Plain that was subject to marine incursions during the Pleistocene separating the south and north palaeo-islands, and Jacmel-Fauche depression (Bond's line), which divides the west and east of the southern palaeo-island. Inset shows location of Hispaniola within the Caribbean and tropical America.

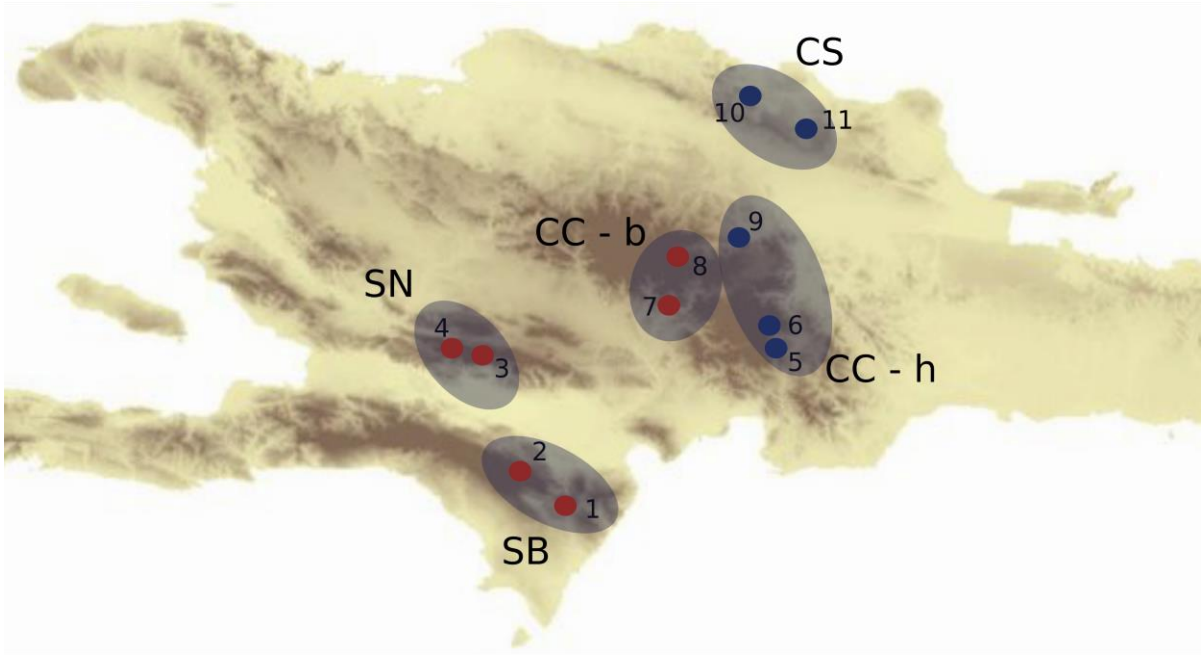


Figure 4.2. Sampling collection sites marked 1 through 11 for *Podocarpus* in Hispaniola. Red dots indicate collection sites of *P. buchii* and blue dots collections of *P. hispaniolensis*. Grey circles indicate populations designated for DIYABC analysis as follows: SB = Sierra Bahoruco; SN = Sierra Neiba; CC-b = Cordillera Central *P. buchii* individuals; CC-h = Cordillera Central *P. hispaniolensis* individuals; and CS = Cordillera Septentrional.

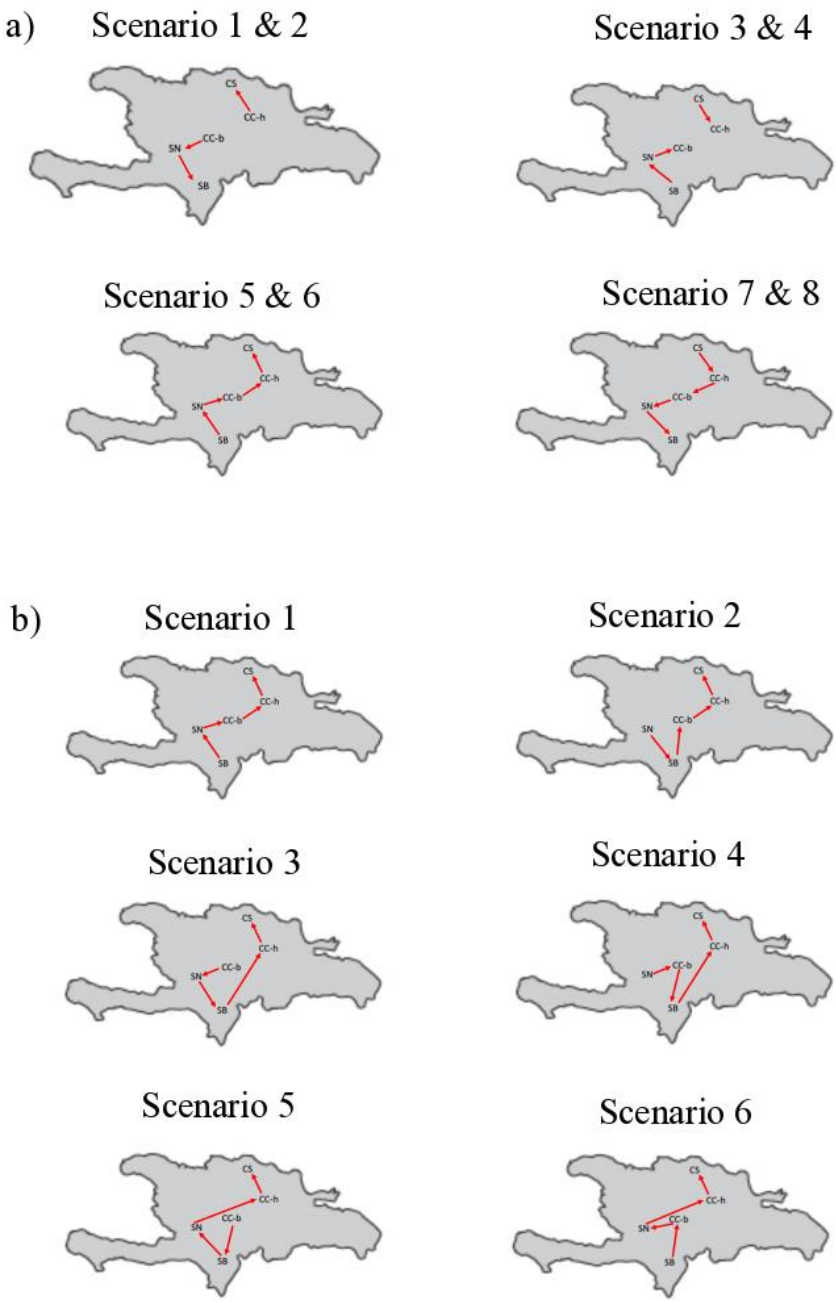
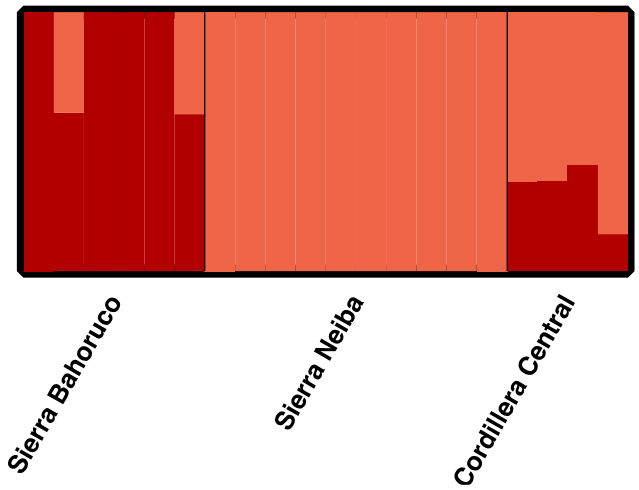


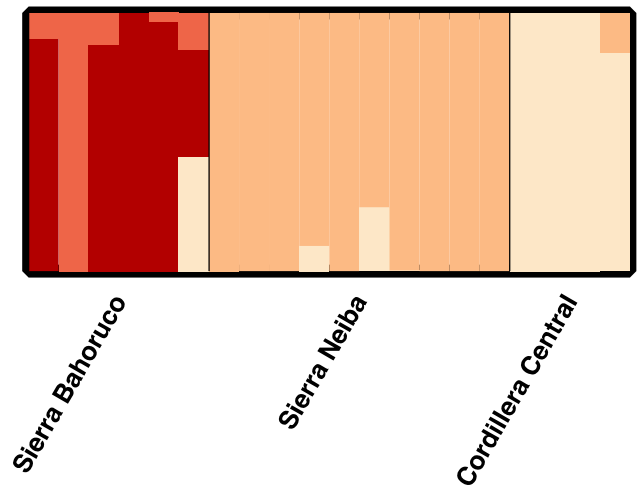
Figure 4.3. a) Schematic diagram showing the demographic scenarios compared in the first DIYABC for *Podocarpus* in Hispaniola. Scenarios 2, 4, 6 and 8 have the same topology as scenarios 1, 3, 5 and 7 but with bottlenecks. b) Schematic diagram showing the demographic scenarios compared in the second DIYABC with bottlenecks.

