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

María Esther Nieto-Blázquez1 | Alexandre Antonelli2,3,4 | Julissa Roncal1

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2017 The Authors. Ecology and Evolution published by John Wiley & Sons Ltd.

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. We 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, we 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 our sampling, we 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.

KEYWORDS

Antilles, Cenozoic, colonization, dispersal, island biogeography, West Indies

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 endemity 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 (Lomolino, Riddle, & Whittaker, 2017; Whittaker, Fernández-Palacios, Matthews, Borregaard, & Triantis, 2017; Whittaker, Triantis, & Ladle, 2008). According to the new synthesis in island biogeography theory
(Lomolino et al., 2017), 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, Mittermeier, Mittermeier, Fonseca, & Kent, 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 (Graham, 2003; Iturralde-Vinent & MacPhee, 1999). Moreover, climatic change, through cooling or warming periods (Zachos, Pagani, Sloan, Thomas, & Billups, 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, Steinbauer, Cabral, & Kreft, 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 southeast 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, Crawford, & Bermingham, 2012), invertebrates (Binford et al., 2008), vertebrates (Hulsey, Keck, & Hollingsworth, 2011), and also plants, as shown for the genus Styrax (Styracaceae, Fritsch, 2003), Moacrotor (Euphorbiaceae, van Ee, Berry, Riina, & Gutiérrez Amaro, 2008), and Copernicia (Arecaceae, Bacon, Baker, & Simmons, 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, Cezilly, Moreau, Motreuil, & Wattier, 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., Appelhans, Kessler, Smets, Razafimandimbison, & Janssens, 2012; Jestrow, Rodriguez, & Francisco-Ortega, 2010; Jestrow, Valdés, et al., 2012), revealing a complex biogeographic history (Roncal, Zona, & Lewis, 2008). Some endemic genera have sister taxa that are widely distributed in continental America (Lavin, Wojciechowski, Gasson, Hughes, & Wheeler, 2003; Lavin, Wojciechowski, et al., 2001; Rova, Delprete, Andersson, & Albert, 2002; Wurdack, Hoffmann, & Chase, 2005), others have relatives with a more restricted continental distribution (Lavin, Pennington, et al., 2001; Baldwin et al., 2002; Wojciechowski, Lavin, & Sanderson, 2004), and a few are sister to taxa that are native to regions outside the Neotropics, such as Africa (Lavin, Pennington, et al., 2001; Lavin, Wojciechowski, et al., 2001), Polynesia (Kimball & Crawford, 2004), and New Caledonia (Motley, Wurdack, & Delprete, 2005).

With the aim of providing insights into the origin and evolution of the Caribbean flora, we targeted endemic seed plant genera. We
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 nonendemic 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 speculation like divergent selection and geographic isolation can generate evolutionary significant units above the species level (Barracough, 2010; Barracough & Humphreys, 2015). Plant genera can therefore also be used as units of biodiversity.

We 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? Our hypotheses are (1) GAARlandia played a major role as a migration route in the colonization of the Caribbean Islands. Under this hypothesis, we 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 (Bacon, Silvestro, Jaramillo, Tilston, & Chakrabarty, 2015; Weir, Bermingham, & Schluter, 2009). 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.

2 | MATERIAL AND METHODS

2.1 | Taxon sampling selection

We 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. We 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, we have included 52% endemic genera for which there were molecular phylogenies available at the time of the publication. We 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, we selected species that represented the entire distributional range of the genus, and with complete sequences available in the Zanne et al. (2014) matrix.

2.2 | DNA sequence selection and alignment

Of the seven gene regions available in Zanne et al. (2014), we selected four (18S rDNA, atpB, matK, and rbcL) for our alignment. We excluded the 26S rDNA region because it was not well represented (only 17 sequences were available for our 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. 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, Pfeiffer, & Terri Schwartz, 2010). Manual trimming and concatenation of gene regions were performed in GENEIOUS v. 7.1.9 (Kearse et al., 2012). Our 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 1). We had 37% missing nucleotide data in this final alignment.

2.3 | Phylogenetic reconstruction and dating

We 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. We selected six gymnosperms in the Zamiaceae family to root the tree, which included the monotypic endemic Caribbean genus Microcycas and five Zamia species. We used JMODELTEST2 v.0.1.1 (Darriba, Taboada, Doallo, & Posada, 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 + I’, which was selected for subsequent analyses.

In order to estimate absolute divergence times, we inferred a time-calibrated phylogenetic tree using a Bayesian inference (BI) approach as implemented in BEAST 2.3.1 (Bouckaert et al., 2014). Analysis on the concatenated matrix used the uncorrelated lognormal (UCLN) relaxed clock (Drummond, Ho, Phillips, & Rambaut, 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, we chose the oldest fossil found to constrain the stem of each particular clade (Table 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, Suchard, Xie, & Drummond, 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.
### Table 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.

