

HERBARIUM-BASED PHYLOGENOMICS SHEDS LIGHT ON THE EVOLUTIONARY HISTORY OF MESOAMERICAN TROPICAL FORESTS: ANNONACEAE AS A CASE STUDY

ANDRÉS ERNESTO ORTIZ-RODRIGUEZ^{1*}, LAURA HOLZMEYER², LEOPOLDO HURTADO-REVELES¹, FRANCIS J. NGE³,
 YUYINI LICONA-VERA⁴, THOMAS L. P. COUVREUR²

¹ Departamento de Botánica, Instituto de Biología, Universidad Nacional Autónoma de México, Mexico City, Mexico.

² Plant Diversity, Adaptation and Development, Université de Montpellier, Montpellier, France.

³ National Herbarium of New South Wales, Botanic Gardens of Sydney, Sydney, Australia.

⁴ Departamento de Ecología Evolutiva, Instituto de Ecología, Universidad Nacional Autónoma de México, Mexico City, Mexico.

*Corresponding author: andres.ortiz@ib.unam.mx

Abstract

The tropical forests of Mesoamerica, stretching from Mexico to northern Colombia, represent one of the most diverse biotas on Earth. Yet their evolutionary history and the processes that shape their assembly remain poorly understood, often leading to their treatment as a single biotic unit. In this study, we use the family Annonaceae as a model lineage to reconstruct the region's evolutionary trajectories, integrating hundreds of nuclear markers derived from both historical and contemporary herbarium specimens. This high-resolution phylogenomic framework, encompassing more than 80 % of the ~2,500 species in the family, reveals consistent patterns of floristic and evolutionary differentiation, as well as sharp contrasts between northern Mesoamerica (MegaMexico 2 and the Caribbean) and the southern regions. Northern areas exhibit deeper phylogenetic diversity despite lower species richness, reflecting prolonged isolation and greater environmental heterogeneity. In contrast, the south shows high species richness but phylogenetic redundancy, driven by recent colonizations and climatic filters. Divergence times highlight the central role of Miocene-Pliocene climatic oscillations and more recent orogenic activity in shaping Mesoamerican forests: dry forests emerge as evolutionary refugia and center of radiation. In contrast, wet forests consistently harbor younger lineages. Taken together, these findings challenge the view of Mesoamerica as a single evolutionary unit and demonstrate that its tropical forests are dynamic systems, repeatedly molded by cycles of climatic change and geographic isolation.

Keywords: endemism, phylogenetic diversity, species richness, tropical dry forest, tropical rain forest.

Resumen

Los bosques tropicales de Mesoamérica, desde México hasta el norte de Colombia, constituyen una de las biotas más diversas del planeta. Sin embargo, su historia evolutiva y los procesos que dieron forma a su ensamblaje siguen siendo poco comprendidos, lo que ha llevado a que con frecuencia se les considere como una sola unidad biótica. En este estudio utilizamos a la familia Annonaceae como linaje modelo para reconstruir las trayectorias evolutivas de la región, integrando cientos de marcadores nucleares obtenidos de especímenes de herbario históricos y contemporáneos. Este marco filogenómico de alta resolución, que abarca más del 80 % de las ~2,500 especies de la familia, revela patrones consistentes de diferenciación florística y evolutiva, así como contrastes marcados entre el norte (MegaMéxico 2 y el Caribe) y el sur de Mesoamérica. Las regiones septentrionales muestran una diversidad filogenética más profunda pese a su menor riqueza de especies, reflejo de un prolongado aislamiento y de una mayor heterogeneidad ambiental; en contraste, el sur presenta alta riqueza, pero redundancia filogenética, resultado de colonizaciones recientes y filtros climáticos. Los tiempos de divergencia subrayan el papel central de las oscilaciones climáticas del Mioceno-Plioceno y de la actividad orográfica más reciente en la conformación de los bosques de Mesoamérica, con los bosques secos emergiendo como refugios de historia evolutiva y centros de radiación, mientras que los bosques húmedos albergan de manera consistente linajes más jóvenes. En conjunto, estos hallazgos cuestionan la visión de Mesoamérica como una sola unidad evolutiva y muestran que sus bosques tropicales son sistemas dinámicos, moldeados repetidamente por ciclos de cambio climático y aislamiento geográfico.

Palabras clave: bosque tropical seco, bosque tropical lluvioso, endemismo, diversidad filogenética, riqueza de especies.

Herbarium collections serve as invaluable repositories of botanical knowledge, providing crucial insights into plant phylogeny, ecology, and evolution. Recent advances in molecular techniques and genomic methodologies have enabled the extraction of genetic information from preserved specimens, unlocking previously inaccessible data. This has markedly expanded both the taxonomic and geographic reach of phylogenetic analyses by including species that were historically excluded due to their presence in remote or inaccessible regions, as well as rare, unique, or poorly represented specimens. By incorporating these overlooked sources, researchers have enhanced evolutionary and taxonomic inferences, offering a more comprehensive framework for understanding plant biodiversity and its spatial and temporal distribution patterns.

This study aims to explore the evolutionary history of Mesoamerican tropical forests, stretching from Mexico to northern Colombia, by using the plant family Annonaceae as a model system. Annonaceae represents a key lineage within these forests, and recent advances have enabled the amplification of hundreds of nuclear markers specific to the family (Couvreur *et al.* 2019, Nge *et al.* 2024). These high-resolution markers have significantly improved the ability to infer evolutionary relationships among species and individuals and have proven effective even when applied to both historical and contemporary herbarium specimens (Ortiz-Rodriguez *et al.* 2024). Therefore, a robust phylogenomic backbone now encompasses over 80 % of the approximately 2,500 species in the family, providing a solid framework for exploring phylogenetic relationships and evolutionary timescales (Couvreur 2026). Building on this foundation, the present study investigates the evolutionary trajectories of Mesoamerican Annonaceae species across geographic regions, focusing on their phylogenetic relationships, divergence times, and diversification dynamics.

Rationale. Mesoamerica, extending from south-central Mexico and the Caribbean to northern Colombia, is one of the most biologically diverse regions worldwide. This exceptional richness stems not only from its role as a biotic interchange between North and South America (Marshall *et al.* 1982, Webb 2006), but also from its function as a center of *in situ* diversification (Ortiz-Rodriguez *et al.* 2018). The interplay of geographic features, such as the Isthmus of Panama, the Talamanca Range, the Nicaraguan Depression, and the Isthmus of Tehuantepec, together with a dynamic climatic history, has repeatedly fragmented and reconfigured habitats, fostering evolutionary radiation across multiple scales (*e.g.*, Daza *et al.* 2010, Ramírez-Barahona & Eguiarte 2013, Ornelas *et al.* 2013).

The tropical forests of Mesoamerica, including tropical rain, dry, and montane cloud forests, form a branched configuration encompassing six principal floristic centers: Central America (from southern Nicaragua to Colombia), the Caribbean, the Yucatán Peninsula, the Chiapas-Honduras corridor, and the forest regions of eastern and western Mexico (Magallón *et al.* 2014, Montaña-Arias *et al.* 2016, Prieto-Torres *et al.* 2018). The latter four belong to MegaMexico 2, a biogeographic unit characterized by endemism and ecological complexity (Rzedowski 1991). Nevertheless, defining floristic boundaries across MegaMexico 2 has proven notoriously difficult (Wendt 1993, Cevallos-Ferriz *et al.* 2012, Magallón *et al.* 2014). Palaeobotanical evidence shows that lineages once restricted to specific biomes now occur in other vegetation types, suggesting a shared origin (Cevallos-Ferriz *et al.* 2012). For example, western Mexico's dry forests are thought to derive from moist forests that underwent aridification following the uplift of the Sierra Madre Occidental and the Trans-Mexican Volcanic Belt, which altered precipitation and fragmented humid habitats (Becerra 2005, De-Nova *et al.* 2011, Cevallos-Ferriz *et al.* 2012). Subsequent tectonic pulses created barriers and rain shadows, further isolating and disconnecting regions, particularly between the Gulf and western Mexico (De-Nova *et al.* 2006). Climatic fluctuations during the Miocene and Pleistocene intensified cycles of expansion and contraction, fostering divergence and shaping the more recent distribution ranges of many lineages (Lozano-García *et al.* 2007, Ortiz-Rodriguez *et al.* 2020). Today, the marked climatic seasonality characteristic of northern Mesoamerica (MegaMexico 2) continues to play a decisive role in structuring species distributions across the region (Lozano-García *et al.* 2007, Magallón *et al.* 2014). Notably, numerous South American taxa reach their northernmost limits in Costa Rica, Panama, and southern Nicaragua (Wendt 1993, Ortiz-Rodriguez *et al.* 2018).

Floristic records indicate that most lineages in MegaMexico 2's tropical forests have persisted in this region since the Eocene, supported by prolonged continental isolation and geological stability (Wendt 1993). These lineages trace their origins to taxa that migrated from Asia, Africa, and South America during the Eocene, Neogene, and Pleistocene (Wendt 1993, Ortiz-Rodriguez *et al.* 2018). Over time, many diversified *in situ*, establishing MegaMexico 2 as both a corridor and a cradle of tropical biodiversity (Sosa *et al.* 2018). In contrast, southern Mesoamerica emerged through the uplift of the Central American volcanic arc and the closure of the Isthmus of Panama, processes that positioned it as a transitional zone or biogeographic bridge shaped by overlapping influences from northern Mesoamerica, the Caribbean, and South America (Pérez-Escobar *et al.* 2019). The role of the Caribbean in the evolution of Mesoamerican flora remains relatively understudied, yet hypotheses such as the GAARlandia land bridge underscore its importance as a pathway connecting South America and the Caribbean (Tong *et al.* 2019). The older geological history of the major Caribbean islands, compared with southern Mesoamerica, suggests that their flora reflects both ancient processes of colonization and diversification as well as more recent *in situ* radiations (Tong *et al.* 2019).

Based on these biogeographic scenarios, a spatially heterogeneous pattern of phylogenetic diversity, species richness, and endemism is expected across Mesoamerica. Specifically, in MegaMexico 2 and the Caribbean, prolonged isolation and older geological histories should favor high levels of endemism and deeper phylogenetic diversity, reflecting ancient lineages that diversified *in situ*. By contrast, southern Mesoamerica, shaped as a transitional zone, is expected to harbor greater species richness due to the overlap of lineages from northern Mesoamerica, the Caribbean, and South America. This richness may either reflect concentration within a few dominant lineages or genuinely high phylogenetic diversity resulting from the convergence of unrelated lineages.