a) *P. buchii*

K = 2

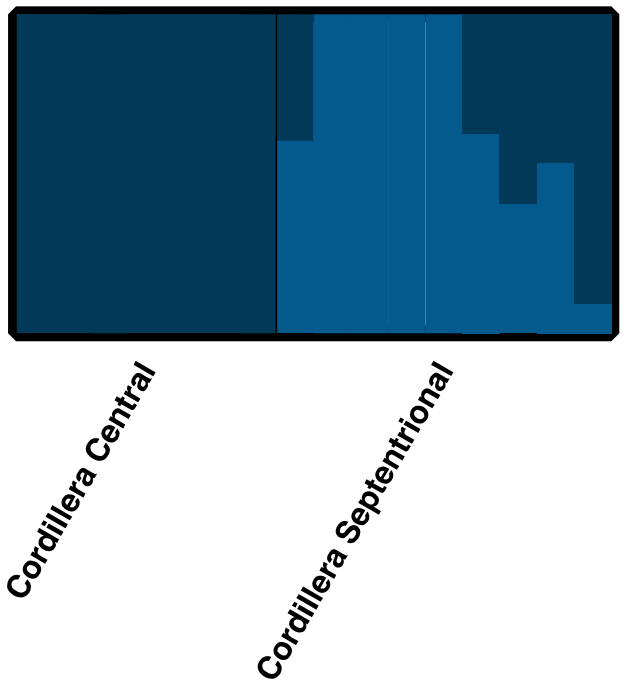


K = 4

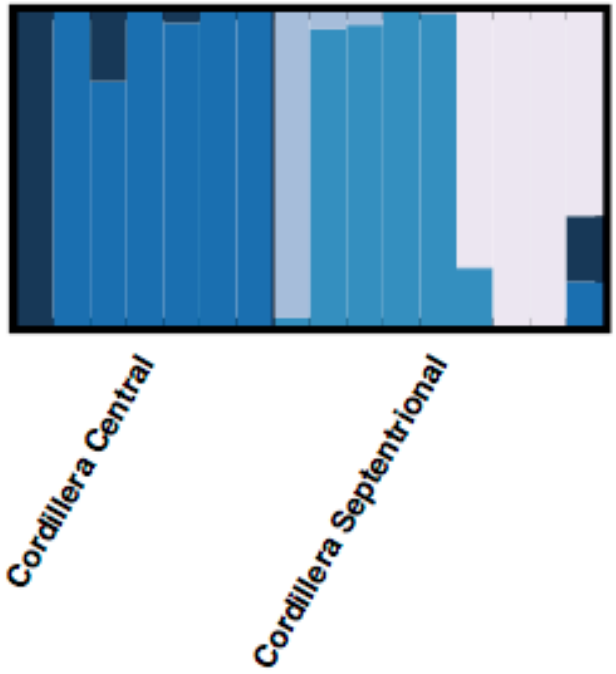


b) *P. hispaniolensis*

K = 2



K = 5



c) both species combined. Numbers indicate collection sites.

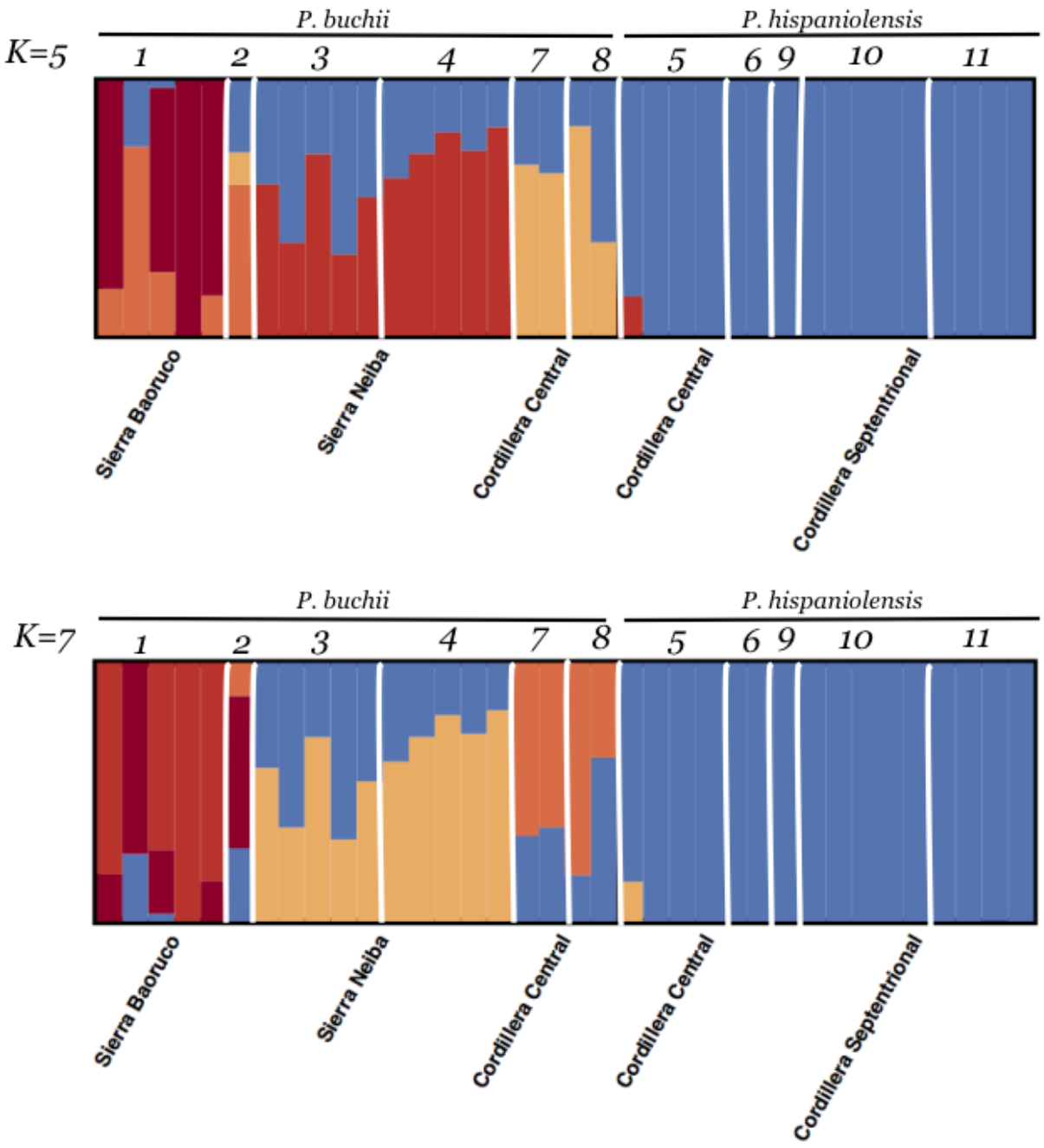


Figure 4.4. Genetic STRUCTURE of *Podocarpus* in Hispaniola. Bar plots show: a) $k = 2$ and $k = 4$ for *P. buchii*; b) $k = 2$ and $k = 5$ for *P. hispaniolensis;* and c) $k = 5$ and $k = 7$ for both species combined, numbers indicate collection sites.

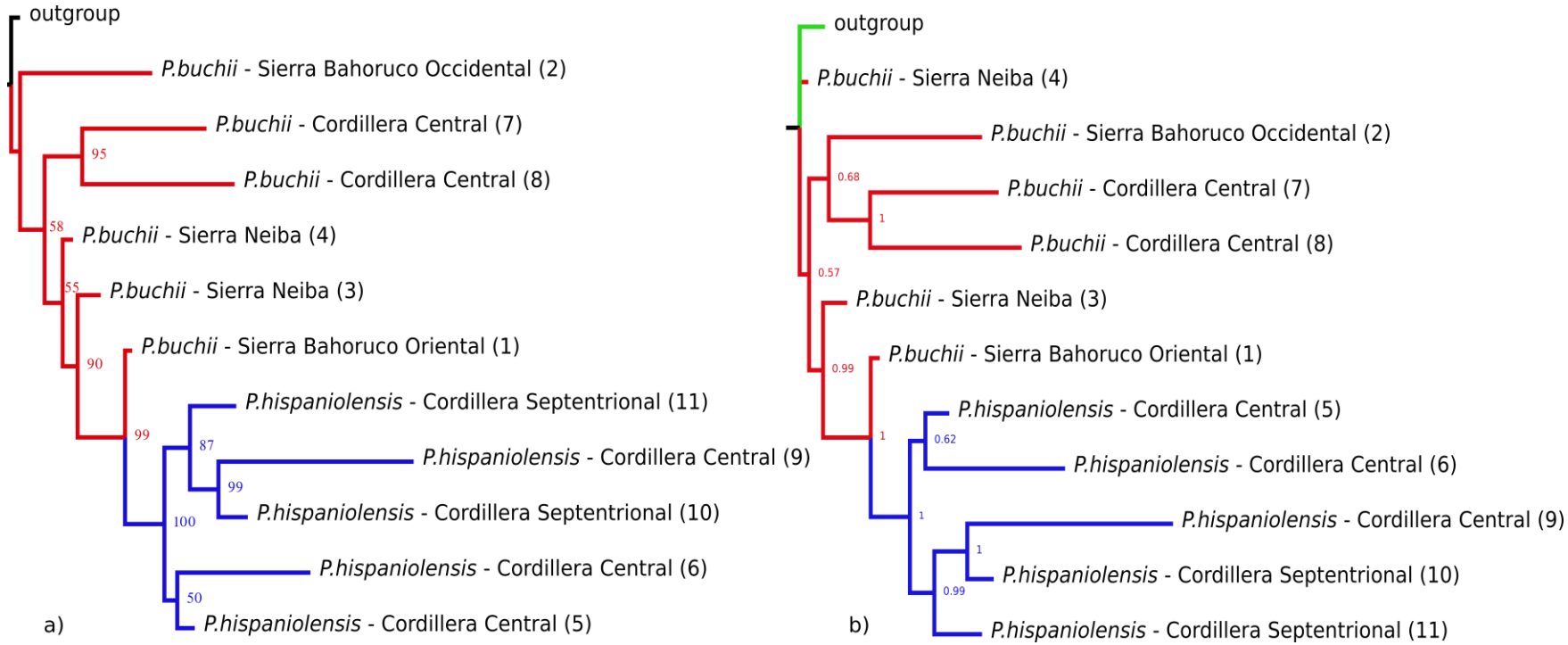


Figure 4.5. a) Maximum likelihood (RAxML) phylogenetic tree of the 11 collection sites. Numbers at nodes are bootstrap support values. b) Bayesian (MrBayes) phylogenetic tree of the 11 collection sites. Numbers at nodes are posterior probability branch support values. Numbers in parenthesis indicate collection sites. *P. buchii* in red and *P. hispaniolensis* in blue

CHAPTER 5

General Conclusions

5.1. Summary

My research has explored the evolutionary history of endemic plant lineages in the Caribbean region, a hotspot of biodiversity (Mittermeier et al., 2004), with a high level of endemism of genera and species (Acevedo-Rodríguez & Strong, 2008). The rich Caribbean flora has provided material for numerous floristic studies that have established links between the island biota and that of continental America. Hypotheses to explain the origin of the Caribbean biota have two main processes, vicariance and dispersal, which act at spatial scales from local and regional to continental. I have explored both processes in a hypothesis-testing framework.