<table>
<thead>
<tr>
<th>Endemic genus</th>
<th>Endemic genus sampling</th>
<th>Suprageneric sampling</th>
<th>Suprageneric rank (number of species)</th>
<th>Family</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acidocroton</td>
<td>1/3</td>
<td>7/10</td>
<td>Tribe Crotoneae (16)</td>
<td>Euphorbiaceae</td>
</tr>
<tr>
<td>Acidoton</td>
<td>1/8</td>
<td>6/12</td>
<td>Tribe Pluknetiaceae (9)</td>
<td>Euphorbiaceae</td>
</tr>
<tr>
<td>Anacaona</td>
<td>1/1</td>
<td>12/13</td>
<td>Tribe Cucurbiteae (20)</td>
<td>Cucurbitaceae</td>
</tr>
<tr>
<td>Arcoa</td>
<td>1/1</td>
<td>46/56</td>
<td>Tribe Caesalpinieae (79)</td>
<td>Fabaceae</td>
</tr>
<tr>
<td>Bonania</td>
<td>1/8</td>
<td>20/23</td>
<td>Tribe Hippomaneae (30)</td>
<td>Euphorbiaceae</td>
</tr>
<tr>
<td>Broughtonia</td>
<td>4/6</td>
<td>50/54</td>
<td>Subtribe Laeliinae (117)</td>
<td>Orchidaceae</td>
</tr>
<tr>
<td>Brya</td>
<td>1/4</td>
<td>30/48</td>
<td>Tribe Dalbergieae (53)</td>
<td>Fabaceae</td>
</tr>
<tr>
<td>Calycogonium</td>
<td>1/36</td>
<td>3/19–23</td>
<td>Tribe Miconieae (13)</td>
<td>Melastomataceae</td>
</tr>
<tr>
<td>Chascotheca</td>
<td>1/2</td>
<td>3/3</td>
<td>Subtribe Astrocasinae (5)</td>
<td>Phyllanthaceae</td>
</tr>
<tr>
<td>Cubanola</td>
<td>1/2</td>
<td>9/28</td>
<td>Tribe Chiococceae (19)</td>
<td>Rubiaceae</td>
</tr>
<tr>
<td>Dendropemon</td>
<td>1/36</td>
<td>4/9</td>
<td>Subtribe Psittacanthineae (4)</td>
<td>Loranthaceae</td>
</tr>
<tr>
<td>Dilomilis</td>
<td>1/5</td>
<td>50/54</td>
<td>Subtribe Laeliinae (117)</td>
<td>Orchidaceae</td>
</tr>
<tr>
<td>Ditta</td>
<td>1/2</td>
<td>7/7</td>
<td>Tribe Adenoclineae (9)</td>
<td>Euphorbiaceae</td>
</tr>
<tr>
<td>Doerpfeldia</td>
<td>1/1</td>
<td>1/1</td>
<td>Tribe Doerpfeldieae (1)</td>
<td>Rhamnaceae</td>
</tr>
<tr>
<td>Domingoa</td>
<td>2/3</td>
<td>50/54</td>
<td>Subtribe Laeliinae (117)</td>
<td>Orchidaceae</td>
</tr>
<tr>
<td>Espadacea</td>
<td>1/1</td>
<td>2/4</td>
<td>Subfamily Goetzeoideae (2)</td>
<td>Solanaceae</td>
</tr>
<tr>
<td>Fuertesia</td>
<td>1/1</td>
<td>2/4</td>
<td>Subfamily Gronovioideae (2)</td>
<td>Lamiaceae</td>
</tr>
<tr>
<td>Goetzea</td>
<td>1/2</td>
<td>2/4</td>
<td>Subfamily Goetzeoideae (2)</td>
<td>Solanaceae</td>
</tr>
<tr>
<td>Grimmeodendron</td>
<td>1/2</td>
<td>20/23</td>
<td>Tribe Hippomaneae (30)</td>
<td>Euphorbiaceae</td>
</tr>
<tr>
<td>Haenianthus</td>
<td>1/2</td>
<td>15/18</td>
<td>Tribe Oleaeae (53)</td>
<td>Oleaceae</td>
</tr>
<tr>
<td>Hebestigma</td>
<td>1/1</td>
<td>10/13</td>
<td>Tribe Robinieae (23)</td>
<td>Fabaceae</td>
</tr>
<tr>
<td>Hemithrinax</td>
<td>3/3</td>
<td>9/10</td>
<td>Tribe Cryosophileae (14)</td>
<td>Areceae</td>
</tr>
<tr>
<td>Lasiocroton</td>
<td>3/5</td>
<td>5/6</td>
<td>Tribe Adelieae (18)</td>
<td>Euphorbiaceae</td>
</tr>
<tr>
<td>Leptocereus</td>
<td>1/12</td>
<td>14/27</td>
<td>Tribe Echinocereae (22)</td>
<td>Cactaceae</td>
</tr>
<tr>
<td>Leucocroton</td>
<td>3/28</td>
<td>5/6</td>
<td>Tribe Adelieae (18)</td>
<td>Euphorbiaceae</td>
</tr>
<tr>
<td>Microcycas</td>
<td>1/1</td>
<td>2/8</td>
<td>Family Zamiaceae (6)</td>
<td>Zamiaceae</td>
</tr>
<tr>
<td>Moacronot</td>
<td>1/8</td>
<td>7/10</td>
<td>Tribe Crotoneae (16)</td>
<td>Euphorbiaceae</td>
</tr>
<tr>
<td>Neobaceae</td>
<td>3/8</td>
<td>2/2</td>
<td>Subtribe Pachypodiinae (7)</td>
<td>Apocynaceae</td>
</tr>
<tr>
<td>Neocogniauxia</td>
<td>1/2</td>
<td>50/54</td>
<td>Subtribe Laeliinae (117)</td>
<td>Orchidaceae</td>
</tr>
<tr>
<td>Penelopeia</td>
<td>1/1</td>
<td>12/13</td>
<td>Tribe Cucurbiteae (20)</td>
<td>Cucurbitaceae</td>
</tr>
<tr>
<td>Petitia</td>
<td>1/2</td>
<td>2/7</td>
<td>Subfamily Viticloideae (2)</td>
<td>Lamiaceae</td>
</tr>
<tr>
<td>Picrodendron</td>
<td>1/1</td>
<td>13/19</td>
<td>Family Picrodendraceae (13)</td>
<td>Picrodendraceae</td>
</tr>
<tr>
<td>Pictetia</td>
<td>1/8</td>
<td>30/48</td>
<td>Tribe Dalbergieae (53)</td>
<td>Fabaceae</td>
</tr>
<tr>
<td>Poitea</td>
<td>3/12</td>
<td>10/13</td>
<td>Tribe Robinieae (23)</td>
<td>Fabaceae</td>
</tr>
<tr>
<td>Psychilis</td>
<td>2/15</td>
<td>50/54</td>
<td>Subtribe Laeliinae (117)</td>
<td>Orchidaceae</td>
</tr>
<tr>
<td>Quisqueya</td>
<td>1/4</td>
<td>50/54</td>
<td>Subtribe Laeliinae (117)</td>
<td>Orchidaceae</td>
</tr>
<tr>
<td>Rhodopis</td>
<td>1/2</td>
<td>47/84</td>
<td>Tribe Phaseoleae (86)</td>
<td>Fabaceae</td>
</tr>
<tr>
<td>Stahlia</td>
<td>1/1</td>
<td>46/56</td>
<td>Tribe Caesalpinieae (79)</td>
<td>Fabaceae</td>
</tr>
<tr>
<td>Synapsis</td>
<td>1/1</td>
<td>2/3</td>
<td>Family Schlegeliaceae (2)</td>
<td>Schlegeliaceae</td>
</tr>
<tr>
<td>Tetramicra</td>
<td>1/13</td>
<td>50/54</td>
<td>Subtribe Laeliinae (117)</td>
<td>Orchidaceae</td>
</tr>
<tr>
<td>Zombia</td>
<td>1/1</td>
<td>9/10</td>
<td>Tribe Cryosophileae (14)</td>
<td>Areceae</td>
</tr>
</tbody>
</table>
**Table 2** Calibration points used for divergence time estimation in BEAST. The offset values from the BEAUTI settings column correspond to assigned fossil ages.