To test these hypotheses, we focus on Annonaceae, an ideal study system that integrates ecological representativity, morphological diversity, and a well-resolved taxonomy and phylogeny (Nge *et al.* 2024). As the most species-rich family within Magnoliidae, with 110 genera and more than 2,500 species of trees, lianas, and shrubs, Annonaceae are structurally dominant in tropical forests worldwide (Nge *et al.* 2024). Their trimerous flowers, differentiated perianth, and numerous stamens and carpels sustain diverse pollination networks, while their varied fruits support dispersal by birds, mammals, and other vertebrates, underscoring their ecological importance (Gottsberger 2012, Onstein *et al.* 2019). Beyond this, their cultural and economic value is considerable, with species cultivated for edible fruits and widely used in traditional medicine. Recent taxonomic revisions and phylogenomic data now covering over 80 % of species have established a robust evolutionary framework, making Annonaceae uniquely suited for biogeographic and diversification studies (Couvreur 2026). In Mesoamerica (from south-central Mexico and the Caribbean to northern Colombia), the family is not especially species-rich but exhibits striking endemism, with ~280 species across 20 genera, high turnover among subregions, and limited overlap (Schatz 2025, Hurtado-Reveles *et al.* 2026). This combination of ecological dominance, species richness, and robust taxonomic and phylogenomic resolution provides a powerful lens for disentangling the distinct evolutionary processes shaping MegaMexico 2, southern Mesoamerica, and the Caribbean, subregions recurrently treated as a single biotic and evolutionary unit, a perspective that obscures their internal history and dynamic trajectories of tropical biodiversity.

Materials and methods

Species sampling, geographic distribution, and time-calibrated phylogeny. The study focuses on species occurring in the Mesoamerican region, as defined above. To compile the dataset, we consulted the most recent publications on Mesoamerican flora, along with references specific to the flora of Mexico and the Caribbean (Schatz 2025, Hurtado-Reveles *et al.* 2026). These sources included monographs, species descriptions, taxonomic treatments, and checklists (*e.g.* Maas *et al.* 2011, Ortiz-Rodriguez *et al.* 2024). Based on this literature, we constructed a list of species present in the region, which was taxonomically curated to verify accepted names and to identify synonyms or invalid names for the region (Rainer 2026). For each species, we obtained geographic data from the Global Biodiversity Information Facility (GBIF 2024, GBIF 2026 a,b), complemented with herbarium specimens. The distributional range of each species was further constrained based on the geographic information reported in the consulted taxonomic treat-

ments (Schatz 2025, Hurtado-Reveles *et al.* 2026). In addition, records of introduced species (*e.g.*, *Cananga odorata* (Lam.) Hook.f. & Thomson, *Polyalthia longifolia* (Sonn.) Thwaites) as well as *Annona muricata* L., *Annona cherimola* Mill., *Annona reticulata* L., and *Annona squamosa* L. were excluded from the dataset, as they were considered doubtful due to their long history of human management. After verifying the correct application of scientific names and excluding managed species, the final dataset was reduced from 318 names and 26,787 records to 281 accepted names and 22,958 georeferenced records in Mesoamerica.

For each species occurring in Mesoamerica, DNA was extracted directly from herbarium-preserved leaf material, with type specimens prioritized whenever possible. This formed part of the broader strategy of the GLOBAL project, which aims to sequence all species of Annonaceae (Couvreur 2026). When type material was unavailable, samples were obtained from specimens that clearly conformed to the species as currently circumscribed and recognized by taxonomic specialists in the respective groups.

Phylogenetic relationships among species were inferred using a hybrid-capture phylogenomic approach with the Annonaceae-specific baiting kit (Couvreur *et al.* 2019), designed to capture 469 orthologous exons in the family. Extraction, library preparation, sequencing protocols, and bioinformatic pipelines followed Soulé *et al.* (2023, 2024). Briefly, paired reads were processed using the HybPiper v. 1 pipeline (Johnson *et al.* 2016) to eliminate potential paralogs and generate cleaned alignments for phylogenetic analyses. Reads were mapped to a reference to recover corresponding exons, which were assembled into contigs. Exons and introns for each gene were combined into supercontigs. Each locus was aligned using MAFFT v. 7.305 (Katoh *et al.* 2002) with the -auto function, and alignments were trimmed with Gblocks v. 0.91 (Castresana 2000) to remove poorly aligned regions (introns and intergenic regions). Potential paralogs (multiple contigs mapping to a single reference sequence) were flagged by constructing gene trees for each locus in RAxML v. 8.2.12 (Stamatakis 2014). Loci identified with multiple copies were classified as paralogs and excluded from downstream analyses (Joyce *et al.* 2025). With this strategy, our final dataset included 231 Mesoamerican species with validated genomic data suitable for phylogenetic analyses.

Phylogenetic relationships were then inferred in a global context of the family, including more than 80 % of all Annonaceae species and 100 % of the recognized genera (Nge *et al.* 2024). Two complementary strategies were employed: a concatenated matrix approach and a gene tree approach (Dagallier *et al.* 2023). In both cases, datasets were filtered to retain exons with at least 75 % coverage across 75 % of sampled individuals (the 75/75 ratio; Couvreur *et al.* 2019). For the concatenated dataset, phylogenetic trees were inferred using Maximum Likelihood (ML) in RAxML v. 8.2.12 (Stamatakis 2014) with the -f a option, 100 bootstrap replicates, and the GTRGAMMA model. For the gene tree approach, phylogenetic trees were inferred for each supercontig in RAxML v. 8.2.9 under the GTRGAMMA model with 100 bootstrap replicates. A summary species tree was then inferred in ASTRAL-III v. 5.5.11 (Zhang *et al.* 2018) under the multi-species coalescent model. Branch support was assessed using quartet support (QS; -t 1 option) and local posterior probabilities (LPP; default parameters). The species tree was rooted at the divergence between *Magnolia* sp. and Annonaceae (Helmstetter *et al.* 2024). Both phylogenetic hypotheses were visualized and annotated using the ape and phytools packages in R v. 4.2.0 (Paradis & Schliep 2019, Revell 2024).

A species-level dated phylogenetic tree for Annonaceae was generated using a backbone-and-patch approach (Up- ham *et al.* 2019), following the Bayesian sequential-subtree methodology previously applied to mammals (Álvarez-Carretero *et al.* 2022). For the backbone, we relied on a recently published phylogenomic analysis of magnoliids (Helmstetter *et al.* 2025) based on the Angiosperms353 baiting kit (Johnson *et al.* 2019). This dataset included 235 unique species representing 199 genera and all 18 families across the orders Canellales, Laurales, Magnoliales (which includes Annonaceae), and Piperales. Within Annonaceae, 78 genera were represented, covering all subfamilies and most tribes. Calibration was achieved using five expert-validated and phylogenetically placed magnoliid fossils (Massoni *et al.* 2015). To ensure computational feasibility, we divided the sequenced samples into 18 smaller monophyletic patches, each ranging from 25 to 224 species. Each patch included at least one taxon overlapping with the backbone phylogeny to allow tree merging. Phylogenetic trees were inferred from concatenated sequence data using RAxML v. 8.2.9 (Stamatakis 2014) under the GTRGAMMA model with 100 bootstrap replicates, and species trees were estimated under the coalescent model using ASTRAL v. 5.7.8 (Zhang *et al.* 2017). Constrained starting

chronograms were generated with treePL v. 1.0 (Smith & O'Meara 2012), and 30 clock-like loci were selected using SortaDate (Smith *et al.* 2018). Bayesian divergence-time analyses were conducted in BEAST v. 2.6.7 (Bouckaert *et al.* 2019) via CIPRES (Miller *et al.* 2011), employing a relaxed lognormal clock and birth–death priors, with convergence assessed in Tracer v. 1.7 (Rambaut *et al.* 2018). Patch analyses incorporated two to three secondary calibrations per clade, derived from the fossil-dated backbone. The 18 clade-specific trees were subsequently patched onto the backbone using modified scripts from Upham *et al.* (2019), resulting in a Maximum Clade Credibility dated phylogenetic tree of Annonaceae. To account for uncertainty, we merged 500 posterior trees (after a 20 % burn-in) from each patch sequentially with 500 posterior backbone topologies, thereby generating a posterior distribution of 500 dated Annonaceae trees.

Floristic Regionalization of Mesoamerica and Climatic Characterization of Annonaceae Assemblages. Based on the geographic and taxonomic dataset, we constructed a presence-absence matrix at a resolution of 0.5 degrees. This resolution represents an intermediate scale between finer grids (*e.g.*, $< 0.25^\circ$), which can exaggerate gaps caused by uneven sampling, and coarser grids (*e.g.*, $\geq 1^\circ$), which can obscure ecological patterns. The 0.5° resolution was therefore chosen to balance interpretability and computational feasibility. To minimize biases from unequal sampling effort, we retained only a single record of each species per grid cell, ensuring that the matrix reflects species presence rather than sampling intensity. Floristic similarity among cells was analyzed to characterize the distribution of the main assemblages of Annonaceae species in the region, with all analyses conducted in the R environment (R Core Team 2022). Floristic similarity among spatial cells was quantified using the Jaccard index implemented in the *vegan* package (Faith *et al.* 1987, Oksanen *et al.* 2022), which provides a robust measure of shared species composition derived from presence-absence data. The resulting similarity matrix was converted into a distance matrix, forming the basis for hierarchical clustering. Hierarchical clustering was conducted with the Ward.D2 method, implemented through the *hclust* function (Murtagh & Legendre 2014, R Core Team 2022), which minimizes within-cluster variance and produces compact, ecologically interpretable clusters that capture the distributional patterns of Annonaceae species across the region. This procedure is a standard approach in floristic characterization, as it enables the identification of groups of sites with comparable species assemblages and facilitates the visualization of their relationships. To identify the optimal number of clusters in a way that was both statistically supported and biologically interpretable, several complementary approaches were implemented in the R environment (R Core Team 2022). Silhouette analysis was first conducted using the *cluster* package (Maechler *et al.* 2022). For each candidate solution ($k = 2\text{--}15$), clusters were generated with *cutree* function, and silhouette widths were calculated with the *silhouette* function based on the Jaccard distance matrix, implemented through the *factoextra* package (Kassambara & Mundt 2020). The average silhouette width was then plotted across values of k , and the solution with the highest mean silhouette was identified as the most parsimonious partition. Because silhouette analysis often favors larger numbers of clusters, segmented regression was applied using the *segmented* package to refine the selection of k (Muggeo 2008). A linear model was fitted to silhouette values across k , and breakpoints were estimated to detect changes in slope that indicate thresholds in clustering structure. In addition, cluster stability was evaluated with the Adjusted Rand Index (ARI) implemented in the *mclust* package (Scrucca *et al.* 2023). Successive clustering solutions (k and $k-1$) were compared, and ARI values were plotted to assess the consistency of partitions across different numbers of clusters. High ARI values indicated stable solutions, whereas abrupt declines suggested less reliable partitions (Scrucca *et al.* 2023). The selected clustering solution was mapped by converting cells into polygons with the *sf* and *purrr* packages (Pebesma & Bivand 2023, Wickham & Henry 2023) and overlaid on regional shapefiles obtained from the *rnaturalearth* package (Massicotte & South 2023). To characterize the climatic space of clusters, values of four WorldClim bioclimatic variables (Fick & Hijmans 2017)-mean annual temperature (bio1), temperature seasonality (bio4), annual precipitation (bio12), and precipitation seasonality (bio15)-were extracted from raster layers using the *raster* and *sp* packages (Pebesma & Bivand 2005, Bivand *et al.* 2013, Hijmans 2023). These values were standardized and subjected to Principal Component Analysis (PCA) using the base R function *prcomp* (R Core Team 2022). PCA scores (PC1 and PC2) were then integrated into the cluster dataset to characterize the ecological space occupied by each floristic cluster.