My overall objective was to elucidate the geographic origin and colonization times of endemic Caribbean genera. Within this broad context, I wanted to reconstruct the biogeographical history and diversification patterns of the tropical conifer genus *Podocarpus* at the scale of the Caribbean archipelago, and in particular an experimental elucidation of the genetic structure and historical demography of the two species of *Podocarpus* endemic to the island of Hispaniola. I tested various biogeographical hypotheses including vicariance, dispersal, and the progression rule of speciation. This thesis therefore combines original fieldwork with published DNA sequences and new large-scale NGS sequence data, combined in a phylogenetic / biogeographical analysis of the origin and assembly of Caribbean flora.

This thesis is to my knowledge the first attempt to elucidate the biogeographic origin of endemic plant genera to the Caribbean based on a single combined phylogenetic analysis. The large fraction of unsampled endemic genera in molecular phylogenies

illustrates the gaps in systematic and biogeographic knowledge of the Caribbean flora (Francisco-Ortega et al., 2007).

My research does not support the landbridge GAARlandia as a colonization route to the Antilles for plants, because crown and stem ages of endemic genera with ancestors distributed in South America do not correspond to the hypothesized timeframe. Additional geological evidence on the timing and geomorphology of this proposed land bridge awaits to be combined with molecular dating and biogeographic modeling approaches for a larger number of endemic lineages.

As a specific example, I consider the biogeography of Caribbean *Podocarpus*. The islands distribution of the genus is the result of a single Oligocene colonization from South America that dates back ca. 30 Ma. According to this hypothesis, *Podocarpus* species of the Greater Antilles diversified in the early Miocene (ca. 20 Ma), and those of the Lesser Antilles originated from an ancestor in the Greater Antilles through over-water dispersal. The progression rule hypothesis is not supported by the data and analysis presented here, as the Lesser Antilles (or parts of them) were already above water when colonization of the Greater Antilles (older island system) occurred. Thus, colonization of the younger Lesser Antilles did not happen as they emerged. The inter-island diversification of the genus in the Greater Antilles exemplifies the complexity of Caribbean biota assembly processes, where vicariance of Cuba and Hispaniola and over-water dispersal to Jamaica both played a role. Contrary to predictions based on the availability of empty niches on islands, no higher diversification rates were found for

insular taxa, and island colonization did not trigger shifts in diversification rates for Caribbean *Podocarpus*.

Finally, I explored the within-island diversification pattern of *Podocarpus* at a population level on the smaller spatial scale of the island of Hispaniola, and the role of dispersal barriers on such diversification. Genetic diversity within the two endemic species, *P. buchii* and *P. hispaniolensis*, is geographically structured, with distinct genetic clusters occurring in each of the main Cordillera systems. This structure does not seem to have been influenced by marine incursions, the fusion of two palaeo-islands, or mountain uplift events as shown for other Hispaniolan taxa. Instead, the genetic structure seems to be due to more recent topographical barriers such as the dry valleys that separate the different Cordilleras. An ancestor of Hispaniolan *Podocarpus* colonized the island in the south and moved northwards by a series of stepping-stones. A speciation event occurred in Central Cordillera leading to the formation of *P. hispaniolensis*, which is a monophyletic clade nested within a paraphyletic *P. buchii*. This is an example of the Progenitor-Derivative Hypothesis, that is believed to explain the evolution of the rich biodiversity in the Neotropics.

My research exemplifies the complexity of biota assemblage in the Caribbean, as the result of multiple processes like vicariance, dispersal and *in situ* speciation. The geological complexity of the Caribbean has provided opportunities for lineages to diverge and diversify, and the proximity of the islands to continental America has most likely played an important role on the assembly of Caribbean flora through dispersal.

5.2. Future directions

This research has added to the empirical knowledge of the historical biogeography of plants in the Caribbean. The power of NGS molecular data to test alternative biogeographical hypotheses in the region has been shown. A better understanding of the geological formation of the islands and more accurate timing of island emergence and the rise and fall of land bridges will be key to draw more accurate conclusions on the origin of Caribbean flora.

Further molecular analysis of these unsampled taxa and their continental relatives would allow tests of the mixture of relatively recent (Oligocene–Miocene) but also older (Paleocene–Eocene) lineages that gave rise to the extant Caribbean endemic flora, and of predominantly Antillean ancestors of endemic genera.

Biome evolution has recently received more attention from biogeographers and evolutionary biologists. Biomes are broad biogeographical regions with a particular biota. On a geological scale, biomes originate and are subject to ecological, geological and climatological conditions that will shape their biota composition. Because of this dynamism to which they are subject, they can be used to answer macro-evolutionary questions. In this sense, biomes are useful systems to understand the biota assembly of a region because, as organismal lineages, they will reflect an evolutionary response to environmental, geological, and biological processes (Crisp, 2006). One approach is the use of a time-calibrated phylogenetic framework. Studies that have used phylogenies to test hypotheses about biome evolution show that species biome conservatism through time is more common than traditional expectations of niche shifts (Prinzing et al., 2001; Wiens & Graham, 2005). However, it has been also shown that species can shift from an

ancestral biome through time, as occurs in Amazonia (Antonelli et al., 2018). Areas for future research on biome evolution of endemic Caribbean taxa are identification of particular biomes where endemic plant lineages are found and use of a phylogenetic framework to test for shifts from an ancestral biome and direction of migration of those shifts. Also, elucidating if endemic plant lineages show biome conservatism.

An understanding of inter-island migrations requires fully-resolved phylogenies of Caribbean lineages. Conventional single- or multi-locus DNA sequence data might not be able to fully resolve the phylogenetic relationships of close relatives, whereas newer, high-throughput DNA sequencing methods should be used to generate fully-resolved topologies. Ecological niche modeling (ENM) could also aid in the study of speciation events, such as that shown here in the Central Cordillera. ENM might detect niche divergence between the two *Podocarpus* species, which might shed light into speciation mechanisms. Also, testing gene flow explicitly between populations and species could help elucidating reproductive isolation of populations.

5.3. References

Acevedo-Rodríguez P. & Strong M.T. (2008) Floristic richness and affinities in the West Indies. *Botanical Review*, **74**, 5–36.

Antonelli A., Zizka A., Carvalho F.A., Scharn R., Bacon C.D., Silvestro D., & Condamine F.L. (2018) Amazonia is the primary source of Neotropical biodiversity. *Proceedings of the National Academy of Sciences*, **115**, 6034 LP-6039.

Crisp M. (2006) Biome assembly: what we know and what we need to know. *Journal of Biogeography*, **33**, 1332–1333.

Francisco-Ortega J., Santiago-Valentín E., Acevedo-Rodríguez P., Lewis C., Pipoly III J., Meerow A.W., & Maunder M. (2007) Seed plant genera endemic to the Caribbean Island Biodiversity Hotspot: A review and a molecular phylogenetic perspective. *The Botanical Review*, **73**, 183–234.

Mittermeier R.A., Gil R.R., Hoffman M., Pilgrim J., T. B., Mittermeier C.G., Lamoreux J., & da Fonseca G.A.B. (2004) *Hotspots revisited: Earth's biologically richest and most threatened terrestrial ecoregions.* CEMEX, Mexico D.F.

Prinzing A., Durka W., Klotz S., & Brandl R. (2001) The niche of higher plants: evidence for phylogenetic conservatism. *Proceedings. Biological sciences*, **268**, 2383–2389.

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.

Appendix A

Appendix for Chapter 2: Historical Biogeography of endemic seed plant genera in the Caribbean: Did GAARlandia play a role?

Figure A1. Maximum clade credibility tree obtained in BEAST showing mean ages in Ma and 95% HPDs
https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1002%2Fece3.3521&file=ece33521-sup-0001-FigS1.pdf

Figure A2. MCC tree obtained in BEAST showing node numbers
https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1002%2Fece3.3521&file=ece33521-sup-0002-FigS2.pdf

Figure A3. Ancestral area reconstruction using DEC*j* model for complex model 1
https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1002%2Fece3.3521&file=ece33521-sup-0003-FigS3.pdf

Appendix B

Appendix for Chapter 4: Evolutionary history of *Podocarpus* in Hispaniola

Table B1. Voucher specimen information for *Podocarpus* samples collected in Hispaniola Island and outgroup. Herbarium acronym where specimens were deposited indicated in parenthesis. JBSD = Jardín Botánico Santo Domingo; E = Royal Botanic Garden Edinburgh Herbarium; RBGE = Royal Botanic Garden Edinburgh living collections; MBC = Montgomery Botanical Garden; HAC = Herbario de la Academia de Ciencias, La Habana.