<table>
<thead>
<tr>
<th>Fossil name</th>
<th>Clade constrained</th>
<th>Plant organs and synapomorphies</th>
<th>Primary reference</th>
<th>BEAUTI settings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Machaerium</td>
<td>Stem of Tribe Dalbergieae (Leguminosae)</td>
<td>Fossil leaflets. Strong marginal vein, poorly organized higher order venation, numerous closely spaced craspedodromous secondary veins, and epidermal cell structure are diagnostic characters for <em>Machaerium</em> (Tribe Dalbergieae)</td>
<td>(Herendeen, Crepet, &amp; Dilcher, 1992)</td>
<td>Offset = 40, Mean = 1.0, SD = 0.5</td>
</tr>
<tr>
<td>Fraxinus excelsior</td>
<td>Crown of family Oleaceae</td>
<td>Fruit fossils. Winged (samara type) fruit that resembles <em>Fraxinus</em> in peduncle, vein structure and shape, and position of seed</td>
<td>(Jung &amp; Lee, 2009)</td>
<td>Offset = 5.33, Mean = 1.0, SD = 0.5 (as used in Magallón et al., 2015)</td>
</tr>
<tr>
<td>Sabalites carolinensis</td>
<td>Stem of Tribe Cryosophileae (Subfamily Coryphoideae, Arecaceae)</td>
<td>Leaf fossil. Oldest known palm fossil assignable to Subfamily Coryphoideae with costapalmate leaf</td>
<td>(Dransfield et al., 2008)</td>
<td>Offset = 86.7, Mean = 1.7, SD = 0.3 (as used in Bacon et al., 2012)</td>
</tr>
<tr>
<td>Micrantheum spinyspora</td>
<td>Stem of family Picrodendraceae</td>
<td>Pollen fossils</td>
<td>(Christophel, Harris, &amp; Syber, 1987)</td>
<td>Offset = 35.55, Mean = 1.0, SD = 0.5</td>
</tr>
<tr>
<td>Acalypha</td>
<td>Stem of tribes Adelieae and Pluketenieae (Subfamily Acalyphoideae, Euphorbiaceae)</td>
<td>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</td>
<td>(Sun et al., 1989)</td>
<td>Offset = 61.0, Mean = 1.0, SD = 0.5 (as used in Davis, Webb, Wurdack, Jaramillo, &amp; Donoghue, 2005)</td>
</tr>
<tr>
<td>Solanispernum reniforme</td>
<td>Stem of family Solanaceae</td>
<td>Fossil seeds; one of the earliest fossils assigned to Solanaceae</td>
<td>(Chandler, 1962)</td>
<td>Offset = 47.0, Mean = 1.0, SD of 0.5 (as used in Martínez-Millán, 2010)</td>
</tr>
<tr>
<td>Trithrinax dominicana,</td>
<td>Stem of genus Trithrinax (Arecaceae)</td>
<td>Flower fossils. Stamen filaments exerted and tips bent inwards are diagnostic characters for Trithrinax</td>
<td>(Poinar, 2002)</td>
<td>Offset = 24.5, Mean = 1.0, SD of 0.5</td>
</tr>
<tr>
<td>Prosopis linearifolia</td>
<td>Stem of Umtiza clade (Fabaceae)</td>
<td>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 with Arcoa (<em>Umtiza</em> clade)</td>
<td>(Herendeen, Lewis, &amp; Bruneau, 2003)</td>
<td>Offset = 34.0, Mean = 1.0, SD of 0.5 (as used in Lavin, Herendeen, &amp; Wojciechowski, 2005)</td>
</tr>
<tr>
<td></td>
<td>Stem of Angiosperms</td>
<td>Secondary calibration point</td>
<td>(Silvestro et al., 2015)</td>
<td>Laplace prior distribution, Offset = 143.7, $\mu = 1.0$, scale = 4.36</td>
</tr>
<tr>
<td></td>
<td>Stem of Spermatophyta</td>
<td>Secondary calibration point</td>
<td>(Silvestro et al., 2015)</td>
<td>Gamma prior distribution, Offset = 366.0, Mean = 1.0 SD = 0.5</td>
</tr>
</tbody>
</table>
2.4 | Ancestral area estimation

We 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, Moore, Webb, & Donoghue, 2005; Ree & Smith, 2008), dispersal–vicariance (DIVA, Ronquist, 1997), and Bayesian biogeographic inference (BayArea, Landis, Matzke, Moore, & Huelsenbeck, 2013). BioGeoBears requires an ultrametric tree (we used the MCC tree from BEAST) and a matrix of geographic distributions in presence–absence format. As BioGeoBears requires positive branch lengths (Matzke, 2013), we manually edited the only negative branch length by adding 0.3 nucleotide substitution per site. To prepare the presence–absence matrix, we obtained species distributions from the Global Biodiversity Information Facility (GBIF, https://www.gbif.org/, accessed 12 June, 2015). We defined five biogeographic operational areas: (A) Antilles; (B) Central America; (C) South America; (D) North America; and (E) rest of the world (Figure 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.