To facilitate ecological interpretation, the PCA-defined environmental space was divided into four quadrants based on the sign of PC1 and PC2 values: Humid-Warm (PC1 < 0, PC2 < 0), Humid-Cold (PC1 < 0, PC2 > 0), Dry-Warm (PC1 > 0, PC2 < 0), and Dry-Cold (PC1 > 0, PC2 > 0). This classification provided a simplified framework for comparing the distribution of clusters across major climatic gradients. We further quantified the relative position of each cluster within the PCA-defined environmental space. First, the global centroid was calculated as the mean of all PCA scores. Centroids for each cluster were then computed, and their Euclidean distances to the global centroid were used as a measure of climatic divergence. This approach allowed us to identify more central clusters (*i.e.*, closer to average regional conditions) versus those that are more divergent, occupying extreme or less represented environments. Then, differences in climatic dispersion among clusters were evaluated using *betadisper* function, implemented in the *vegan* package (Oksanen *et al.* 2022), which quantifies variation in distances around cluster centroids. The resulting model was tested with ANOVA, and significant pairwise differences among clusters were identified using Tukey's HSD (R Core Team 2022).

Spatial distribution of species richness and phylogenetic diversity. To evaluate the spatial distribution of species richness and phylogenetic diversity of Annonaceae in Mesoamerica, we standardized occurrence records by aggregating them into 0.5° grid cells and retaining only one record per species per cell, ensuring that richness was measured as the number of distinct species rather than the number of observations. From this presence-absence matrix, species richness (SR) was calculated directly as the sum of species per cell, while phylogenetic diversity (PD) was estimated using the pruned phylogenetic tree, restricted to the 231 species with validated genomic data, implemented in the R package *picante* (Kembel *et al.* 2010). Because SR and PD are inherently correlated, we fitted a loess regression of PD against SR and extracted the residuals (rPD), which represent the surplus or deficit of evolutionary history relative to species richness. Thus, the results are expressed not as global values but as spatial patterns of SR, and rPD, including a bivariate visualization of SR versus rPD, where high values of both metrics indicate assemblages that are simultaneously species-rich and evolutionarily diverse, while negative rPD values suggest phylogenetically redundant communities (Faith 1992, Forest *et al.* 2007).

We also evaluated floristic similarity among countries where Annonaceae species occur, in order to assess whether their similarity is determined by the evolutionary history of the regions in which they are distributed (MegaMexico 2, southern Mesoamerica, and the Caribbean). To accomplish this, we assembled a database of occurrence records containing species names and geographic coordinates (latitude and longitude). Each record was spatially assigned to a country by overlaying point coordinates onto national polygons obtained from the *rnaturalearth* dataset (Massicotte & South 2023), using spatial functions of the *sf* package (Pebesma & Bivand 2023). When coordinates fell outside polygon boundaries (*e.g.*, near borders), the nearest country polygon was assigned with the function *st_nearest_feature*. From these assignments, we generated a presence-absence matrix of species by country, implemented with *dplyr* (Wickham *et al.* 2023) and *tidyr* (Wickham *et al.* 2024), producing a binary table in which rows corresponded to species and columns to countries. Floristic similarity among countries was quantified using Jaccard distances, calculated with the *vegan* package (Oksanen *et al.* 2022), and hierarchical clustering was performed with the *average* linkage method to generate dendrograms of floristic relationships.

To complement these analyses, we represented the data as a bipartite network linking species and countries. Using the *igraph* package (Csardi & Nepusz 2006), we constructed a graph in which nodes represented either species or countries, and edges indicated the occurrence of a species in a given country. The network was visualized with *ggraph* (Pedersen 2024), employing a force-directed layout (Fruchterman-Reingold algorithm) that naturally grouped countries according to shared species. This dual approach, clustering dendrograms and bipartite networks, allowed us to visualize both hierarchical patterns of floristic similarity and the structural relationships underlying species distributions across Mesoamerican countries.

Evolutionary time scales of Annonaceae in Mesoamerica. To explore both the temporal depth and the phylogenetic structure of Annonaceae assemblages in Mesoamerica, we applied two complementary approaches. First, to evaluate the

temporal depth of assemblages, we estimated the mean divergence age of species present in each grid cell. The dated phylogenetic tree was processed in R using the package *ape* (Paradis & Schliep 2019), converted into a standard phylo object, and tip labels were harmonized with the species in the presence–absence matrix. Node depths were calculated with the *node.depth.edgelen* function and subtracted from the maximum tree height to obtain relative node ages. For each species, a custom function identified its parental node and assigned the age of divergence with its sister lineage, thereby generating a vector of divergence ages across all species. These lineage ages were integrated with the presence–absence matrix using *dplyr* (Wickham *et al.* 2023), allowing the computation of the mean divergence age of species present in each grid cell. Spatial polygons representing 0.5° grid cells were constructed with *sf* and *purrr* (Pebesma & Bivand 2023, Wickham & Henry 2023), intersected with the study region, and mapped with *ggplot2* (Wickham 2016) to visualize the geographic distribution of divergence ages. A continuous color gradient was applied to highlight spatial variation in the temporal depth of assemblages. Mean divergence ages per cell were then combined with climatic variables (bio1, bio4, bio12, bio15), forest types, and floristic clusters. These data were analyzed through a principal component analysis (PCA) on climatic variables, implemented with *stats::prcomp* (R Core Team 2022), and visualized with *ggplot2* (Wickham 2016). Points in PCA space were scaled by divergence age, colored by floristic cluster, and annotated with forest-type centroids, thereby linking the distribution of temporally deep lineages to climatic gradients and floristic structure.

Second, to assess the phylogenetic structure of communities, we calculated two complementary indices of phylogenetic diversity using the package *picante* (Kembel *et al.* 2010), based on the cophenetic distance matrix obtained with *ape* (Paradis & Schliep 2019): Mean Pairwise Distance (MPD) and Mean Nearest Taxon Distance (MNTD). MPD represents the average phylogenetic distance between all pairs of species in a community, providing a measure of overall evolutionary depth and broad divergence patterns. In contrast, MNTD estimates the average distance between each species and its closest relative, making it sensitive to the presence of closely related clades and useful for detecting clustering or repulsion among near lineages. While MPD reflects deeper divergences and global phylogenetic structure, MNTD captures more recent evolutionary proximity. Comparing both metrics thus complements the lineage age analysis, offering a comprehensive view of community assembly by distinguishing between diversity patterns shaped by ancient divergences and those driven by recent relationships. Finally, maps of MPD and MNTD were visualized with *ggplot2* (Wickham 2016), employing color gradients to represent spatial variation in the phylogenetic structure of Annonaceae communities.

Lastly, to elucidate the evolutionary and biogeographic dynamics of lineages, we calculated the differences between genus divergence ages and species divergence ages, an approach we term the *Genus-Species Age Difference (GSAD) framework*. High GSAD values indicate that species diverged long after the origin of their genera, suggesting later diversification events. When such species are consistently clustered in specific geographic regions, this pattern may reflect spatial shifts in diversification, highlighting potential corridors or nodes of lineage expansion. Conversely, low GSAD values correspond to species whose divergence ages are close to those of their genera, pointing to early radiations that occurred near the time of genus origin and potentially anchoring taxa to their ancestral biogeographic areas. Thus, the GSAD framework serves as a proxy for reconstructing the geographic history of lineages, linking divergence times to spatial distributions. Using *ape* (Paradis & Schliep 2019), we calculated node and tip ages from the BEAST tree to obtain divergence ages for each species and genus. We then computed GSAD values and averaged them across cells of a presence–absence matrix to obtain representative values per spatial unit. The *sf* and *purrr* R packages (Pebesma & Bivand 2023, Wickham & Henry 2023) were used to construct polygons for each grid cell and to intersect them with the study region to delimit the analysis area. Finally, *ggplot2* (Wickham 2016) was employed to generate spatial gradient maps of GSAD values, thereby visualizing comparative evolutionary divergence and highlighting biogeographic patterns across the landscape.

Results

Floristic Regionalization of Mesoamerica and Climatic Characterization of Annonaceae Assemblages. Silhouette analysis showed a steady increase in average silhouette width across candidate solutions, with values improving from 0.067 at $k = 3$ to 0.152 at $k = 9$, and continuing to rise toward 0.190 at $k = 15$. Although higher k values yielded slight-

ly greater silhouette widths, segmented regression identified a breakpoint at approximately $k = 7$, suggesting that solutions beyond this threshold provided diminishing returns in terms of structural improvement. Cluster stability was further assessed with the Adjusted Rand Index (ARI). Comparisons between successive solutions revealed consistently high ARI values around $k = 7-9$ (0.96-0.97), indicating strong agreement between partitions and confirming the robustness of solutions in this range. In contrast, solutions with fewer clusters showed lower stability (*e.g.*, ARI = 0.63 at $k = 5$), while those with larger numbers of clusters tended to fragment biologically coherent assemblages. Based on these complementary criteria, the seven-cluster solution was selected (Figure 1). Statistically, it balanced high silhouette width with strong ARI stability, while biologically it provided ecologically meaningful groupings that captured the major distributional patterns of Annonaceae species across the region without over-partitioning the data. This choice ensured that the resulting clusters were both quantitatively supported and interpretable in terms of floristic composition and ecological differentiation.