Species	Collector number	Herbarium or living collection accession number	Location
P. buchii	Nieto-Blázquez, 150	128142 (JBSD)	Dominican Republic, Barahona Province
P. buchii	Nieto-Blázquez, 151	128179 (JBSD)	Dominican Republic, Barahona Province
P. buchii	Nieto-Blázquez, 152	128144 (JBSD)	Dominican Republic, Barahona Province
P. buchii	Nieto-Blázquez, 153	128143 (JBSD)	Dominican Republic, Barahona Province
P. buchii	Nieto-Blázquez, 154	128178 (JBSD)	Dominican Republic, Barahona Province
P. buchii	Nieto-Blázquez, 155	128149 (JBSD)	Dominican Republic, Pedernales Province
P. buchii	Nieto-Blázquez, 156	128184 (JBSD)	Dominican Republic, Independencia Province
P. buchii	Nieto-Blázquez, 157	128166 (JBSD)	Dominican Republic, Independencia Province
P. buchii	Nieto-Blázquez, 158	128157 (JBSD)	Dominican Republic, Independencia Province
P. buchii	Nieto-Blázquez, 159	128146 (JBSD)	Dominican Republic, Independencia Province
P. buchii	Nieto-Blázquez, 160	128145 (JBSD)	Dominican Republic, Independencia Province
P. buchii	Nieto-Blázquez, 161	128158 (JBSD)	Dominican Republic, Independencia Province
P. buchii	Nieto-Blázquez, 162	128148 (JBSD)	Dominican Republic, Independencia Province
P. buchii	Nieto-Blázquez, 163	128176 (JBSD)	Dominican Republic, Independencia Province

P. buchii	Nieto-Blázquez, 164	128186 (JBSD)	Dominican Republic, Independencia Province
P. buchii	Nieto-Blázquez, 165	128189 (JBSD)	Dominican Republic, Independencia Province
P. buchii	Nieto-Blázquez, 172	128164 (JBSD)	Dominican Republic, La Vega Province
P. buchii	Nieto-Blázquez, 173	128185 (JBSD)	Dominican Republic, La Vega Province
P. buchii	Nieto-Blázquez, 174	128159 (JBSD)	Dominican Republic, La Vega Province
P. buchii	Nieto-Blázquez, 175	128154 (JBSD)	Dominican Republic, La Vega Province
P. hispaniolensis	Nieto-Blázquez, 166	128188 (JBSD)	Dominican Republic, San Jose de Ocoa
P. hispaniolensis	Nieto-Blázquez, 167	128173 (JBSD)	Dominican Republic, San Jose de Ocoa
P. hispaniolensis	Nieto-Blázquez, 168	128187 (JBSD)	Dominican Republic, San Jose de Ocoa
P. hispaniolensis	Nieto-Blázquez, 169	128141 (JBSD)	Dominican Republic, San Jose de Ocoa
P. hispaniolensis	Nieto-Blázquez, 170	128181 (JBSD)	Dominican Republic, San Jose de Ocoa
P. hispaniolensis	Nieto-Blázquez, 171	128183 (JBSD)	Dominican Republic, San Jose de Ocoa
P. hispaniolensis	Nieto-Blázquez, 176	128170 (JBSD)	Dominican Republic, La Vega Province
P. hispaniolensis	Nieto-Blázquez, 177	waiting for info (JBSD and E)	Dominican Republic, Puerto Plata
P. hispaniolensis	Nieto-Blázquez, 178	waiting for info (JBSD and E)	Dominican Republic, Puerto Plata
P. hispaniolensis	Nieto-Blázquez, 179	128150 (JBSD)	Dominican Republic, Puerto Plata
P. hispaniolensis	Nieto-Blázquez, 180	128156 (JBSD)	Dominican Republic, Puerto Plata
P. hispaniolensis	Nieto-Blázquez, 181	128153 (JBSD)	Dominican Republic, Puerto Plata
P. hispaniolensis	Nieto-Blázquez, 182	128151 (JBSD)	Dominican Republic, Salcedo Province
P. hispaniolensis	Nieto-Blázquez, 183	128155 (JBSD)	Dominican Republic, Salcedo Province
P. hispaniolensis	Nieto-Blázquez, 184	128168 (JBSD)	Dominican Republic, Salcedo Province
P. hispaniolensis	Nieto-Blázquez, 185	128169 (JBSD)	Dominican Republic, Salcedo Province
P. urbanii	Gardner & Knees, 6381	20011359* (RBGE)	Jamaica

P. purdieanus	(donated by RBGE)	20080621*B (MCB)	Jamaica
P. aristulatus	Nieto-Blázquez, 186	43224 (HAC)	Cuba
P. ekmanii	Nieto-Blázquez, 187	43222 (HAC)	Cuba
P. angustifolius	R. Oviedo	43221 (HAC)	Cuba
P. coriaceus	Gardner & Knees, 6622	20030490*A7 (RBGE)	Trinidad and Tobago
P. trinitensis	Gardner & Knees, ????	20030492*A1 (RBGE)	Trinidad and Tobago
P. guatemalensis	unknown source	20140203 (MCB)	Belize
P. matudae	waiting for info	20140481*A (RBGE)	Wild collected unknown
P. oleifolius	GANIAL, 66	E00617101 (E)	Bolivia
P. sellowii	waiting for info	20071743*A (RBGE)	Brazil

Table B2. Pairwise F_{ST} values among the 11 *Podocarpus* collection sites as obtained from *Stacks* v. 1.47 in upper side of matrix diagonal. Distance between collection sites in meters in lower side of matrix diagonal as obtained from Geographic Distance Matrix Generator (Ersts, 2011).

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
1		0.0692	0.0786	0.0815	0.0653	0.0502	0.0591	0.0618	0.0805	0.1246	0.0938
2	35,936.39		0.0422	0.0444	0.0263	0	0	0	0	0.1043	0.1034
3	84,905.23	53,781.37		0.0475	0.0851	0.0581	0.0431	0.0693	0.0640	0.1370	0.1311
4	92,192.86	59079.7	10827.34		0.0967	0.0750	0.0424	0.0604	0.0672	0.1602	0.1544
5	121527.04	136031.98	139542.17	150232.21		0.0174	0.0329	0.0544	0.0556	0.1161	0.0730
6	124504.69	137412.38	138345.21	148946.99	6610.97		0.0036	0.0042	0	0.0771	0.0829
7	117818.01	116472.43	99273.32	108888.36	56371.01	51702.65		0.0026	0	0.1093	0.0801
8	90346.24	92485.66	86169.06	96733.78	53900.07	52230.31	27986.62		0	0.1066	0.0892
9	117818.01	116472.43	99273.32	108888.36	56371.01	51702.65	0	27986.62		0.0731	0.0549
10	189292.03	183603.43	152839.5	159824.23	105219.77	98614.92	72083.34	100069.88	72083.34		0.1143
11	192633.04	192488.86	168976.32	177128.51	91728.84	85307.72	76127.32	102434.33	76127.32	29961.91	

Table B3. DIYABC model comparison of eight demographic scenarios under direct and logistic approaches. Posterior probabilities [95% highest posterior densities]

Direct approach

closest	scenario 1	scenario 2	scenario 3	scenario 4	scenario 5	scenario 6	scenario 7	scenario 8
50	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.46 [0.02,0.89]	0.00 [0.00,0.00]	0.54 [0.10,0.97]
100	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.48 [0.04,0.91]	0.00 [0.00,0.00]	0.52 [0.08,0.95]
150	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.00 [0.00,0.00	0.00 [0.00,0.00]	0.42 [0.00,0.86]	0.00 [0.00,0.00]	0.57 [0.13,1.00]
200	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.43 [0.00,0.86]	0.00 [0.00,0.00]	0.57 [0.13,1.00]
250	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.00 [0.00,0.00	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.44 [0.00,0.87]	0.00 [0.00,0.00]	0.56 [0.12,0.99]
300	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.45 [0.01,0.88]	0.00 [0.00,0.00]	0.54 [0.11,0.98]
350	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.46 [0.03,0.90]	0.00 [0.00,0.00]	0.53 [0.09,0.96]
400	0.00 [0.00,0.00]	0.00 [0.00,0.00	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.46 [0.03,0.90]	0.00 [0.00,0.00]	0.53 [0.09,0.96]
450	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.47 [0.03,0.91]	0.00 [0.00,0.00]	0.52 [0.08,0.96]
500	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.47 [0.02,0.91]	0.00 [0.00,0.00]	0.52 [0.08,0.95]

Logistic approach

n	scenario 1	scenario 2	scenario 3	scenario 4	scenario 5	scenario 6	scenario 7	scenario 8
8000	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.96 [0.94,0.98]	0.00 [0.00,0.00]	0.03 [0.01,0.05]
16000	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.97 [0.96,0.98]	0.00 [0.00,0.00]	0.02 [0.01,0.03]
24000	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.97 [0.97,0.98]	0.00 [0.00,0.00]	0.02 [0.01,0.02]
32000	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.98 [0.97,0.98]	0.00 [0.00,0.00]	0.01 [0.01,0.02]
40000	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.98 [0.97,0.98]	0.00 [0.00,0.00]	0.01 [0.01,0.02]
48000	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.98 [0.97,0.98]	0.00 [0.00,0.00]	0.01 [0.01,0.02]
56000	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.98 [0.98,0.98]	0.00 [0.00,0.00]	0.01 [0.01,0.06]
64000	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.98 [0.98,0.98]	0.00 [0.00,0.00]	0.01 [0.01,0.01]
72000	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.98 [0.98,0.98]	0.00 [0.00,0.00]	0.01 [0.01,0.01]
80000	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.00 [0.00,0.00]	0.98 [0.98,0.98]	0.00 [0.00,0.00]	0.01 [0.01,0.01]

Table B4. Model checking for best-fit model (scenario 6) for first DIYABC analysis. HMO=mean of complete genic diversity, FMO= mean of complete F_{ST} distances between two populations, NMO= mean of complete Nei's genetic distances.