We 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, DECj; BAYEAREALIKE; BAYAREALIKEj; DIVALIKE; DIVALIKEj). 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). We 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 (Bacon et al., 2015; Weir et al., 2009), a third complex stratified model was run using unequal dispersal probabilities for the period 0–15 Ma. We 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 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 (Jaramillo et al., 2017; Montes et al., 2015) facilitating biotic interchange between North and South America as shown for wide range of taxonomic groups in Bacon et al. (2015). We 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, Antonelli, & Donoghue, 2016). During this period, the Antilles were already above water; therefore, dispersal from/to the Antilles and the surrounding landmasses was possible; we 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, Oleinik, Marincovich, & Barinov, 2002). The North Atlantic land bridge connecting Europe to North America was hypothesized...
TABLE 3 Dispersal matrices used in BioGeoBears for complex biogeographic modeling

| Complex model 1 and 2 | AN | CA | NA | RW | SA | AN | CA | NA | RW | SA | AN | CA | NA | RW | SA |
|-----------------------|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| 0–15 Ma               | 1  | 0.5| 0.5| 0.1| 0.5| 1  | 0.5| 0.1| 0.5| 0.1| 1  | 0.5| 0.5| 0.1| 1  |
| 15–33 Ma              | 1  | 0.5| 0.5| 0.1| 0.5| 1  | 0.5| 0.1| 0.5| 0.1| 1  | 0.5| 0.5| 0.1| 1  |
| 1                    | 0.5| 1  | 1  | 0.5| 1  | 1  | 0.5| 1  | 0.5| 1  | 1  | 0.5| 1  | 0.5| 1  |
| 0.5                  | 1  | 0.5| 0.5| 0.1| 0.5| 1  | 0.5| 1  | 0.5| 1  | 1  | 0.5| 1  | 0.5| 1  |
| 0.1                  | 0.5| 1  | 1  | 0.1| 0.1| 0.1| 0.1| 0.1| 0.1| 0.1| 0.1| 0.1| 0.1| 0.1| 0.1|
| 0.5 (0.75)           | 1  | 0.5(0.75)| 0.1| 1  | 0.5| 0.1| 0.1| 0.1| 1  | 0.1| 0.1| 0.1| 0.1| 1  |
| 35–50 Ma             | 1  | 0.5| 0.5| 0.01| 0.5| 1  | 0.01| 0.1| 0.01| 0.1| 1  | 0  | 0  | 0  | 0  |
| 50–130 Ma            | 0.5| 1  | 0.1| 0.01| 0.1| 0.01| 0.1| 0.01| 0.01| 0.01| 0  | 1  | 0.01| 0.01| 0.01|
| 130–378 Ma           | 0.5| 0.1| 0.1| 0.01| 0.1| 0.01| 0.1| 0.01| 0.01| 0.01| 0  | 1  | 0.01| 0.01| 0.01|
| 1                    | 0.5| 0.1| 0.1| 0.01| 0.1| 0.01| 0.1| 0.01| 0.01| 0.01| 1  | 0  | 0.01| 0.01| 0.01|
| 0.5                  | 0.1| 0.1| 0.1| 0.01| 0.1| 0.01| 0.1| 0.01| 0.01| 0.01| 1  | 0  | 0.01| 0.01| 0.01|

Complex model number 1 accounts for equal dispersal probabilities in both directions between areas. 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 (values shown within parentheses), thus favoring dispersal South to North for the 0–15 Ma period.

5. 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 (Poole & Vorren, 1993; Schnitker, 1980). Therefore, we 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: We 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, we 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. We 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, we set a probability of 0.1 for potential dispersal events from/to Central and North America, and Central and South America. We 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, we 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. We 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, we imposed a minimal dispersal probability of 0.1 from/to Antilles to North and South America;

6. 130–378 Ma: From the Middle Devonian to Late Jurassic, land-masses 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. We 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, Soltis, & Soltis, 2010; Silvestro, Cascales-Miñana, Bacon, & Antonelli, 2015), and it is prior to the origin of Zamiaceae (Salas-Leiva et al., 2013). We did not allow any dispersal event from/to the Antilles as those islands had not formed yet.

2.5 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, Hardy, & Rutschmann, 2005; Pirie & Doyle, 2012). We therefore conducted a second approach to contrast and validate our broad-scale results using multiple independently dated phylogenetic trees. We compiled crown and stem ages from published trees that comprised the Caribbean endemic genera included in the
broad-scale analysis. We obtained 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 4).

## 3 | RESULTS

### 3.1 | Phylogenetic reconstruction and molecular dating

The ML and BI analyses recovered congruent tree topologies for the higher relationships of taxa. Figure 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 (2009). In one exception, the BI, but not the ML analysis, recovered a clade of 15 Euphorbiaceae species within the Orchidaceae clade, which we therefore removed from the tree using the "drop.tip ()" function in the R package "ape" (Paradis, Claude, & Strimmer, 2004).

Consequently, the endemic genera Moacroton and Acidocroton were removed from the dated tree and will not be further discussed. In addition, due to our 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 1). Therefore, divergence times and ancestral areas for the endemic genera Doerpfelia, Espadacea, 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 we 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: 