The dendrogram reveals two major floristic assemblages (Figure 1A). The first, comprising Clusters 1, 2, and 3, shows high internal homogeneity (fusion heights 0.00-0.81) and reflects strong similarity in species composition. Cluster 1 dominates the Atlantic slope from Mexico to Colombia and extends into southern Mesoamerica, though it is weakly represented in western Mexico and the Caribbean. Cluster 2 runs discontinuously along the Pacific slope from Sonora to Nicaragua, with a disjunct presence in the Yucatán Peninsula, while Cluster 3 is confined to the Caribbean. Together, these clusters form a continental–Caribbean complex that underscores the close affinity between Central American and Caribbean floras. In contrast, Clusters 4-7 exhibit greater internal heterogeneity (fusion heights 0.92-1.13). Cluster 4 is species-poor and sporadically distributed along coastal regions of Mexico and Nicaragua, whereas Clusters 5, 6, and 7 are concentrated in northern Mesoamerica: Cluster 5 is exclusive to Yucatán, Cluster 6 restricted to northeastern Mexico (Veracruz-Tamaulipas), and Cluster 7 limited to western Mexico. This pattern highlights the distinctiveness of western Mexico and the uniqueness of the Yucatán Peninsula, while southern Mesoamerica is represented by a single cluster and the Caribbean flora remains clearly differentiated from mainland Mesoamerica (Figure 1A). Overall, MegaMexico 2 harbors six of the seven clusters, reflecting both the cohesion of southern and Caribbean floras and the diversification of northern Mexican communities.

The betadisper analysis revealed that floristic clusters differ significantly in the dispersion of their climatic distances (Figure 1B). The ANOVA confirmed highly significant differences among groups ($F = 21.18$, $P < 2.2e-16$). Average distances to the median varied across clusters, with Cluster 1 showing the greatest dispersion (1.81) and Cluster 5 and 3 the lowest (0.83 and 0.61, respectively). This indicates that climatic variability within clusters is not homogeneous: some clusters, particularly Cluster 1, which is broadly distributed across continental Mesoamerica, encompass a wider range of conditions, whereas others, such as Cluster 5 and 3, restricted to the Yucatán Peninsula and the Caribbean, are concentrated in comparatively more limited environments (Figure 1B). Pairwise comparisons using Tukey's HSD further indicated that Cluster 1 differs significantly from Clusters 2, 3, 4, 5, 6, and 7, with Cluster 1 consistently exhibiting greater dispersion. In addition, Cluster 5 showed significantly lower dispersion compared to Cluster 2. Complementary measures of cohesion and separation support these differences. Average silhouette widths ranged from 0.015 (Cluster 2) to 1.000 (Cluster 6), with intermediate values for Cluster 1 (0.178), Cluster 3 (0.044), Cluster 4 (0.837), Cluster 5 (0.262), and Cluster 7 (0.402). Cluster dissimilarities (heights) varied between 0.000 (Cluster 6) and 3.416 (Cluster 3), with values of 2.917 (Cluster 1), 3.365 (Cluster 2), 0.704 (Cluster 4), 3.079 (Cluster 5), and 2.456 (Cluster 7). Together, these metrics show that clusters differ not only in their dispersion but also in their internal cohesion and degree of separation.

Beyond internal variability, the PCA-defined environmental space highlights contrasting positions among clusters relative to the global centroid. Clusters 3, 4, and 5 are located close to the overall centroid (0.118, 0.285, and 0.173, respectively), indicating less divergence from average regional conditions and suggesting greater internal heterogeneity. In contrast, Clusters 6, 7, and 2 are positioned farther from the global centroid (1.76, 1.83, and 1.03, respectively), reflecting greater climatic divergence and association with more extreme or less-represented environments. Cluster 1 occupies an intermediate position (0.552), not central but also not as extreme, consistent with its high internal dispersion. This pattern suggests that Cluster 1 may represent specialized conditions in certain regions, yet still fall within relatively common climatic ranges.

In terms of vegetation types, Clusters 2, 6, and 7 can be grouped as floristic assemblages of tropical dry forests, since they occupy ecological spaces characteristic of this biome (Figure 1B). Yet, our analyses reveal important differentiation within this vegetation type. Clusters 2 and 7 are associated with tropical dry and warm forests, such as those extending along the Pacific coast from western Mexico to Nicaragua and also occurring in the Yucatán Peninsula. Cluster 6, by contrast, corresponds to drier but cooler forests, including those found in eastern Mexico around the Tamaulipas region (Figure 1). The remaining clusters (1, 3, 4, and 5), including the Caribbean cluster, are primarily distributed in humid ecological spaces typical of tropical rainforests and montane cloud forests. These clusters also exhibit greater environmental heterogeneity, encompassing a broader range of climatic conditions compared to the dry forest assemblages (Figure 1B).

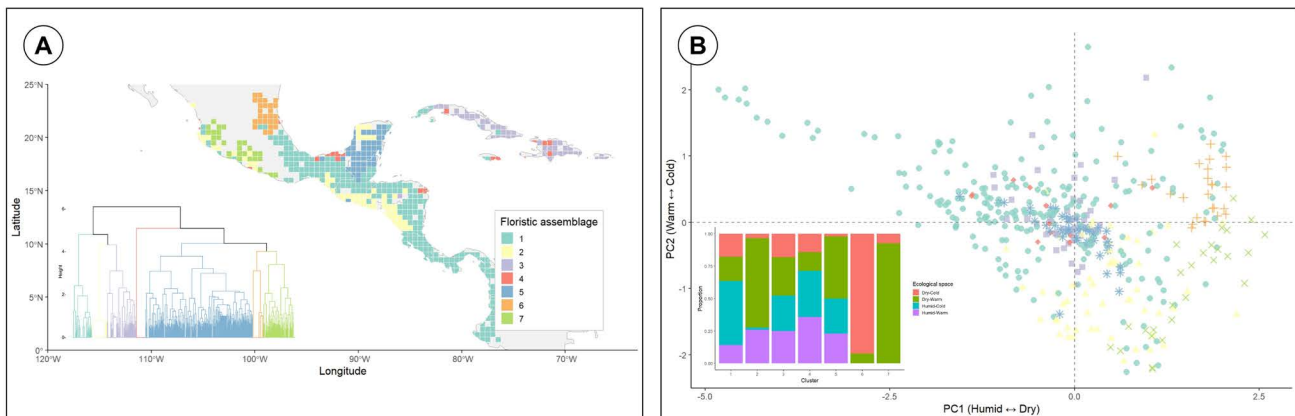


Figure 1. Distribution of floristic clusters (assemblages) based on Annonaceae species across Mesoamerica, with climatic characterization of each cluster. (A) Geographic distribution of Mesoamerican floristic clusters obtained through floristic similarity analysis using the Ward hierarchical clustering method, distinguished by color. Floristic assemblages are notably more diverse in northern Mesoamerica (MegaMexico 2) than in the Caribbean and southern portions of the region. (B) Climatic space defined by four bioclimatic variables (bio1, bio4, bio12, and bio15), illustrating the position of each floristic cluster. Colors correspond to those used in the map, with shapes providing additional differentiation. Clusters 2, 6, and 7 are located away from the center [0,0], occupying ecological niches typical of dry forests, whereas clusters closer to the center occur under more humid conditions. Cluster 1 is the most dispersed and heterogeneous in terms of climatic conditions. The bar chart illustrates the proportion of each cluster associated with each quadrant, thereby reflecting the distribution of assemblages along the humidity–temperature gradient of the ecological space.

Spatial distribution of species richness and phylogenetic diversity. A total of 281 species across 20 genera with geographic records were compiled for the Mesoamerican region, among which the phylogenetic placement of 231 species was successfully established. Species richness is not uniformly distributed across Mesoamerica, reflecting spatial heterogeneity in diversity patterns (Figure 2A). The southern portion of the region, particularly southern Nicaragua and the Chocó area of Colombia, is by far the most species-rich (181 species). A secondary zone of high richness occurs in MegaMexico 2 (99 species), extending from Mexico to Honduras along the Atlantic slope (Figure 2A). In contrast, the Caribbean, with 36 species, can be considered comparatively species-poor (Figure 2A).

Floristic similarity analyses support the distinction between MegaMexico 2, southern Mesoamerica, and the Caribbean (Supplementary Figure 1). The dendrogram based on Jaccard distances revealed strong clustering among geographically proximate continental countries, while Caribbean islands formed their own clusters, reflecting their distinctive floristic composition. The bipartite network provided a complementary view of the structural relationships between species and countries. Continental countries with higher richness (Mexico, Colombia, Costa Rica, Panama) emerged as central nodes with numerous connections and a substantial proportion of endemic taxa (Supplementary Figure 1). In contrast, Caribbean islands appeared as peripheral nodes, linked to fewer species and characterized by

Evolutionary history of the Mesoamerican forest

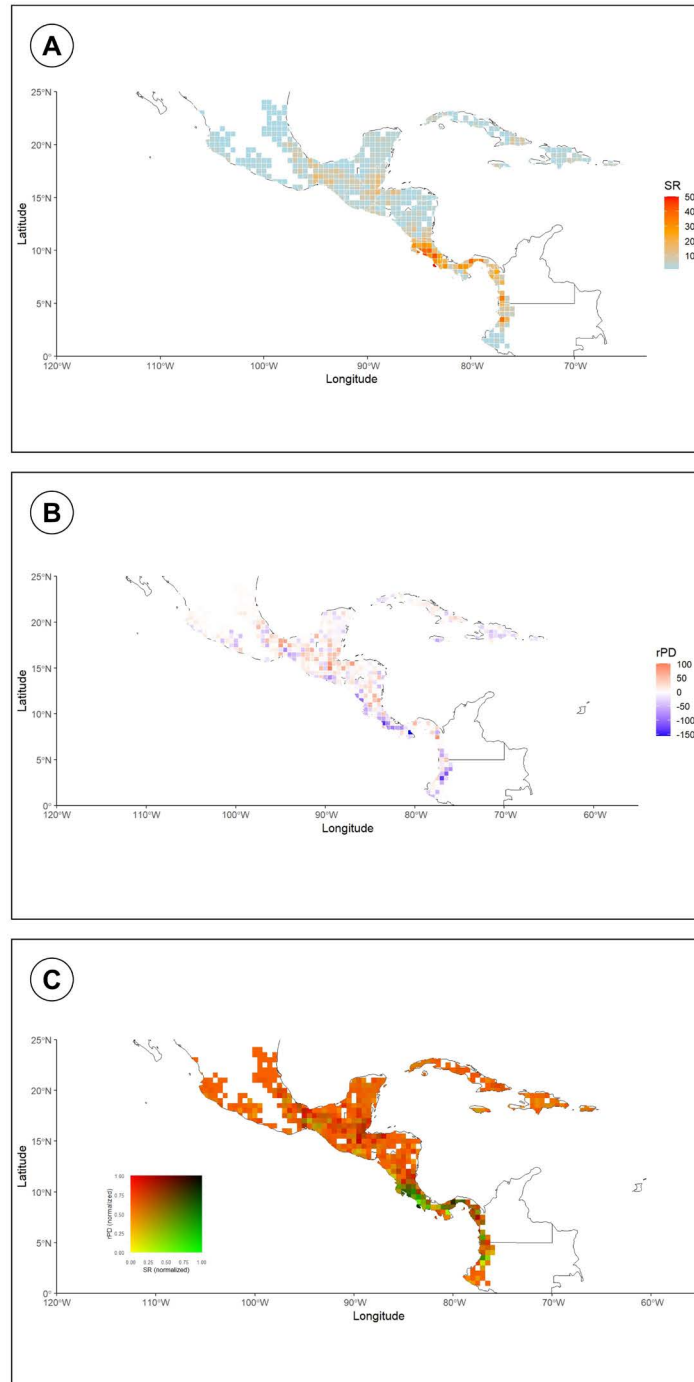


Figure 2. Variation in species richness and phylogenetic diversity of Annonaceae assemblages across Mesoamerica. (A) Species richness (SR). The highest richness is concentrated in southern Mesoamerica (Costa Rica to northern Colombia), with secondary centers in Mexico and the Caribbean. (B) Residual phylogenetic diversity (rPD). High values indicate phylogenetically heterogeneous communities, whereas low or negative values reflect phylogenetically redundant communities. (C) Bivariate map of species richness and residual phylogenetic diversity. Northern Mesoamerica (MegaMexico 2) and the Caribbean are characterized by communities where phylogenetic diversity exceeds expectations based on species richness. In contrast, southern Mesoamerica shows communities with high species richness but comparatively lower phylogenetic diversity, except in the Chocó region, which harbors more complex phylogenetic assemblages.