Summary Statistics	Observed Value	Proportion of simulated<observed
HMO_1_SB	0.3902	1.0000 (***)
HMO_1_SN	0.3850	1.0000 (***)
HMO_1_CC - b	0.3280	0.9925 (**)
HMO_1_ CC - h	0.3535	0.9975 (**)
HMO_1_CS	0.3140	0.9760 (*)
FMO_1_ SB & SN	0.0813	0.9365
FMO_1_ SB & CC - b	0.0923	0.9080
FMO_1_ SB & CC - h	0.0906	0.9370
FMO_1_ SB & CS	0.1141	0.9585 (*)
FMO_1_ SN & CC - b	0.1192	0.9970 (**)
FMO_1_ SN & CC - h	0.1012	0.9970 (**)
FMO_1_ SN & CS	0.1719	1.0000 (***)
FMO_1_ CC - b & CC - h	0.1066	0.9980 (**)
FMO_1_ CC - b & CS	0.0973	0.9910 (**)
FMO_1_ CC - h & CS	0.1325	1.0000 (***)
NMO_1_ SB & SN	0.0770	0.9495
NMO_1_ SB & CC - b	0.0850	0.9180
NMO_1_ SB & CC - h	0.0986	0.9505 (*)
NMO_1_ SB & CS	0.1014	0.9635 (*)
NMO_1_ SN & CC - b	0.1137	0.9985 (**)
NMO_1_ SN & CC - h	0.1095	0.9990 (***)
NMO_1_ SN & CS	0.1531	1.0000 (***)
NMO_1_ CC - b & CC - h	0.1070	0.9990 (***)
NMO_1_ CC - b & CS	0.0868	0.9910 (**)
NMO_1_ CC – h & CS	0.1167	1.0000 (***)

Table B5. Parameter estimation for scenario 6 based on 1 million simulated datasets. q = quantiles for mean posterior value; N = effective population sizes; t = coalescent times in number of generations, and NA= ancestral unsampled population size.

Parameter	mean	median	mode	q025	q050	q250	q750	q950	q975
N1	2.86E+03	5.43E+02	3.04E+01	2.25E+01	3.15E+01	1.58E+02	1.85E+03	1.31E+04	2.44E+04
N2	5.11E+03	6.11E+02	1.00E+01	1.00E+01	1.00E+01	1.05E+02	3.09E+03	2.83E+04	4.94E+04
N3	8.68E+03	9.58E+02	1.00E+01	1.00E+01	1.47E+01	1.18E+02	7.03E+03	5.22E+04	7.07E+04
N4	1.35E+04	1.59E+03	1.00E+01	1.53E+01	2.10E+01	1.69E+02	1.49E+04	7.20E+04	8.57E+04
N5	5.92E+03	3.08E+02	1.00E+01	1.00E+01	1.00E+01	3.90E+01	3.51E+03	3.65E+04	5.54E+04
t1	2.35E+01	1.53E+01	1.00E+01	1.00E+01	1.01E+01	1.19E+01	2.37E+01	6.28E+01	8.78E+01
db	1.43E+03	3.13E+02	1.00E+01	1.10E+01	1.24E+01	4.58E+01	1.77E+03	6.95E+03	8.22E+03
Nf5	4.61E+03	3.34E+02	6.18E+01	4.22E+01	5.35E+01	1.24E+02	1.69E+03	2.63E+04	5.36E+04
t2	3.71E+01	2.26E+01	1.35E+01	1.12E+01	1.18E+01	1.57E+01	3.89E+01	1.10E+02	1.47E+02
Nf4	1.58E+04	4.09E+03	2.82E+02	2.43E+02	3.22E+02	1.18E+03	1.88E+04	7.80E+04	9.02E+04
t3	5.99E+01	3.40E+01	1.52E+01	1.26E+01	1.37E+01	2.09E+01	6.53E+01	1.84E+02	2.48E+02
Nf3	3.65E+03	3.93E+02	3.46E+01	2.41E+01	3.21E+01	1.11E+02	1.68E+03	1.79E+04	3.66E+04
t4	6.69E+01	3.04E+01	1.50E+01	1.16E+01	1.24E+01	1.85E+01	5.91E+01	1.89E+02	2.92E+02
Nf2	7.10E+03	1.18E+03	6.87E+01	4.11E+01	6.74E+01	3.20E+02	4.91E+03	4.13E+04	6.31E+04
td	5.45E+03	5.77E+03	9.99E+03	1.82E+02	3.54E+02	2.38E+03	8.59E+03	9.84E+03	9.93E+03
NA	7.88E+04	8.55E+04	9.85E+04	2.64E+04	3.59E+04	6.86E+04	9.46E+04	9.91E+04	9.96E+04

Table B6. First DIYABC analysis measure of performance for scenario 6. N = effective population sizes; t = coalescent times in number of generations, and NA= ancestral unsampled population size.

Parameter	TRUE values	Mean Relative Bias	square root of the relative mean integrated square error (RRMISE)
N1	1.12E+04	3.99E+01	2.67E+02
N2	9.96E+03	3.34E+01	2.52E+02
N3	1.17E+04	6.31E+01	3.79E+02
N4	1.03E+04	8.47E+01	4.30E+02
N5	1.16E+04	9.61E+01	4.71E+02
t1	1.19E+02	1.24E+00	5.22E+00
db	1.60E+03	2.03E+01	7.40E+01
Nf5	1.12E+04	5.30E+01	3.46E+02
t2	4.91E+02	1.28E+00	4.91E+00
Nf4	9.15E+03	5.71E+01	3.52E+02
t3	1.53E+03	1.15E+00	4.40E+00
Nf3	9.46E+03	7.56E+01	4.11E+02
t4	3.93E+03	1.19E+00	5.27E+00
Nf2	1.24E+04	7.45E+01	4.26E+02
td	1.48E+03	1.79E+01	6.96E+01
NA	1.17E+04	2.18E+01	2.14E+02

Table B7. Second DIYABC model comparison of 6 demographic scenarios under direct and logistic approaches. Posterior probabilities [95% highest posterior densities].

Direct approach

closest	scenario 1	scenario 2	scenario 3	scenario 4	scenario 5	scenario 6
50	0.08 [0.00,0.31]	0.24 [0.00,0.61]	0.10 [0.00,0.36]	0.32 [0.00,0.72]	0.20 [0.00,0.55]	0.06 [0.00,0 .26]
100	0.12 [0.00,0.40]	0.24 [0.00,0.61]	0.14 [0.00,0.44]	0.26 [0.00,0.64]	0.15 [0.00,0.46]	0.09 [0.00,0.34]
150	0.14 [0.00,0.44]	0.24 [0.00,0.61]	0.15 [0.00,0.46]	0.24 [0.00,0.62]	0.14 [0.00,0.44]	0.08 [0.00, 0.31]
200	0.15 [0.00,0.47]	0.24 [0.00,0.61]	0.15 [0.00,0.47]	0.21 [0.00,0.56]	0.15 [0.00,0.46]	0.09 [0.00,0.34]
250	0.15 [0.00,0.46]	0.21 [0.00,0.57]	0.15 [0.00,0.47]	0.22 [0.00,0.58]	0.14 [0.00,0.45]	0.10 [0.00, 0.37]
300	0.14 [0.00,0.45]	0.21 [0.00,0.56]	0.16 [0.00,0.48]	0.23 [0.00,0.59]	0.15 [0.00,0.46]	0.10 [0.00,0.37]
350	0.14 [0.00,0.46]	0.21 [0.00,0.57]	0.15 [0.00,0.46]	0.22 [0.00,0.58]	0.14 [0.00,0.46]	0.11 [0.00 ,0.39]
400	0.15 [0.00,0.46]	0.22 [0.00,0.58]	0.13 [0.00,0.43]	0.21 [0.00,0.57]	0.14 [0.00,0.45]	0.12 [0.00, 0.41]
450	0.14 [0.00,0.46]	0.21 [0.00,0.57]	0.13 [0.00,0.43]	0.20 [0.00,0.56]	0.15 [0.00,0.47]	0.13 [0.00, 0.43]
500	0.15 [0.00,0.46]	0.21 [0.00,0.56]	0.13 [0.00,0.43]	0.21 [0.00,0.57]	0.15 [0.00,0.47]	0 .13 [0.00,0.43]

Logistic approach

n	scenario 1	scenario 2	scenario 3	scenario 4	scenario 5	scenario 6
6000	0.02 [0.01,0.03]	0.89 [0.86,0.91]	0.00 [0.00,0.00]	0.07 [0.05,0.09]	0.00 [0.00,0.00]	0.00 [0.00,0.00]
12000	0.03 [0.00,0.21]	0.89 [0.87,0.91]	0.00 [0.00,0.00]	0.06 [0.00,0.26]	0.00 [0.00,0.18]	0.00 [0. 00,0.18]
18000	0.03 [0.03,0.04]	0.89 [0.88,0.91]	0.00 [0.00,0.00]	0.06 [0.05,0.07]	0.00 [0.00,0.00]	0.00 [0.00,0.00]
24000	0.04 [0.03,0.04]	0.89 [0.88,0.91]	0.00 [0.00,0.00]	0.05 [0.04,0.06]	0.00 [0.00,0.00]	0.00 [0. 00, 0.00]
30000	0.04 [0.03,0.05]	0.89 [0.88,0.90]	0.00 [0.00,0.00]	0.05 [0.04,0.06]	0.00 [0.00, 0.00]	0.00 [0.00, 0.00]
36000	0.04 [0.04,0.05]	0.89 [0.89,0.90]	0.00 [0.00,0.00]	0.05 [0.04,0.05]	0.00 [0.00, 0.00]	0.00 [0. 00, 0.00]
42000	0.04 [0.04,0.05]	0.89 [0.89,0.90]	0.00 [0.00,0.00]	0.04 [0.04,0.05]	0.00 [0.00, 0.00]	0.00 [0. 00, 0.00]
48000	0.05 [0.04,0.05]	0.89 [0.89,0.90]	0.00 [0.00,0.00]	0.04 [0.04,0.05]	0.00 [0.00, 0.00]	0 .00 [0.00, 0.00]
54000	0.05 [0.04,0.05]	0.89 [0.88,0.90]	0.00 [0.00, 0.00]	0.04 [0.04,0.05]	0.00 [0.00, 0.00]	0.00 [0. 00, 0.00]
60000	0.05 [0.05,0.05]	0.89 [0.88,0.90]	0.00 [0.00,0.00]	0.04 [0.04,0.05]	0.00 [0.00, 0.00]	0.00 [0.00, 0.00]

Table B8. Model checking for best-fit model (scenario 2) for second DIYABC analysis. HMO=mean of complete genic diversity, FMO= mean of complete F_{ST} distances between two populations, NMO= mean of complete Nei's genetic distances.