<table>
<thead>
<tr>
<th>Endemic genus</th>
<th>Crown age in Ma</th>
<th>Stem age in Ma</th>
<th>Ancestral area estimation</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acidoton</td>
<td>1.7</td>
<td>2.6</td>
<td>SA</td>
<td>(Cervantes et al., 2016)</td>
</tr>
<tr>
<td>Anacaona</td>
<td>13</td>
<td>17</td>
<td>SA</td>
<td>(Schaefer et al., 2009)</td>
</tr>
<tr>
<td>Arcoa</td>
<td>34</td>
<td></td>
<td>SA</td>
<td>(Lavin et al., 2005)</td>
</tr>
<tr>
<td>Bononia</td>
<td>41.6</td>
<td>46.39</td>
<td>MX, SA, MS</td>
<td>(Cervantes et al., 2016)</td>
</tr>
<tr>
<td>Broughtonia</td>
<td>15.68</td>
<td>20.74</td>
<td></td>
<td>(Sosa et al., 2016)</td>
</tr>
<tr>
<td>Brya</td>
<td>41.9</td>
<td>47.2</td>
<td></td>
<td>(Lavin et al., 2005)</td>
</tr>
<tr>
<td>Cubanola</td>
<td>19.2</td>
<td>47.3</td>
<td>AN</td>
<td>(Antonelli et al., 2009)</td>
</tr>
<tr>
<td>Dilomilis</td>
<td>16.01</td>
<td>46.72</td>
<td></td>
<td>(Sosa et al., 2016)</td>
</tr>
<tr>
<td>Ditta</td>
<td>95</td>
<td>105</td>
<td></td>
<td>(van Ee et al., 2008)</td>
</tr>
<tr>
<td>Dominoa</td>
<td>19.34</td>
<td>20.9</td>
<td></td>
<td>(Sosa et al., 2016)</td>
</tr>
<tr>
<td>Hebestigma</td>
<td>38.1</td>
<td>48.3</td>
<td></td>
<td>(Lavin et al., 2003)</td>
</tr>
<tr>
<td>Hemithrinax</td>
<td>6.99</td>
<td>17.54</td>
<td>AN</td>
<td>(Cano et al., unpublished data)</td>
</tr>
<tr>
<td>Leptocereus</td>
<td>2.8</td>
<td>4.8</td>
<td>SA</td>
<td>(Hernández-Hernández et al., 2014)</td>
</tr>
<tr>
<td>Lasiocroton</td>
<td>1.17</td>
<td>10.76</td>
<td>AN</td>
<td>(Cervantes et al., 2016)</td>
</tr>
<tr>
<td>Leucocroton</td>
<td>5.27</td>
<td>10.76</td>
<td>AN</td>
<td>(Cervantes et al., 2016)</td>
</tr>
<tr>
<td>Microcycas</td>
<td>36.5</td>
<td>60.32</td>
<td>AF-CA</td>
<td>(Salas-Leiva et al., 2013)</td>
</tr>
<tr>
<td>Neocogniauxia</td>
<td>16.01</td>
<td>46.72</td>
<td></td>
<td>(Sosa et al., 2016)</td>
</tr>
<tr>
<td>Penelopoeia</td>
<td>13</td>
<td>17</td>
<td>SA</td>
<td>(Schaefer et al., 2009)</td>
</tr>
<tr>
<td>Pictetia</td>
<td>14.5</td>
<td>45.6</td>
<td></td>
<td>(Lavin et al., 2005)</td>
</tr>
<tr>
<td>Poitea</td>
<td>9.2</td>
<td>16.4</td>
<td></td>
<td>(Lavin, Wojciechowski, et al., 2001)</td>
</tr>
<tr>
<td>Psychilis</td>
<td>15.68</td>
<td>20.74</td>
<td></td>
<td>(Sosa et al., 2016)</td>
</tr>
<tr>
<td>Quisqueya</td>
<td>15.68</td>
<td>20.74</td>
<td></td>
<td>(Sosa et al., 2016)</td>
</tr>
<tr>
<td>Tetramicra</td>
<td>15.68</td>
<td>20.74</td>
<td></td>
<td>(Sosa et al., 2016)</td>
</tr>
<tr>
<td>Zombia</td>
<td>3.75</td>
<td>21.7</td>
<td>AN</td>
<td>(Cano et al., unpublished data)</td>
</tr>
</tbody>
</table>

Ancestral area abbreviations: SA, South American ancestor; AN, Antillean ancestor; AF-CA, African Caribbean ancestor; MX, Mexico; and MS, Mesoamerica.
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 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 (Acidoten [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 5 (see also Figs S1 and S2 for BEAST MCC tree and node numbers in MCC tree). The four-gene concatenated matrix, and BEAST MCC tree are available in Dryad (http://dx.doi.org/10.5061/dryad.gq93s).

### 3.2 Ancestral area estimation

The likelihood values for the null and the two complex stratified models can be found in Table 6 for the 18 models ran in BioGeoBears. Model selection did not support the hypothesizeddirectionality 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 our data set (\(\text{Lnl} = -1221.2\) and \(\text{AIC} = 2,448.3\)), while the second-best model was the DEC\(j\) from the complex model 2 (\(\text{Lnl} = -1235.1\) and \(\text{AIC} = 2,476.3\)).

The complete ancestral area estimation using the DEC\(j\) model (complex model 1) is shown in Fig. S3. To plot the most likely ancestral distribution of endemic genera, we selected their corresponding stem nodes. When an endemic genus was sister to another endemic...
genus, we 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. We considered Bonania–Grimmeodendron also as an endemic clade, because the sister relationship of Grimmeodendron e glandulosum and Sebastiana 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).

Our results show that nine endemic genera or clades and their sister groups had ancestors distributed in the Antilles (i.e., highest