marked endemism that clearly distinguishes them from their continental counterparts. The force-directed layout emphasized the modularity of the network, showing that Caribbean connections are more closely aligned with Central America and northern South America than with Mexico ([Figure S1](#)). This pattern underscores the singularity of the Mexican flora, highlights the role of Central America as a zone of overlap and a biogeographic bridge, and reveals the Caribbean's affinities with southern countries ranging from Costa Rica to Ecuador.

The spatial distribution of residual phylogenetic diversity (rPD) reveals generally heterogeneous patterns across Mesoamerica ([Figure 2B](#)). Notably, in most communities of southern Mesoamerica, high species richness exceeds phylogenetic diversity, resulting in predominantly negative rPD values and indicating phylogenetically redundant assemblages ([Figure 2C](#)). In contrast, rPD values are markedly higher and more broadly distributed in MegaMexico 2 and the Caribbean, suggesting that communities in these regions are phylogenetically more heterogeneous than expected given their species richness ([Figure 2B](#)). The bivariate spatial distribution clearly shows that communities in the Caribbean and MegaMexico 2 include areas where high species richness coincides with unexpectedly high evolutionary diversity, as well as regions with low richness but predominantly positive rPD values, where a few species represent substantial evolutionary diversity ([Figure 2C](#)). To a lesser extent, these regions also contain areas of high richness coupled with negative rPD values. In contrast, most communities in southern Mesoamerica are characterized by high species richness but negative rPD values ([Figure 2C](#)). An exception is the Chocó region, where cells with high richness and positive rPD values predominate, suggesting that, unlike the rest of southern Mesoamerica, the Chocó hosts communities that are less redundant ([Figure 2C](#)). This distinctive pattern likely reflects the role of the Chocó region as a biogeographic convergence zone between Central and South America, its function as a floristic refugium, and the persistence of ancient lineages under long-term geological and climatic stability.

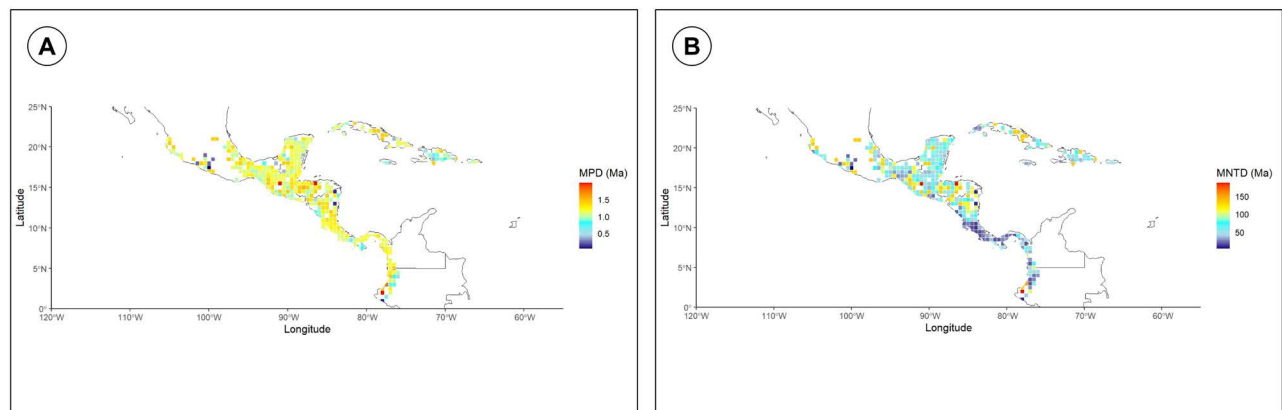


Figure 3. Phylogenetic structure of Annonaceae assemblages in Mesoamerica, measured by Mean Pairwise Distance (MPD) and Mean Nearest Taxon Distance (MNTD). (A) Mean phylogenetic distance among all pairs of species within each community (MPD), reflecting overall evolutionary depth and broad-scale phylogenetic structure. (B) Mean distance between each species and its closest relative (MNTD), capturing more recent evolutionary relationships and revealing clear geographic differentiation among regions.

Evolutionary time scales of Annonaceae in Mesoamerica. Mean Pairwise Distance (MPD), expressed in millions of years, showed relatively homogeneous values across Mesoamerica, with medium to high divergence times broadly distributed throughout the region ([Figure 3A](#)). In contrast, Mean Nearest Taxon Distance (MNTD) exhibited greater spatial heterogeneity, with shorter divergence times concentrated in certain areas ([Figure 3B](#)). The spatial distribution of these two evolutionary time-scale metrics revealed contrasting patterns across the region. This reflects the fact that MPD averages all pairwise distances, thereby highlighting deep evolutionary splits within assemblages while smoothing over local variation. In contrast, Mean Nearest Taxon Distance (MNTD) revealed clear geographic differentiation: very recent divergences (low values) were concentrated within assemblages in southern Mesoamerica, from Nicaragua to northern Colombia, indicating assemblages dominated by closely related lineages ([Figure 3B](#)).

Meanwhile, MegaMexico 2 (Mexico to Honduras) and the Caribbean displayed medium to high values, consistent with communities composed of more distantly related taxa (Figure 3B).

The spatial distribution of average species age divergence indicates that present-day Annonaceae have only recently become part of Mesoamerica's forests (Figure 4A). Between the late Miocene and throughout the Pliocene and Pleistocene [Mean = 4.14 Mya; min = 0.57 Mya; max = 17.8 Mya], these species integrated into the forests they now inhabit. Our results indicate that species in southern Mesoamerica are younger than those found in other parts of the region (Figure 4A). Our results reveal that, on average, communities in western Mexico harbor species with the deepest divergence times relative to their sister species, followed by communities in the Yucatán Peninsula and the Caribbean. We also show that species inhabiting the dry forests of western Mexico are, on average, older than those found in the dry forests of the Yucatán Peninsula and the Gulf of Mexico. In contrast, species assemblages in wetter forests are composed, on average, of comparatively younger species (Figure S2). GSAD values align with this pattern. Southern Mesoamerica is dominated by higher values, suggesting that its assemblages originated long after their respective genera (Figure 4B). This supports the hypothesis of recent colonization or geographic expansion into the area. In contrast, low to very low GSAD values are common across MegaMexico 2 and the Caribbean, suggesting that species in these regions largely trace their origins to the putative geographic origin of their genera (Figure 4B). This supports hypotheses of early *in situ* diversification. When examined in climatic space, low GSAD values occur mainly in the dry forests of continental Mesoamerica (clusters 2, 4, 5, 7) and in some warm and humid clusters, such as cluster 4, restricted to the Mesoamerican costas (Figure 4C). High GSAD values, by contrast, are concentrated in the warm and humid continental forests (primarily cluster 1). The Caribbean represents a particular case: although composed of warm and humid forests, it hosts communities with both low and high GSAD values (Figure 4C). This reveals a mixed assemblage structure, with evidence of both old colonization and *in situ* diversification, alongside more recent colonizations or origins.

Discussion

We reconstruct the evolutionary trajectories of Mesoamerican Annonaceae within a geographic framework and historical context derived from genetic data obtained from herbarium specimens. Our analysis uncovers consistent patterns of regional differentiation, reflecting the evolutionary history of the lineages examined. As emblematic elements of tropical forests, Annonaceae provide a critical perspective on the processes that generated Mesoamerica's exceptional species richness and structured the contemporary distribution of its plant communities.

Floristic differentiation across Mesoamerica. The floristic clusters recovered in our analysis reveal both sharply delimited regions and diffuse transitions, underscoring the complex biogeographic structure of Mesoamerica. Clear differentiation among the Caribbean, southern Mesoamerica, and MegaMexico 2 aligns with regionalization patterns reported for other plant groups (Montaño-Arias *et al.* 2016, Prieto-Torres *et al.* 2018), in which these areas are recognized as the Caribbean subregion (encompassing MegaMexico 2 and the Caribbean islands) and the Brazilian subregion (including southern Mesoamerica). Consistent with these studies, species in MegaMexico 2 are associated with more seasonal and drier tropical forests, whereas those in the Caribbean and southern Mesoamerica show stronger affinities with humid, less seasonal environments (Magallón *et al.* 2014, Montaño-Arias *et al.* 2016, Prieto-Torres *et al.* 2018). This ecological contrast likely contributes to the high proportion of species restricted to each region, further reinforcing their differentiation.