Summary Statistics	Observed value	Proportion (simulated<observed)	
HMO_1_SB	0.3902	1	(***)
HMO_1_SN	0.385	1	(***)
HMO_1_CC - b	0.328	0.8875	
HMO_1_CC - h	0.3535	0.986	(*)
HMO_1_CS	0.314	0.7945	
FMO_1_SB&SN	0.0813	0.693	
FMO_1_SB&CC - b	0.0923	0.9	
FMO_1_SB&CC - h	0.0906	0.979	(*)
FMO_1_SB&CS	0.1141	0.9845	(*)
FMO_1_SN&CC - b	0.1192	0.87	
FMO_1_SN&CC - h	0.1012	0.7555	
FMO_1_SN&CS	0.1719	0.9995	(***)
FMO_1_CC - b&CC - h	0.1066	0.96	(*)
FMO_1_CC - b&CS	0.0973	0.9295	
FMO_1_CC - h&CS	0.1325	0.999	(***)
NMO_1_SB&SN	0.077	0.7695	
NMO_1_SB&CC - b	0.085	0.8625	
NMO_1_SB&CC - h	0.0986	0.9805	(*)
NMO_1_SB&CS	0.1014	0.9865	(*)
NMO_1_SN&CC - b	0.1137	0.9285	
NMO_1_SN&CC - h	0.1095	0.918	
NMO_1_SN&CS	0.1531	1	(***)
NMO_1_CC - b&CC - h	0.107	0.968	(*)
NMO_1_CC - b&CS	0.0868	0.906	
NMO_1_CC - h&CS	0.1167	0.9975	(**)

Table B9. Parameter estimation for scenario 2 based on 1 million simulated datasets. q = quantiles for mean posterior value; N = effective population sizes; t = coalescent times in number of generations, and NA= ancestral unsampled population size.

Parameter	mean	median	mode	q025	q050	q250	q750	q950	q975
N1	9.42E+03	1.22E+03	1.00E+01	1.00E+01	1.59E+01	1.64E+02	8.12E+03	5.35E+04	7.39E+04
N2	7.25E+02	3.09E+02	8.18E+01	3.61E+01	4.83E+01	1.37E+02	6.75E+02	2.11E+03	3.37E+03
N3	6.60E+03	5.03E+02	1.00E+01	1.00E+01	1.00E+01	6.30E+01	4.43E+03	3.95E+04	6.09E+04
N4	6.44E+03	3.87E+02	1.00E+01	1.00E+01	1.00E+01	4.65E+01	4.02E+03	3.94E+04	6.09E+04
N5	1.62E+04	2.38E+03	1.80E+01	1.86E+01	2.61E+01	2.31E+02	2.07E+04	8.10E+04	9.02E+04
t1	2.01E+01	1.44E+01	1.00E+01	1.00E+01	1.00E+01	1.16E+01	2.08E+01	4.73E+01	6.48E+01
db	1.20E+03	2.36E+02	1.00E+01	1.07E+01	1.19E+01	4.04E+01	1.29E+03	6.37E+03	8.02E+03
Nf5	3.15E+03	2.01E+02	5.04E+01	3.05E+01	3.76E+01	7.84E+01	9.52E+02	1.53E+04	3.56E+04
t2	2.58E+01	1.81E+01	1.25E+01	1.05E+01	1.09E+01	1.35E+01	2.78E+01	6.50E+01	8.99E+01
Nf4	1.45E+04	3.45E+03	3.32E+02	2.12E+02	2.87E+02	1.00E+03	1.57E+04	7.48E+04	8.90E+04
t3	4.20E+01	2.70E+01	1.56E+01	1.17E+01	1.25E+01	1.77E+01	4.60E+01	1.19E+02	1.69E+02
Nf3	3.19E+03	2.90E+02	4.50E+01	2.41E+01	3.02E+01	8.86E+01	1.33E+03	1.53E+04	3.24E+04
t4	9.26E+01	5.41E+01	2.26E+01	1.50E+01	1.72E+01	3.02E+01	1.07E+02	2.90E+02	4.09E+02
Nf2	1.19E+04	3.08E+03	2.45E+02	1.32E+02	2.06E+02	9.09E+02	1.19E+04	6.23E+04	8.12E+04
td	8.26E+02	3.50E+02	5.19E+01	3.01E+01	4.14E+01	1.41E+02	8.73E+02	3.37E+03	4.95E+03
NA	8.61E+04	9.17E+04	9.97E+04	4.39E+04	5.38E+04	8.04E+04	9.69E+04	9.95E+04	9.98E+04

Table B10. Second DIYABC analysis measure of performance for scenario 2. N = effective population sizes; t = coalescent times in number of generations, and NA= ancestral unsampled population size.

Parameter	TRUE values	Mean Relative Bias	square root of the relative mean integrated square error (RRMISE)
N1	1.04E+04	5.86E+01	3.61E+02
N2	1.07E+04	2.19E+00	1.82E+01
N3	9.26E+03	8.71E+01	4.73E+02
N4	1.12E+04	9.88E+01	4.98E+02
N5	9.41E+03	1.09E+02	5.11E+02
t1	6.31E+01	9.04E-01	3.67E+00
db	1.54E+03	2.02E+01	7.21E+01
Nf5	1.06E+04	3.57E+01	2.53E+02
t2	2.54E+02	1.18E+00	4.25E+00
Nf4	1.13E+04	5.24E+01	3.40E+02
t3	7.63E+02	1.24E+00	4.23E+00
Nf3	1.19E+04	6.02E+01	3.73E+02
t4	1.98E+03	1.47E+00	5.09E+00
Nf2	1.05E+04	7.28E+01	4.14E+02
td	4.66E+03	1.18E+00	5.77E+00
NA	1.03E+04	5.60E+01	3.64E+02

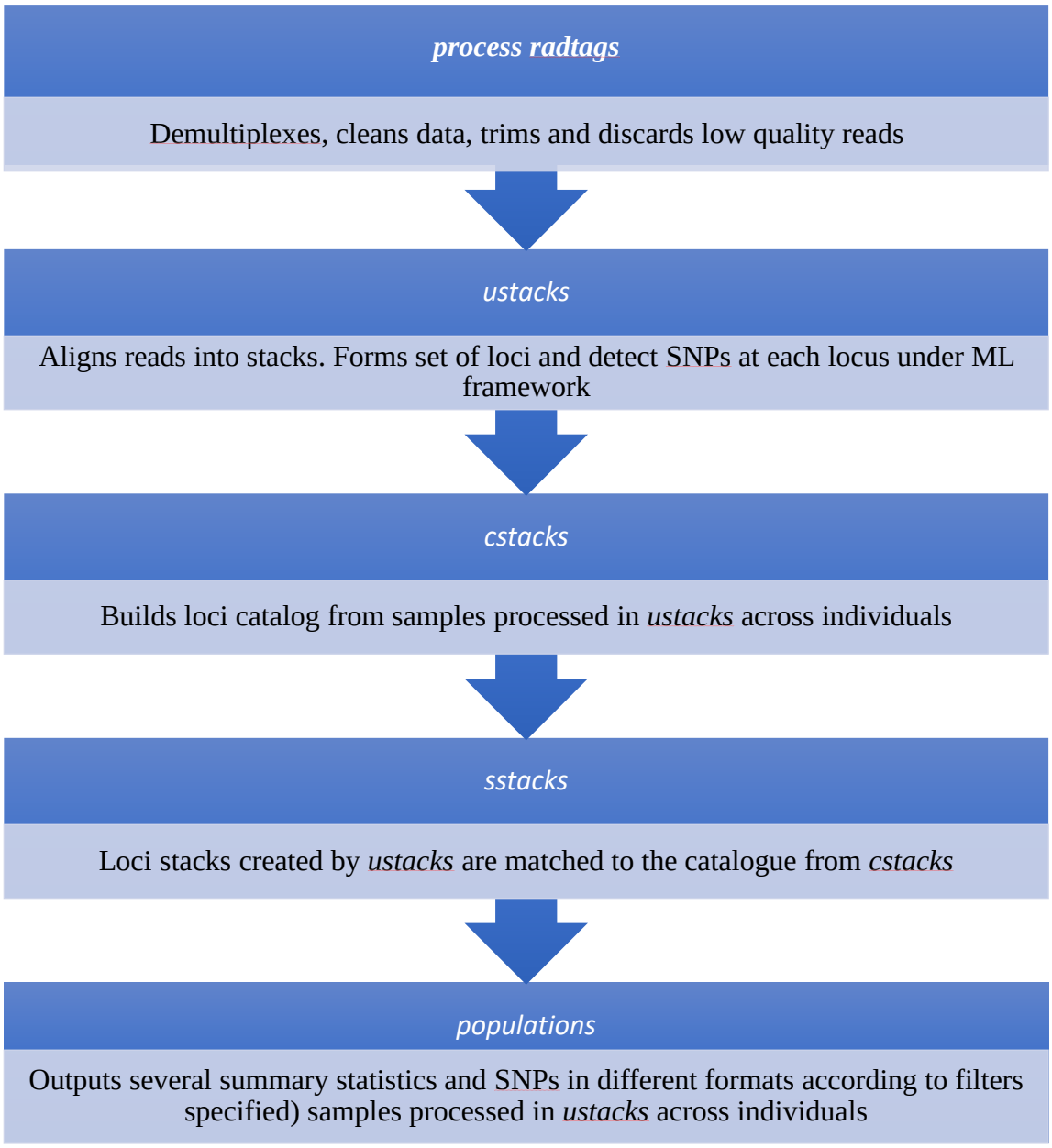


Figure B1. *Stacks* workflow for the *de novo* assembly of *Podocarpus* DNA reads from Hispaniola island.