### TABLE 5

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

<table>
<thead>
<tr>
<th>Endemic genus</th>
<th>Crown node number</th>
<th>Mean ages at crown nodes in Ma (95% HPD)</th>
<th>Stem node number</th>
<th>Mean ages at stem nodes in Ma (95% HPD)</th>
<th>Ancestral Reconstruction probabilities (at stem nodes)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acidoton</td>
<td>880</td>
<td>31.74 (13.65–50.99)</td>
<td>879</td>
<td>42.49 (23.65–60.72)</td>
<td>RW 0.38</td>
</tr>
<tr>
<td>Anacaona</td>
<td>628</td>
<td>12.71 (4.05–22.29)</td>
<td>627</td>
<td>27.67 (17.54–36.97)</td>
<td>AN 0.65</td>
</tr>
<tr>
<td>Arcoa</td>
<td>640</td>
<td>34.09 (26.56–40.05)</td>
<td>637</td>
<td>38.24 (34.68–41.72)</td>
<td>RW 0.80</td>
</tr>
<tr>
<td>Pouteria</td>
<td>926</td>
<td>9.92 (2.81–17.60)</td>
<td>925</td>
<td>19.41 (8.32–30.44)</td>
<td>AN 0.19; ANNA 0.16*</td>
</tr>
<tr>
<td>Broughtonia</td>
<td>1,136</td>
<td>19.42 (10.91–27.63)</td>
<td>1,135</td>
<td>26.73 (18.45–36.05)</td>
<td>AN 0.83*</td>
</tr>
<tr>
<td>Brya</td>
<td>851</td>
<td>13.72 (3.52–24.82)</td>
<td>850</td>
<td>44.17 (28.59–57.89)</td>
<td>SA 0.94</td>
</tr>
<tr>
<td>Calycogonium</td>
<td>957</td>
<td>26.81 (9.67–44.44)</td>
<td>956</td>
<td>47.27 (26.64–67.79)</td>
<td>SA 0.53</td>
</tr>
<tr>
<td>Chascotheca</td>
<td>952</td>
<td>32.86 (15.37–52.74)</td>
<td>951</td>
<td>56.43 (31.72–80.80)</td>
<td>RWO 0.26; ANRW 0.22</td>
</tr>
<tr>
<td>Cubanola</td>
<td>1,004</td>
<td>17.65 (9.73–26.12)</td>
<td>1,002</td>
<td>20.25 (11.98–28.28)</td>
<td>AN 0.97</td>
</tr>
<tr>
<td>Dendropemon</td>
<td>1,083</td>
<td>20.45 (5.34–37.22)</td>
<td>1,082</td>
<td>44.19 (20.76–70.21)</td>
<td>CA 0.65</td>
</tr>
<tr>
<td>Dilomilis</td>
<td>1,199</td>
<td>18.74 (5.10–35.04)</td>
<td>1,111</td>
<td>52.80 (36.69–69.27)</td>
<td>AN 0.93*</td>
</tr>
<tr>
<td>Ditta</td>
<td>937</td>
<td>15.09 (5.32–25.90)</td>
<td>936</td>
<td>28.50 (13.42–43.68)</td>
<td>RW 0.80</td>
</tr>
<tr>
<td>Domingoa</td>
<td>1,144</td>
<td>15.009 (5.61–24.63)</td>
<td>1,116</td>
<td>30.67 (24.48–34.96)</td>
<td>AN 0.57*</td>
</tr>
<tr>
<td>Grimmeodendron</td>
<td>927</td>
<td>6.61 (0.99–12.81)</td>
<td>925</td>
<td>19.41 (8.32–30.44)</td>
<td>AN 0.19; ANNA 0.16*</td>
</tr>
<tr>
<td>Hebestigma</td>
<td>806</td>
<td>53.12 (33.13–72.98)</td>
<td>716</td>
<td>105.48 (88.57–123.62)</td>
<td>SA 0.29</td>
</tr>
<tr>
<td>Hemithrinax</td>
<td>1,210</td>
<td>6.07 (1.71–10.28)</td>
<td>1,209</td>
<td>9.35 (3.68–14.72)</td>
<td>AN 0.98</td>
</tr>
<tr>
<td>Leptocereus</td>
<td>979</td>
<td>9.07 (2.43–16.32)</td>
<td>978</td>
<td>11.12 (3.35–17.42)</td>
<td>SA 0.89</td>
</tr>
<tr>
<td>Lasiocroton</td>
<td>891</td>
<td>22.59 (11.16–35.23)</td>
<td>890</td>
<td>28.85 (18.81–41.40)</td>
<td>CA 0.54*</td>
</tr>
<tr>
<td>Leucocroton</td>
<td>891</td>
<td>22.59 (11.16–35.23)</td>
<td>890</td>
<td>28.85 (18.81–41.40)</td>
<td>CA 0.54*</td>
</tr>
<tr>
<td>Microcycas</td>
<td>1,216</td>
<td>28.23 (12.43–45.78)</td>
<td>1,215</td>
<td>51 (26.15–79.07)</td>
<td>AN 0.95</td>
</tr>
<tr>
<td>Neobracea</td>
<td>1,015</td>
<td>6.49 (0.81–13.42)</td>
<td>1,014</td>
<td>35.45 (18.29–53.39)</td>
<td>ANRW 0.61</td>
</tr>
<tr>
<td>Neocogniauxia</td>
<td>1,119</td>
<td>18.74 (5.10–35.04)</td>
<td>1,111</td>
<td>52.80 (36.69–69.27)</td>
<td>AN 0.93*</td>
</tr>
<tr>
<td>Penelopeia</td>
<td>627</td>
<td>20.91 (9.58–31.80)</td>
<td>626</td>
<td>27.67 (17.54–36.97)</td>
<td>CA 0.39</td>
</tr>
<tr>
<td>Picrodendron</td>
<td>942</td>
<td>19.13 (12.37–26.11)</td>
<td>941</td>
<td>23.93 (16.27–30.93)</td>
<td>RW 0.48</td>
</tr>
<tr>
<td>Pictetia</td>
<td>842</td>
<td>13.96 (5.15–23.03)</td>
<td>841</td>
<td>17.67 (7.47–27.37)</td>
<td>AN 0.45</td>
</tr>
<tr>
<td>Poitea</td>
<td>826</td>
<td>11.50 (4.63–18.89)</td>
<td>823</td>
<td>17.10 (8.50–26.01)</td>
<td>AN 0.25; ANCA 0.21</td>
</tr>
<tr>
<td>Psychilis</td>
<td>1,140</td>
<td>15.93 (8.51–24.31)</td>
<td>1,135</td>
<td>26.73 (18.45–36.05)</td>
<td>AN 0.83*</td>
</tr>
<tr>
<td>Quisqueya</td>
<td>1,142</td>
<td>12.06 (4.73–19.92)</td>
<td>1,135</td>
<td>26.73 (18.45–36.05)</td>
<td>AN 0.83*</td>
</tr>
<tr>
<td>Rhodopsis</td>
<td>797</td>
<td>17.47 (6.42–29.48)</td>
<td>796</td>
<td>21.89 (10.33–33.92)</td>
<td>SA 0.14</td>
</tr>
<tr>
<td>Stahlia</td>
<td>666</td>
<td>3.40 (0.0078–8.50)</td>
<td>665</td>
<td>8.64 (1.86–15.97)</td>
<td>CA 0.67</td>
</tr>
<tr>
<td>Tetramicra</td>
<td>1,142</td>
<td>12.06 (4.73–19.92)</td>
<td>1,135</td>
<td>26.73 (18.45–36.05)</td>
<td>AN 0.83*</td>
</tr>
<tr>
<td>Zambja</td>
<td>1,209</td>
<td>8.13 (3.03–13.74)</td>
<td>1,208</td>
<td>9.35 (3.68–14.72)</td>
<td>AN 0.79</td>
</tr>
</tbody>
</table>