Within MegaMexico 2 (the Mesoamerican dominion *sensu* Montaño-Arias *et al.* 2016), we identified several floristic clusters that also match patterns documented in previous studies (Montaño-Arias *et al.* 2016, Prieto-Torres *et al.* 2018). In particular, there is a clear differentiation between forests along the Pacific slope and those on the Atlantic slope. On the Atlantic side, our results are consistent with earlier works in showing the floristic distinctiveness of the Yucatán Peninsula, the humid forests from southern Veracruz to northern Nicaragua, and the seasonally dry tropical forests of Tamaulipas and San Luis Potosí (Montaño-Arias *et al.* 2016, Prieto-Torres *et al.* 2018). On the Pacific side,

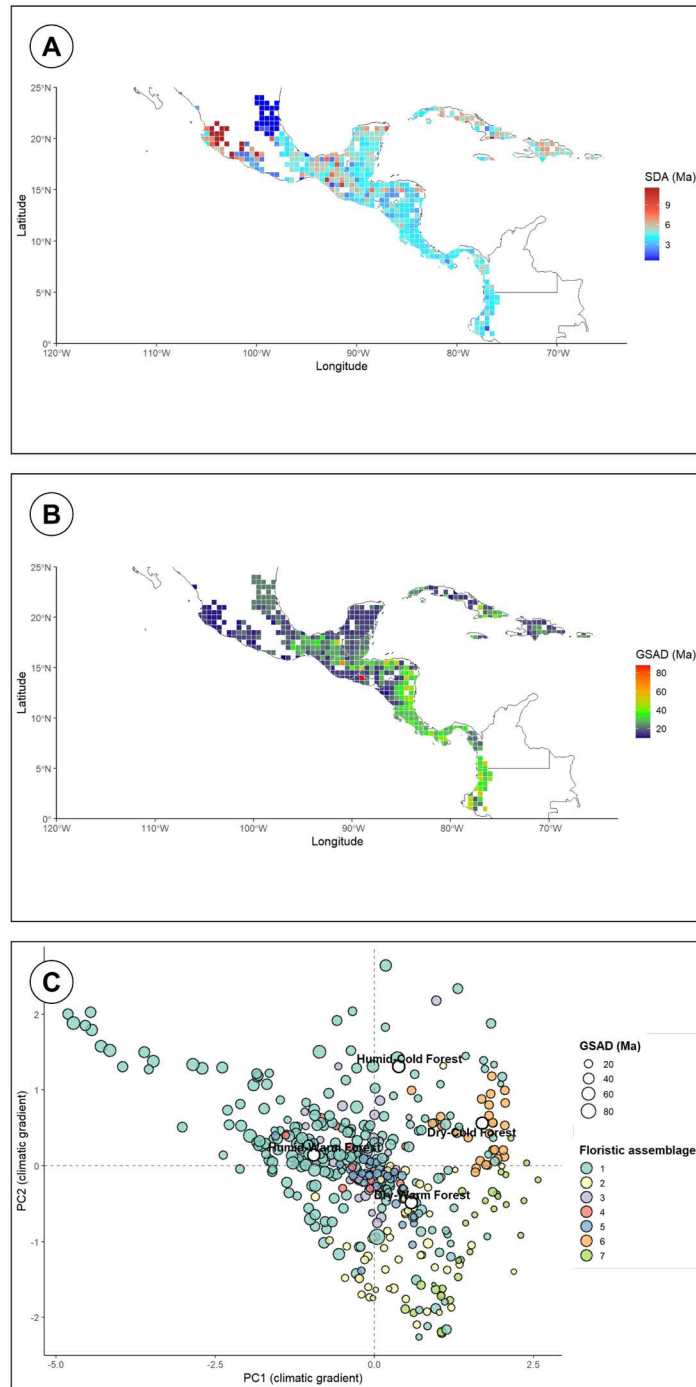


Figure 4. Evolutionary timescales of Mesoamerican Annonaceae. (A) Average divergence time of each community relative to its sister species (SDA). In western Mexico, Annonaceae species are comparatively older. (B) Average divergence time of each species relative to the divergence age of its genus (GSAD). Species with divergence times close to those of their respective genera are mainly distributed in northern Mesoamerica (MegaMexico 2) and the Caribbean. In southern Mesoamerica, species diverged much later than their genera, indicating more recent colonization and diversification events consistent with the evolutionary history of the region. (C) Ecological distribution of GSAD values. Low GSAD values, representing early diversification, are associated with drier environments in MegaMexico, whereas high values are linked to humid forests in southern Mesoamerica and, to a lesser extent, in MegaMexico 2 and the Caribbean. The mixture of low and high GSAD values in MegaMexico 2 and the Caribbean reflects long processes of geographic isolation combined with both early and recent diversification events.

our findings similarly highlight the Balsas region as floristically distinct, as well as the dry forests extending from Sonora to Nicaragua. A novel outcome of our analysis is Cluster 4, restricted to coastal areas on both slopes and the Caribbean, largely determined by the presence of *Annona glabra*. This species is confined to mangrove and flood-plain forest systems, and its presence as a distinct cluster underscores the role of ecological specialization in shaping floristic boundaries (Sánchez-García *et al.* 2026).

The contrasting levels of cohesion and dispersion among clusters reveal how floristic diversity in Mesoamerica is structured by both ecological breadth and biogeographic history (Magallón *et al.* 2014, Crews & Esposito 2020). Clusters with broad internal dispersion, such as Cluster 1, encompass assemblages spanning wide environmental ranges and multiple vegetation types. This heterogeneity is largely driven by the MegaMexico 2 portion of the cluster, where climatic variation and habitat diversity promote greater substructure and phylogenetic differentiation. In southern Mesoamerica, by contrast, Cluster 1 communities remain floristically cohesive and phylogenetically redundant, indicating that climate has played a limited role in their evolutionary assembly and that historical connectivity has sustained compositional uniformity across the region (Gratani 2014). Clusters with low dispersion, such as those restricted to the Yucatán Peninsula or the Caribbean, represent floristic units occurring under relatively homogeneous ecological conditions. Rather than reflecting narrower niches, these assemblages highlight the concentration of species derived from distinct lineages in consistent settings, where climatic variability among sites is reduced.

The positioning of clusters in environmental space highlights complementary dynamics to those observed in their internal cohesion. Groups located near the global centroid, such as the Caribbean and Yucatán clusters (Clusters 3 and 5), represent assemblages embedded within average regional conditions, where distinctiveness arises primarily from geographic isolation rather than climatic extremes (Licona-Vera & Ornelas 2014, Ortiz-Rodriguez *et al.* 2020, Vázquez-López *et al.* 2021, Kolchanova *et al.* 2021). In contrast, clusters positioned farther from the centroid (Clusters 2 and 7) are associated with divergent climatic regimes, reflecting adaptation to less common or more extreme environments. Broadly distributed clusters, particularly Cluster 1, illustrate how environmental gradients interact with geography: in MegaMexico 2, climatic variation and habitat diversity foster substructure and phylogenetic differentiation, whereas in southern Mesoamerica the same cluster remains cohesive and redundant, shaped more by historical connectivity than by climate. Taken together, these spatial patterns show that Mesoamerican floristic diversity is structured not only by geographic separation but also by ecological breadth and environmental homogeneity. This interplay between environmental space and internal dispersion clarifies how plant diversity is organized, reinforcing established regionalizations while revealing the ecological processes that generate both sharp boundaries and gradual transitions.

Contrasting evolutionary trajectories within Mesoamerica. Our results show that the evolutionary history of northern Mesoamerica (MegaMexico 2 and the Caribbean) differs markedly from that of southern Mesoamerica (from southern Nicaragua to northern Colombia). This contrast reflects the prolonged isolation of MegaMexico 2 and the Caribbean relative to southern Mesoamerica, which has served as a recent bridge between North and South America (Wendt 1993, Erkens *et al.* 2007, Crews & Esposito 2020). Consequently, southern Mesoamerica appears as a zone of colonization from the north and the Caribbean, subsequently enriched by intense *in situ* diversification (Pérez-Escobar *et al.* 2019). These findings are consistent with patterns reported in other studies (*e.g.*, Cavers *et al.* 2003, Gutiérrez-García & Vázquez-Domínguez 2013, Lander & Monro 2015, Morales-Jimenez *et al.* 2015, Ortiz-Rodriguez *et al.* 2018, Acha *et al.* 2021, Guzmán-Díaz *et al.* 2025), contradicting the long-standing but erroneous tradition of treating Mesoamerica as a single evolutionary unit (*e.g.*, Antonelli *et al.* 2018, Matos-Maraví *et al.* 2021).

Species richness is not a reliable predictor of phylogenetic diversity in Annonaceae. Across Mesoamerica, we observe a pronounced geographical mismatch between regions. In southern Mesoamerica, high species richness coincides with low residual phylogenetic diversity, resulting in more redundant communities. By contrast, in MegaMexico 2 and the Caribbean, residual phylogenetic diversity is much higher than expected given their relatively low richness species. A few exceptions occur, primarily in wetter forests, where taxonomic richness exceeds phylogenetic diversity. We interpret these patterns as the outcome of prolonged isolation in northern Mesoamerica and the Carib-

bean, combined with the more recent emergence of southern Mesoamerica (Wendt 1993, Crews & Esposito 2020). In MegaMexico 2, high seasonality generates greater biome heterogeneity, which supports both deep phylogenetic diversity and the origin of more recent lineages. In the Caribbean, although climatic conditions resemble those of southern Mesoamerica, geographic isolation from the continent has played a decisive role in shaping its flora (Crews & Esposito 2020). This has produced both ancient and recent lineages that are only loosely related, reflecting processes of diversification at different depths. Finally, southern Mesoamerica, despite its high species richness, forms phylogenetically redundant communities. The distribution of these lineages in wetter, less seasonal environments suggest climatic filtering during colonization, favoring the rapid expansion of closely related lineages.

Average divergence ages highlight sharp regional contrasts. In southern Mesoamerica, from Nicaragua to northern Colombia, low values indicate recent diversification, strong phylogenetic cohesion, and recent colonization. By contrast, MegaMexico 2 and the Caribbean show medium to high values, consistent with communities composed of more distantly related taxa and longer histories of lineage accumulation. This regional contrast is best interpreted as a recent phenomenon, given that Annonaceae species have occupied present-day Mesoamerican forests only within the last few million years (mean = 4.14 Mya; range = 0.57-17.8 Mya). Such timing supports the view that many tropical forest assemblages are of recent origin, while simultaneously underscoring the complex evolutionary dynamics that sustain these ecosystems (Richardson *et al.* 2001, Erkens *et al.* 2007, Antonelli *et al.* 2018, Pérez-Escobar *et al.* 2019). Placed within a broader temporal framework, the Miocene and Pliocene emerge as pivotal intervals in the evolution of Mesoamerican biota, a pattern increasingly documented in plant studies (Richardson *et al.* 2001, Erkens *et al.* 2007, Ortiz-Rodriguez *et al.* 2018, Vázquez-López *et al.* 2021, Antonelli *et al.* 2018, Sosa *et al.* 2018, Pérez-Escobar *et al.* 2019). Episodes of aridification, contraction of humid forests, and expansion of dry forests appear to have played a decisive role in structuring the distribution and assembly of Annonaceae in MegaMexico 2 (Becerra 2005). These findings reinforce the perspective that present-day tropical forests are not static relics of deep time but dynamic systems, repeatedly reshaped by climatic oscillations and geographic barriers.