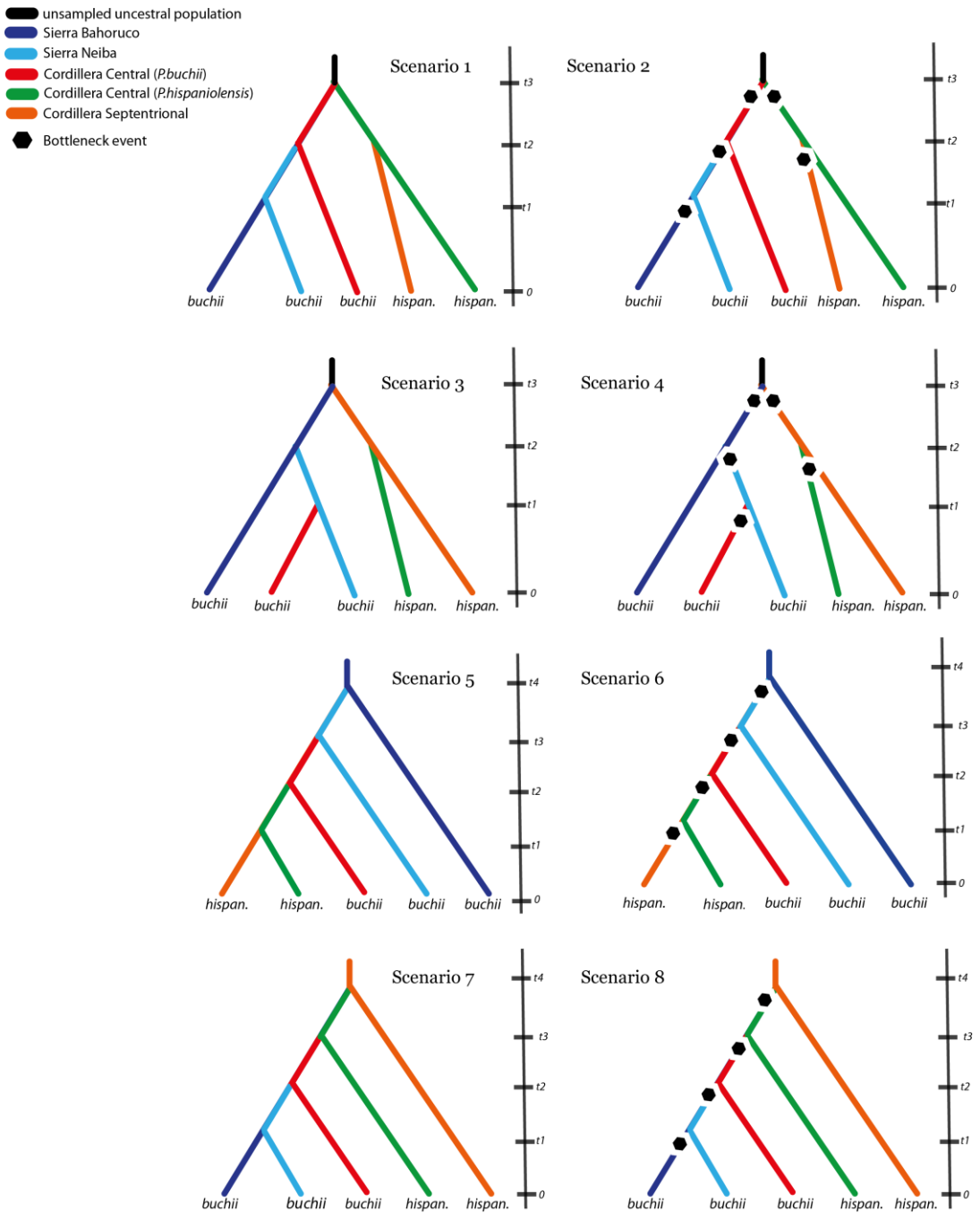


Figure B2. Demographic scenarios used in the first DIYABC analysis for *Podocarpus* in Hispaniola. Colored branches indicate populations from the different mountainous ranges. Bottleneck events indicated by black star. See methods section for description of each scenario.

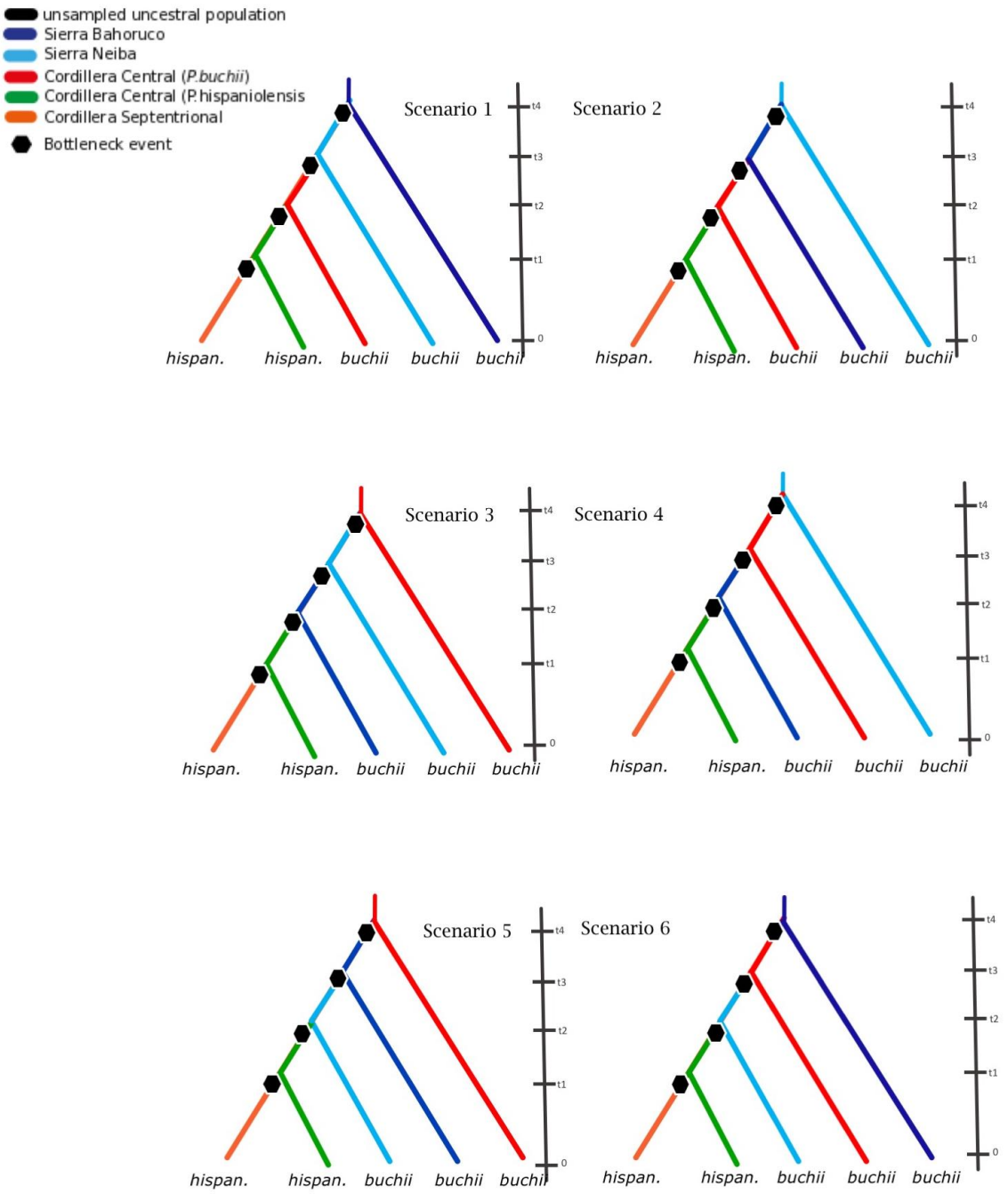
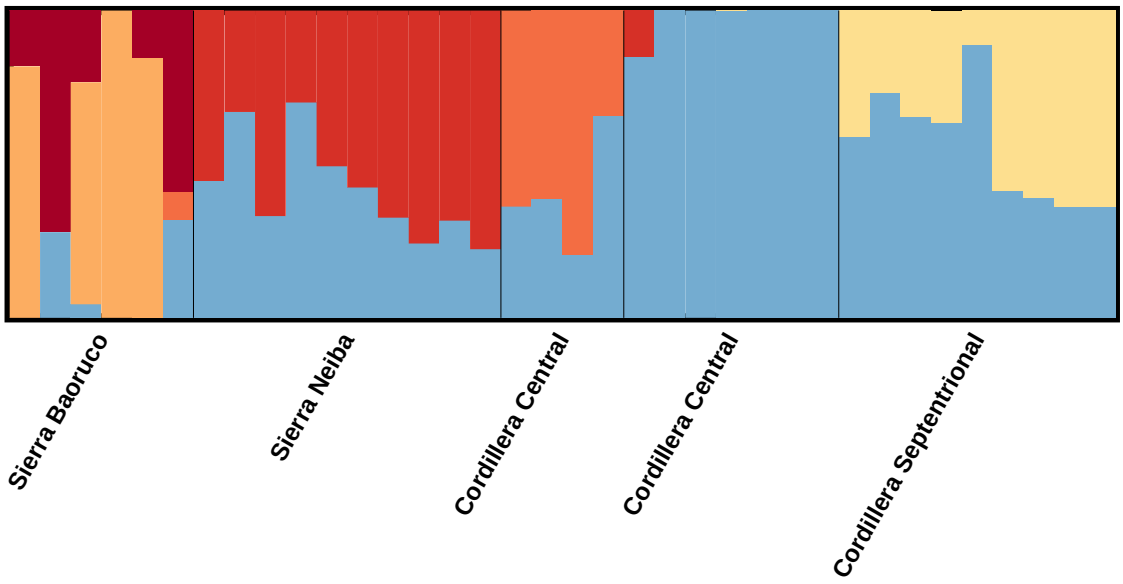