Asterisks indicate that the genus is part of an endemic clade (*). SA, South America; CA, Central America; NA, North America; AN, Antilles; RW, rest of the world.
of uncertainty in the analysis (see Table 5 for ancestral reconstruction probabilities for each genus).

Our analysis also recovered ten instances in which continental taxa appeared nested within a Caribbean clade, suggesting potential island to mainland colonization. More specifically, we 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)].

### 3.3 A compilation of independent evolutionary and biogeographic studies on a subset of Caribbean endemic genera

We found 10 studies that reported crown and stem ages for 24 endemic genera (Table 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 our broad-scale analysis, the rest (18) had crown and/or stem ages inside our 95% HPD interval, and thus, we consider them in agreement with our 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, Penelopea, 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 4).

### 4 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, we 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.

#### 4.1 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 we found in our study. The crown age we recovered for the Angiosperms was also concordant with earlier studies (Bell et al., 2010, 95% HPD 167–199 Ma; Smith, Beaulieu, & Donoghue, 2010, 95% HPD 182–257 Ma; Magallón, Hilu, & Quandt, 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, Gómez-Acevedo, Sánchez-Reyes, and Hernández- (2015, 95% HPD 136–139.95 Ma).

Our 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
TABLE 6 Biogeographic model testing in BioGeoBears

<table>
<thead>
<tr>
<th>LnL</th>
<th># params</th>
<th>d</th>
<th>e</th>
<th>j</th>
<th>AIC</th>
<th>AIC wt</th>
</tr>
</thead>
<tbody>
<tr>
<td>Null model</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BAYAREALIKE</td>
<td>-1,480.99</td>
<td>2</td>
<td>0.00389</td>
<td>0.019351951</td>
<td>0</td>
<td>2,965.992259</td>
</tr>
<tr>
<td>BAYAREALIKEj</td>
<td>-1,260.30</td>
<td>3</td>
<td>0.00228</td>
<td>0.000617994</td>
<td>0.046701184</td>
<td>0</td>
</tr>
<tr>
<td>DEC</td>
<td>-1,318.92</td>
<td>2</td>
<td>0.00449</td>
<td>1.00E-12</td>
<td>0</td>
<td>2,641.850873</td>
</tr>
<tr>
<td>DECj</td>
<td>-1,262.94</td>
<td>3</td>
<td>0.00356</td>
<td>1.00E-12</td>
<td>0.028423902</td>
<td>0</td>
</tr>
<tr>
<td>DIVALIKE</td>
<td>-1,330.90</td>
<td>2</td>
<td>0.00541</td>
<td>1.00E-12</td>
<td>0</td>
<td>2,665.809754</td>
</tr>
<tr>
<td>DIVALIKEj</td>
<td>-1,294.20</td>
<td>3</td>
<td>0.00428</td>
<td>1.08E-09</td>
<td>0.022922813</td>
<td>0</td>
</tr>
<tr>
<td>Complex model 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BAYAREALIKE</td>
<td>-1,465.10</td>
<td>2</td>
<td>0.01131</td>
<td>0.016440112</td>
<td>0</td>
<td>2,934.215946</td>
</tr>
<tr>
<td>BAYAREALIKEj</td>
<td>-1,235.58</td>
<td>3</td>
<td>0.00847</td>
<td>0.000807486</td>
<td>0.126630982</td>
<td>0</td>
</tr>
<tr>
<td>DEC</td>
<td>-1,288.10</td>
<td>2</td>
<td>0.01513</td>
<td>0.000614288</td>
<td>0</td>
<td>2,580.218724</td>
</tr>
<tr>
<td>DECj</td>
<td>-1,221.16</td>
<td>3</td>
<td>0.01144</td>
<td>4.87E-05</td>
<td>0.115873802</td>
<td>0</td>
</tr>
<tr>
<td>DIVALIKE</td>
<td>-1,295.18</td>
<td>2</td>
<td>0.018188</td>
<td>0.000809571</td>
<td>0</td>
<td>2,594.36264</td>
</tr>
<tr>
<td>DIVALIKEj</td>
<td>-1,272.81</td>
<td>3</td>
<td>0.014687</td>
<td>0.000472286</td>
<td>0.020573461</td>
<td>0</td>
</tr>
<tr>
<td>Complex model 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BAYAREALIKE</td>
<td>-1,468.83</td>
<td>2</td>
<td>0.01044</td>
<td>0.016429327</td>
<td>0</td>
<td>2,941.666889</td>
</tr>
<tr>
<td>BAYAREALIKEj</td>
<td>-1,243.37</td>
<td>3</td>
<td>0.00805</td>
<td>0.000673963</td>
<td>0.122983955</td>
<td>0</td>
</tr>
<tr>
<td>DEC</td>
<td>-1,296.90</td>
<td>2</td>
<td>0.01430</td>
<td>0.00060304</td>
<td>0</td>
<td>2,597.809953</td>
</tr>
<tr>
<td>DECj</td>
<td>-1,235.13</td>
<td>3</td>
<td>0.01109</td>
<td>5.33E-05</td>
<td>0.063058035</td>
<td>0</td>
</tr>
<tr>
<td>DIVALIKE</td>
<td>-1,304.63</td>
<td>2</td>
<td>0.01719</td>
<td>0.000805488</td>
<td>0</td>
<td>2,613.262891</td>
</tr>
<tr>
<td>DIVALIKEj</td>
<td>-1,281.03</td>
<td>3</td>
<td>0.01505</td>
<td>0.000812842</td>
<td>0.019658281</td>
<td>0</td>
</tr>
</tbody>
</table>

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 likelihood for each model. Best and second-best model according to AIC values marked as “a” and “b,” respectively.