Dry forests in Mesoamerica as sources of lineages and refuges of evolutionary history. Our results show that the driest forests (clusters 2, 5, and 7), spanning western Mexico to Nicaragua along the Pacific slope and extending into the Yucatán Peninsula, contain species whose divergence times closely match those of their genera. This pattern suggests regions of early radiation and taxa rooted in their ancestral biogeographic areas (Becerra 2005, De-Nova *et al.* 2011, Sosa *et al.* 2018). Such evidence supports the hypothesis of a dry forest origin in western Mexico (Becerra 2005) and the idea that tropical forests were once broadly distributed but progressively fragmented following the uplift of the Sierra Madre Occidental and, more markedly, volcanic activity along the western Trans-Mexican Volcanic Belt (Becerra 2005, Cevallos-Ferriz *et al.* 2012). These processes produced increasingly dry forests in western Mexico compared to those along the Gulf of Mexico. In the Yucatán Peninsula, several studies highlight its long-standing isolation from other forests (*e.g.*, Licona-Vera & Ornelas 2014, Ortiz-Rodriguez *et al.* 2020). The presence of disjunct species shared between Yucatán and western Mexico further reinforces the view of once continuous tropical forests that later fragmented (*e.g.*, De-Nova *et al.* 2006). By contrast, the dry forests of the Balsas Basin, San Luis Potosí, and Tamaulipas harbor communities with species whose divergence times are more distant from those of their genera, reflecting later radiations. This pattern aligns with proposed dynamics of dry forest expansion, indicating that these sites represent more recent formations (Becerra 2005, De-Nova *et al.* 2011).

In contrast to the history of dry forests, the wetter and more heterogeneous forests of Mesoamerica, extending from southern Mexico to Colombia (cluster 1) and in the Caribbean (cluster 3), are generally composed of species with more recent radiations. This challenges the traditional view of humid forests as “museums” of diversity. The Caribbean represents a special case: like MegaMexico 2, its communities reflect both early radiation and more recent diversification, a pattern consistent with insular systems. Thus, forests should be understood as products of dynamic diversification processes rather than static museums or cradles (Koenen *et al.* 2015, Sosa *et al.* 2018). In the Mesoamerican hotspot, climate has been a central driver of evolutionary trajectories, reinforcing the view of these forests as landscapes continually reshaped by environmental oscillations and geographic barriers.

Sources of Uncertainty. Although our interpretations are supported by solid and consistent results, certain considerations must be acknowledged. Phylogenetic relationships were inferred using a single terminal per species, a strategy that may bias divergence time estimates toward older values or obscure more recent divergence events and finer geographic structuring within Mesoamerican Annonaceae. These aspects, however, lie beyond the scope of the present study. Despite this limitation, our findings align with several independent studies that emphasize the role of the Miocene and Pliocene in shaping Mesoamerican biota, as well as the clear differentiation among regions within Mesoamerica. While our interpretations focus on the evolution of Mesoamerican forests, it is important to note that our analyses are based on a single lineage, representing only a fraction of the forest's diversity. Nevertheless, community-level studies and analyses of other lineages reveal congruent patterns, reinforcing the view that Mesoamerica's historical dynamics, particularly those associated with the concept of MegaMexico 2, affected entire plant communities. We therefore recommend conducting ancestral area analyses across multiple Mesoamerican plant lineages to further evaluate the role of MegaMexico 2 as both a center of *in situ* diversification and a source of lineages that dispersed to other regions of the American continent. Such comparative approaches will provide a more comprehensive understanding of the historical processes that shaped the extraordinary diversity of Mesoamerican forests.

Supplementary material

Supplemental data for this article can be accessed here: <https://doi.org/10.17129/botsci.3916>

Acknowledgements

We extend our sincere gratitude to Guadalupe Amancio and Moises Rangel Olguín for their valuable comments during the project discussions and for their encouraging words. We also thank the staff of the MEXU herbarium for their generous support during the revision of Annonaceae specimens.

Literature Cited

- Acha S, Linan A, MacDougal J, Edwards C. 2021. The evolutionary history of vines in a neotropical biodiversity hotspot: Phylogenomics and biogeography of a large passion flower clade (*Passiflora* section *Decaloba*). *Molecular Phylogenetics and Evolution* **164**: 107260. DOI: <https://doi.org/10.1016/j.ympev.2021.107260>
- Álvarez-Carretero S, Tamuri AU, Battini M, Nascimento FF, Carlisle E, Asher RJ, Yang Z, Donoghue PCJ, dos Reis M. 2022. A species-level timeline of mammal evolution integrating phylogenomic data. *Nature* **602**: 263-267. DOI: <https://doi.org/10.1038/s41586-021-04341-1>
- Antonelli A, Zizka A, Carvalho FA, Condamine FL. 2018. Amazonia is the primary source of Neotropical biodiversity. *Proceedings of the National Academy of Sciences of the United States of America (PNAS)* **115**: 6034-6039. DOI: <https://doi.org/10.1073/pnas.1713819115>
- Becerra JX. 2005. Timing the origin and expansion of the Mexican tropical dry forest. *Proceedings of the National Academy of Sciences of the United States of America (PNAS)* **102**: 10919-10923. DOI: <https://doi.org/10.1073/pnas.0409127102>
- Bivand RS, Pebesma E, Gómez-Rubio V. 2013. Applied Spatial Data Analysis with R. 2nd ed. UseR! Series. New York: Springer. xviii+405 pp. ISBN: 978-1-4614-7617-7.
- Bouckaert R, Vaughan TG, Barido-Sottani J, Duchêne S, Fourment M, Gavryushkina A, Heled J, Jones G, Kühnert D, De Maio N, Matschiner M. 2019. BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. *Plos Computational Biology* **15**: e1006650. DOI: <https://doi.org/10.1371/journal.pcbi.1006650>
- Castresana J. 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular Biology and Evolution* **17**: 540-552. DOI: <https://doi.org/10.1093/oxfordjournals.molbev.a026334>
- Cavers S, Navarro C, Lowe AJ. 2003. Chloroplast DNA phylogeography reveals colonization history of a Neotropi-

- cal tree, *Cedrela odorata* L., in Mesoamerica. *Molecular Ecology* **12**: 1451-1460. DOI: <https://doi.org/10.1046/j.1365-294X.2003.01810.x>
- Csardi G, Nepusz T. 2006. The igraph software package for complex network research. *InterJournal, Complex Systems* 1695. <https://igraph.org>
- Cevallos-Ferriz SRS, González-Torres EA, Calvillo-Canadell L. 2012. Perspectiva paleobotánica y geológica de la biodiversidad en México. *Acta Botanica Mexicana* **100**: 317-350. DOI: <https://doi.org/10.21829/abm100.2012.39>
- Couvreur TLP. 2026. *Tropical rain forests a GLOBAL approach*. Couvreur Lab. <https://couvreurlab.org/global.html> (accessed March, 2026)
- Couvreur TLP, Helmstetter AJ, Koenen EJM, Bethune K, Brandão RDS, Little SA, Sauquet H, Erkens RHJ. 2019. Phylogenomics of the major tropical plant family Annonaceae using targeted enrichment of nuclear genes. *Frontiers in Plant Science* **9**: 1-15. DOI: <https://doi.org/10.3389/fpls.2018.01941>
- Crews SC, Esposito LA. 2020. Towards a synthesis of the Caribbean biogeography of terrestrial arthropods. *BMC Evolutionary Biology* **20**: 1-27. DOI: <https://doi.org/10.1186/s12862-019-1576-z>
- Dagallier L-PMJ, Mbago FM, Couderc M, Gaudeul M, Grall A, Loup C, Wieringa JJ, Sonké B, Couvreur TLP. 2023. Phylogenomic inference of the African tribe Monodoreae (Annonaceae) and taxonomic revision of *Dennettia*, *Uvariadendron* and *Uvariopsis*. *PhytoKeys* **233**: 1-200. DOI: <https://doi.org/10.3897/phytokeys.233.103096>
- Daza JM, Castoe TA, Parkinson CL. 2010. Using regional comparative phylogeographic data from snake lineages to infer historical processes in Middle America. *Ecography* **33**: 343-354. DOI: <https://doi.org/10.1111/j.1600-0587.2010.06281.x>
- De-Nova JA, Medina R, Montero JC, Weeks A, Rosell JA, Olson ME, Eguiarte LE, Magallón S. 2011. Insights into the historical construction of species-rich Mesoamerican seasonally dry tropical forests: The diversification of *Bursera* (Burseraceae, Sapindales). *New Phytologist* **193**: 276-287. DOI: <https://doi.org/10.1111/j.1469-8137.2011.03909.x>
- De-Nova JA, Sosa V, Wurdack KJ. 2006. Phylogenetic relationships and the description of a new species of *Enriquebeltrania* (Euphorbiaceae s.s.): An enigmatic genus endemic to Mexico. *Systematic Botany* **31**: 533-546. DOI: <https://doi.org/10.1043/05-70.1>
- Erkens RHJ, Chatrou LW, Maas JW, van der Niet T, Savolainen V. 2007. A rapid diversification of rainforest trees (*Guatteria*; Annonaceae) following dispersal from Central into South America. *Molecular Phylogenetics and Evolution* **44**: 399-411. DOI: <https://doi.org/10.1016/j.ympev.2007.03.020>
- Faith DP. 1992. Conservation evaluation and phylogenetic diversity. *Biological Conservation* **61**: 1-10. DOI: [https://doi.org/10.1016/0006-3207\(92\)91201-3](https://doi.org/10.1016/0006-3207(92)91201-3)
- Faith DP, Minchin PR, Belbin L. 1987. Compositional dissimilarity as a robust measure of ecological distance. *Vegetatio* **69**: 57-68. DOI: <https://doi.org/10.1007/BF00038687>
- Fick SE, Hijmans RJ. 2017. WorldClim 2: New 1 km spatial resolution climate surfaces for global land areas. *International Journal of Climatology* **37**: 4302-4315. DOI: <https://doi.org/10.1002/joc.5086>
- Forest F, Grenyer R, Rouget M, Davies TJ, Cowling RM, Faith DP, Balmford A, Manning JC, Procheş Ş, van der Bank M, Reeves G, Hedderson TAJ, Savolainen V. 2007. Preserving the evolutionary potential of floras in biodiversity hotspots. *Nature* **445**: 757-760. DOI: <https://doi.org/10.1038/nature05587>
- GBIF [Global Biodiversity Information Facility]. 2024. GBIF Occurrence Download. <https://doi.org/10.15468/dl.64tyna> (accessed February 26, 2024).
- GBIF [Global Biodiversity Information Facility]. 2026a. GBIF Occurrence Download. <https://doi.org/10.15468/dl.64qbjt> (accessed January 20, 2026).
- GBIF [Global Biodiversity Information Facility]. 2026b. GBIF Occurrence Download. <https://doi.org/10.15468/dl.9ctkx9> (accessed January 20, 2026).
- Gottsberger G. 2012. How diverse are Annonaceae with regard to pollination? *Botanical Journal of the Linnean Society* **169**: 245-261. DOI: <https://doi.org/10.1111/j.1095-8339.2011.01209.x>
- Gratani L. 2014. Plant phenotypic plasticity in response to environmental factors. *Advances in Botany* **2014**: 1-17.