Figure B3. Demographic scenarios used in the second DIYABC analysis for *Podocarpus* in Hispaniola. Colored branches indicate populations from the different mountainous ranges. Bottleneck events indicated by black star. See methods section for description of each scenario.

k = 9



k = 10

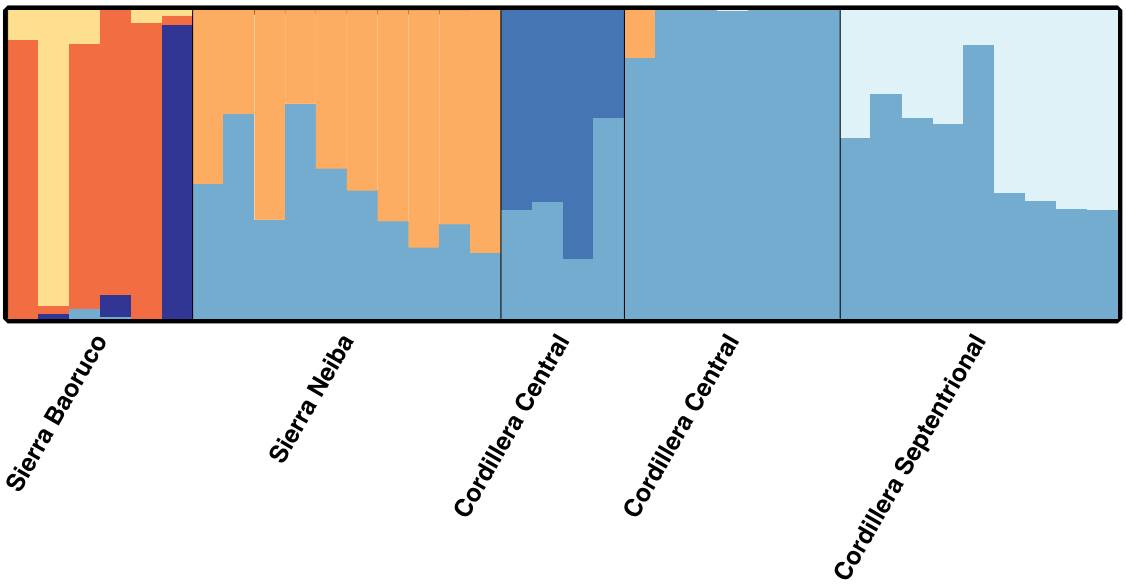


Figure B4. Genetic STRUCTURE of *Podocarpus* in Hispaniola island. Bar plots show *k* = 9 and *k* = 10 for both species combined.

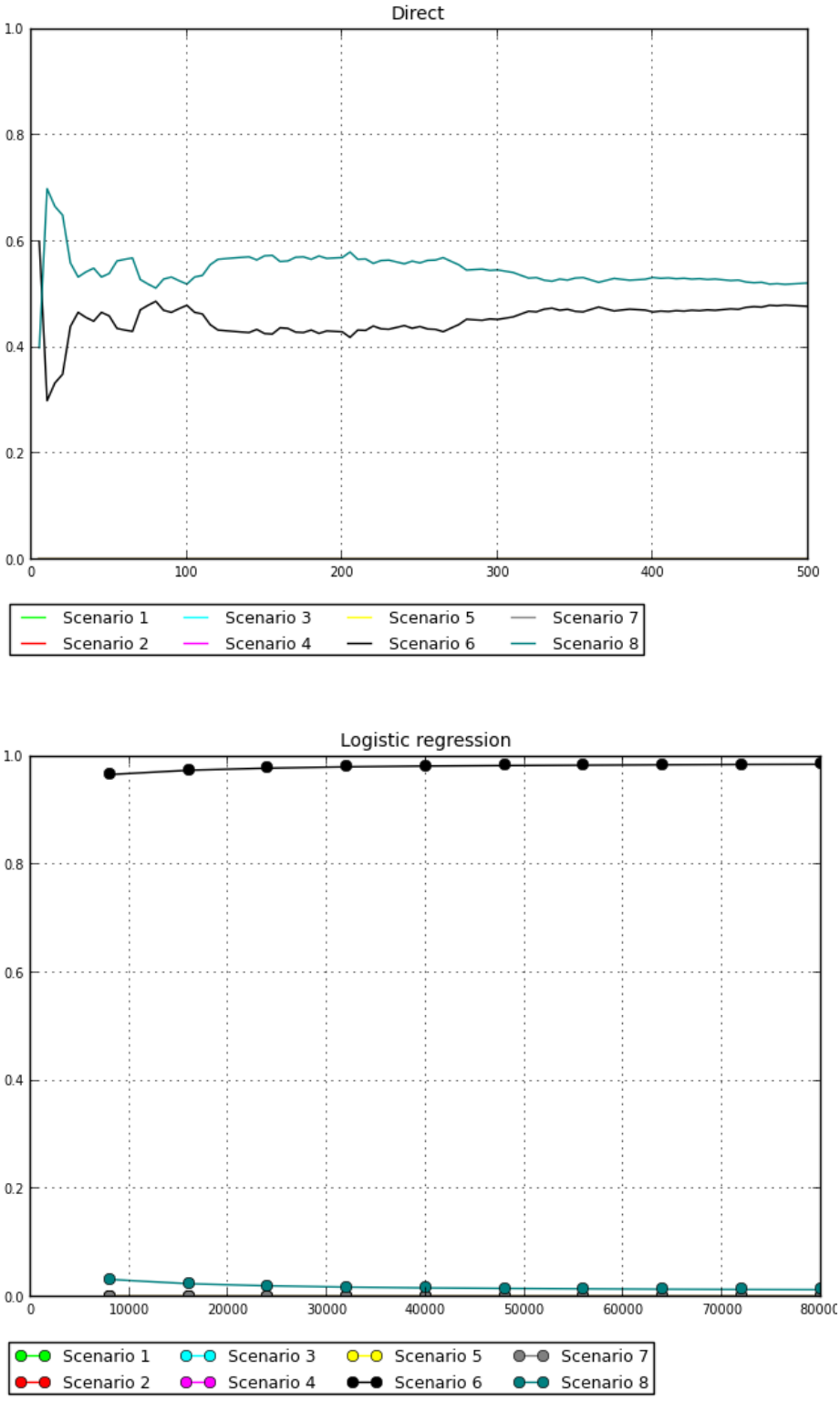


Figure B5. Model comparison for first DIYABC analysis under direct and logistic approach.

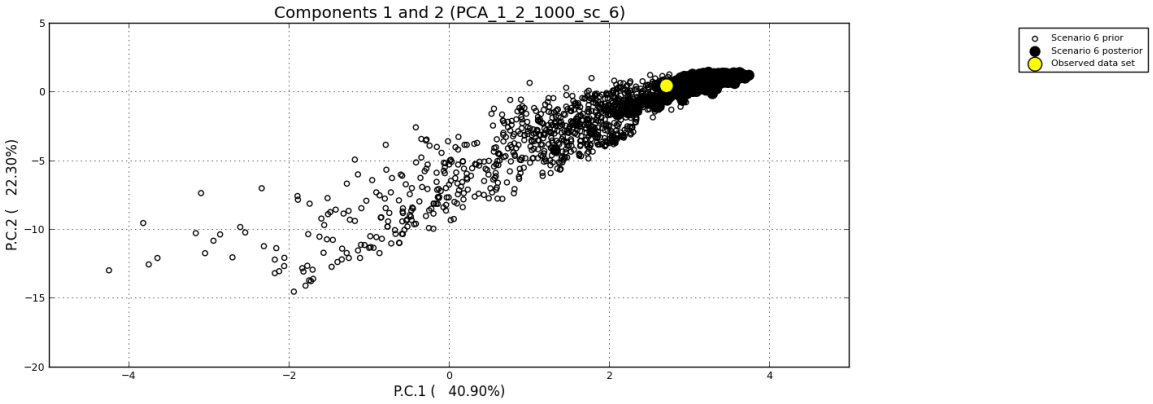


Figure B6. Model checking for scenario 6 (first DIYABC analysis).

Appendix References

Ersts P.J. (2011) Geographic Distance Matrix Generator v1.2.3. American Museum of Natural History, Center for Biodiversity and Conservation. Available at http://biodiversityinformatics.amnh.org/open_source/gdmg. Data accesed September 2018.