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). Our 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 our age confidence intervals (Figure 3). Divergence times for Acidotot, Leucocroton, and Lasiocroton estimated here were older than those reported in Cervantes, Fuentes, Gutiérrez, Magallón, and Borsch (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, our 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.

We 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, we discuss below their relevance in light of the range evolution analysis.

4.2 | 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 our results support this. About 32% of the endemic genera had ancestors distributed in continental America. Our 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).
Our 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 (Cervantes et al., 2016; Renner, 2004), 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). Our results of a Central American ancestor disagree with those reported in the literature in two cases. The first is the endemic clade Lasiocron–Leucocron (Euphorbiaceae), for which Jestrow, Gutiérrez, and Francisco-Ortega (2012) and Cervantes et al. (2016) suggested an ancestor in eastern Cuba and the Antilles, respectively. The second is Penelopeia (Cucurbitaceae), for which Schaefer, Heibl, and Renner (2009) proposed a South American origin for the subfamily Cucurbitoideae. We 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), we only found five endemic genera with South American ancestors. Of these five, the mean stem ages of Leptocereus and Rhodopsis (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 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 our SA ancestral area result (Hernández-Hernández, Brown, Schlumpberger, Eguiarte, & Magallón, 2014).

Our analyses revealed Old World ancestors for five endemic genera. For example, Acidoton (Euphorbiaceae) formed a clade with two North American species of Tragio, 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, Zhang, & Renner, 2010; Wei et al., 2015). However, Cervantes et al. (2016) proposed a South American ancestor for Acidoton, which we also attribute to their different taxon sampling. We found no formal biogeographic analysis for any of the other four genera in the literature.

Our 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 our analyses might explain the disagreement between the Antillean ancestor we 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 (Hermansen, Taylor, Taylor, & Stevenson, 2006). Antillean-distributed ancestors were recovered for all endemic orchids (seven genera within three clades). For the genus *Banonia* (Euphorbiaceae), our results showed an Antillean and North American distributed ancestor, whereas Cervantes et al. (2016), who did not include the endemic sister genus *Grimmeodeon*, recovered an ancestor distributed in Mesoamerica, SA, and the Caribbean. Our finding of Antillean-distributed ancestors for the endemic legume genera *Pictetia* and *Poitea* was not in agreement with that of Lavin, Wojciechowski, et al. (2001) who using a cladistic vicariance analysis hypothesized on a boreotropical origin for these two endemic genera. We attribute this disagreement to differences in taxon sampling and biogeographic analysis method. Our analysis recovered the Antilles as the most likely ancestral range for the two palm genera *Hemithrinax* and *Zonbola* and for *Cubanola* (Rubiaceae) corroborating the results of Cano et al. (unpublished data), and Antonelli, Nylander, Persson, and Sanmartin (2009), respectively.

Our phylogenetic framework also identified nine instances for which Antillean taxa acted as source for continental taxa. This result is in line with Bellemain and 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, Trusty, Santos-Guerra, Jansen, & Francisco-Ortega, 2004; Carine, Russell, Santos-Guerra, & Francisco-Ortega, 2004; Condamine, Leslie, & Antonelli, 2016; Patiño et al., 2015). Our 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 overwater, bird, or wind dispersal the most plausible explanation.

5 | CONCLUSIONS

This is, to our knowledge, the first attempt to synthesize the evolutionary history of unsampled endemic plant lineages from the Caribbean. The high percentage of unsampled endemic genera in molecular phylogenies illustrates the knowledge gap of systematic and biogeographic studies on the Caribbean flora, and efforts should focus on sequencing more endemic taxa and their continental relatives. This would allow to further test our results of a mixture of primarily recent (Oligocene–Miocene) but also old (Paleocene–Eocene) lineages giving origin to the extant Caribbean endemic flora, and of predominantly Antillean ancestors of endemic genera. In addition, little is known about the reproductive biology of most seed plants endemic to the Caribbean region. This limited amount of natural history studies prevents us from generalizing on the dispersal patterns or mechanisms that ancestral lineages used to colonize and diversify in the Caribbean.

Our research does not support a role for GAARlandia as a colonization route to the Antilles for plants, and further 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.

ACKNOWLEDGMENTS

We thank Westgrid’s (Compute Canada) for the computational resources and Dr. Oliver Stueker from ACENET (Compute Canada) at MUN for his assistance during the use of the computer cluster for the BEAST analysis. We are also grateful to the editor and an anonymous reviewer whose comments and suggestions greatly improved this manuscript. This work was funded by an NSERC-Discovery grant (RGPIN-2014-03976) to J.R. and MUN’s President’s Doctoral Student Investment Fund to M.E.N.B. A.A. is funded by the Swedish Research Council (B0539601), the European Research Council under the European Union’s Seventh Framework Programme (FP/2007-2013, ERC Grant Agreement n. 331024), the Swedish Foundation for Strategic Research, and a Wallenberg Academy Fellowship.

CONFLICT OF INTEREST

None declared.

AUTHOR CONTRIBUTIONS

M.E.N.B., J.R., and A.A. designed the study. M.E.N.B. compiled and analyzed the data. M.E.N.B. wrote the article with contributions from J.R. and A.A.

ORCID

Maria Esther Nieto-Blázquez http://orcid.org/0000-0002-6201-282X

REFERENCES