- <https://doi.org/10.1155/2014/208747>
- Gutiérrez-García TA, Vázquez-Domínguez E. 2013. Consensus between genes and stones in the biogeographic and evolutionary history of Central America. *Quaternary Research* **79**: 311-324. DOI: <https://doi.org/10.1016/j.yqres.2012.12.007>
- Guzmán-Díaz S, Aldaba Núñez FA, Veltjen E, Asselman P, Jiménez JE, Valdés Sánchez J, Pino Infante G, Callejas Posada R, Vázquez García JA, Larridon I, Park S, Kim S, Martínez Salas EM, Samain M-S. 2025. There and back again: Historical biogeography of neotropical magnolias based on high-throughput sequencing. *BMC Ecology and Evolution* **25**: 1-18. DOI: <https://doi.org/10.1186/s12862-025-02379-7>
- Helmstetter AJ, Ezedin Z, de Lirio EJ, de Oliveira SM, Chatrou LW, Erkens RHJ, Larridon I, Leempoel K, Maurin O, Roy S, Baker WJ, Couvreur TLP, Forest F, Sauquet H. 2025. Toward a phylogenomic classification of magnoliids. *American Journal of Botany* **112**: e16451. <https://doi.org/10.1002/ajb2.16451>
- Helmstetter AJ, Ezedin Z, Lirio EJ, Oliveira SMD, Chatrou LW, Erkens RHJ, Larridon I, Leempoel K, Maurin O, Roy S, Zuntini AR, Baker WJ, Couvreur TLP, Forest F, Sauquet H. 2024. Towards a phylogenomic classification of Magnoliidae. *American Journal of Botany* **112**: e16451. DOI: <https://doi.org/10.1002/ajb2.16451>
- Hijmans R. 2023. *raster: Geographic data analysis and modeling*. R package version 3.6-26. Available at: <https://CRAN.R-project.org/package=raster> (accessed May 2026)
- Hurtado-Reveles L, Rangel-Olguín M, Amancio G, Ortiz-Rodríguez AE. 2026. Diversity, distribution and priority conservation areas for Annonaceae in Mexico. *Botanical Sciences*. DOI: <https://doi.org/10.17129/botsci.3834>
- Johnson MG, Gardner EM, Liu Y, Medina R, Goffinet B, Shaw AJ, Zerega NJC, Wickett NJ. 2016. HybPiper: Extracting coding sequence and introns for phylogenetics from high-throughput sequencing reads using target enrichment. *Applications in Plant Sciences* **4**: 1600016. DOI: <https://doi.org/10.3732/apps.1600016>
- Johnson MG, Pokorny L, Dodsworth S, Botigué LR, Cowan RS, Devault A, Eiserhardt WL, Epiwalage N, Forest F, Kim JT, Leitch IJ, Maurin O, Soltis DE, Soltis PS, Wong GK-S, Baker WJ. 2019. A universal probe set for targeted sequencing of 353 nuclear genes from any flowering plant designed using k-medoids clustering. *Systematic Biology* **68**: 594-606. DOI: <https://doi.org/10.1093/sysbio/syy086>
- Joyce EM, Schmidt-Lebuhn AN, Orel HK, Nge FJ, Anderson BM, Hammer TA, McLay TGB. 2025. Navigating phylogenetic conflict and evolutionary inference in plants with target capture data. *Australian Systematic Botany* **38**: SB24011. DOI: <https://doi.org/10.1071/SB24011>
- Kassambara A, Mundt F. 2020. *factoextra: Extract and visualize the results of multivariate data analyses*. R package version 1.0.7.
- Katoh K, Misawa K, Kuma KI, Miyata T. 2002. MAFFT: A novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research* **30**: 3059-3066. DOI: <https://doi.org/10.1093/nar/gkf436>
- Kembel SW, Cowan PD, Helmus MR, Cornwell WK, Morlon H, Ackerly DD, Blomberg SP, Webb CO. 2010. Picante: R tools for integrating phylogenies and ecology. *Bioinformatics* **26**: 1463-1464. DOI: <https://doi.org/10.1093/bioinformatics/btq166>
- Koenen EJM, Clarkson JJ, Pennington TD, Chatrou LW. 2015. Recently evolved diversity and convergent radiations of rainforest mahoganies (Meliaceae) shed new light on the origins of rainforest hyperdiversity. *New Phytologist* **207**: 327-339. DOI: <https://doi.org/10.1111/nph.13490>
- Kolchanova S, Komissarov A, Kliver S, Mazo-Vargas A, Afanador Y, Velez-Valentín J, de la Rosa RV, Castro-Marquez S, Rivera-Colon I, Majeske AJ, Wolfsberger WW, Hains T, Corvelo A, Martinez-Cruzado J-C, Glenn TC, Robinson O, Koepfli K-P, Oleksyk TK. 2021. Molecular phylogeny and evolution of Amazon parrots in the Greater Antilles. *Genes* **12**: 608. DOI: <https://doi.org/10.3390/genes12040608>
- Lander TA, Monro A. 2015. Conservation of *Brosimum alicastrum*, an underutilized crop and keystone forest tree species; a potential win-win for conservation and development in Latin America. *Biodiversity and Conservation* **24**: 1917-1930. DOI: <https://doi.org/10.1007/s10531-015-0913-9>
- Licona-Vera Y, Ornelas JF. 2014. Genetic, ecological and morphological divergence between populations of the endangered Mexican sheartail hummingbird (*Doricha eliza*). *Plos One* **9**: e101870. DOI: <https://doi.org/10.1371/journal.pone.0101870>

- Lozano-García M del S, Caballero M, Ortega B, Sosa S. 2007. Tracing the effects of the Little Ice Age in the tropical lowlands of eastern Mesoamerica. *Proceedings of the National Academy of Sciences* **104**: 16200-16203. DOI: <https://doi.org/10.1073/pnas.0707896104>
- Maas PJM, Westra LYT, Rainer H, Lobão AQ, Erkens RHJ. 2011. An updated index to genera, species, and infraspecific taxa of Neotropical Annonaceae. *Nordic Journal of Botany* **29**: 257-356. DOI: <https://doi.org/10.1111/j.1756-1051.2011.01092.x>
- Maechler M, Rousseeuw P, Struyf A, Hubert M, Hornik K. 2022. *cluster: Cluster Analysis Basics and Extensions*. R package version 2.1-3. <https://CRAN.R-project.org/package=cluster> (accessed May 2026)
- Magallón S, Pennington RT, Villaseñor JL, Rodríguez P, Honorio-Coronado E, Foster R, Harris D, Wieringa J. 2014. The influence of regional history and ecological affinity in the angiosperm composition of Mexican lowland tropical rain forests. In: Stevens WD, Montiel OM, Raven PH, eds. *Paleobotany and Biogeography: A Festschrift for Alan Graham in His 80th Year*. St. Louis, Missouri, Missouri: Botanical Garden Press, pp. 287-325. ISBN: 978-0-915279-97-5.
- Marshall LG, Webb SD, Sepkoski JJ, Raup DM. 1982. Mammalian evolution and the Great American Interchange. *Science* **215**: 1351-1357. DOI: <https://www.science.org/doi/10.1126/science.215.4538.1351>
- Massicotte P, South A. 2023. *rnaturalearth: World map data from Natural Earth*. R package version 1.0.1. <https://CRAN.R-project.org/package=rnaturalearth> (accessed May 2026)
- Massoni J, Couvreur TLP, Sauquet H. 2015. Five major shifts of diversification through the long evolutionary history of Magnoliidae (angiosperms). *BMC Evolutionary Biology* **15**: 1-14. DOI: <https://doi.org/10.1186/s12862-015-0320-6>
- Matos-Maraví P, Wahlberg N, Freitas AVL, Devries P, Antonelli A, Penz CM. 2021. Mesoamerica is a cradle and the Atlantic Forest is a museum of Neotropical butterfly diversity: Insights from the evolution and biogeography of *Brassolini* (Lepidoptera: Nymphalidae). *Biological Journal of the Linnean Society* **133**: 704-724. DOI: <https://doi.org/10.1093/biolinnean/blab034>
- Miller MA, Pfeiffer W, Schwartz T. 2011. The CIPRES Science Gateway. http://www.phylo.org/sub_sections/portal/portal_papers.php (accessed April 2026)
- Montaño-Arias G, Luna-Vega I, Morrone JJ, Espinosa D. 2016. Biogeographical identity of the Mesoamerican dominion with emphasis on seasonally dry tropical forests. *Phytotaxa* **376**: 277-290. DOI: <https://doi.org/10.11646/phytotaxa.376.6.3>
- Muggeo VMR. 2008. *segmented: An R package to fit regression models with broken-line relationships*. *R News* **8**: 20-25. <https://cran.r-project.org/doc/Rnews/> (accessed May 2026)
- Murtagh F, Legendre P. 2014. Ward's hierarchical agglomerative clustering method: Which algorithms implement Ward's criterion? *Journal of Classification* **31**: 274-295. DOI: <https://doi.org/10.1007/s00357-014-9161-z>
- Nge FJ, Chaowasku T, Damthongdee A, Wiya C, Soulé VRC, Rodrigues-Vaz C, Bruy D, Mariac C, Chatrou LW, Chen J, Choo LM, Dagallier L-PMJ, Erkens RHJ, Johnson DM, Leeratiwong C, Lobão AQ, Lopes JC, Martínez-Velarde MF, Munzinger J, Murray NA, Neo WL, Rakotoarinivo M, Ortiz-Rodriguez AE, Sonké B, Thomas DC, Wieringa JJ, Couvreur TLP. 2024. Complete genus-level phylogenomics and new subtribal classification of the pantropical plant family Annonaceae. *Taxon* **73**: 1341-1369. DOI: <https://doi.org/10.1002/tax.13260>
- Oksanen J, Simpson G, Blanchet F, Kindt R, Legendre P, Minchin P, O'Hara R, Solymos P, Stevens M, Szoecs E, Wagner H, Barbour M, Bedward M, Bolker B, Borcard D, Carvalho G, Chirico M, De Cáceres M, Durand S, Evangelista H, FitzJohn R, Friendly M, Furneaux B, Hannigan G, Hill M, Lahti L, McGlinn D, Ouellette M, Ribeiro Cunha E, Smith T, Stier A, Ter Braak C, Weedon J. 2022. *vegan: Community Ecology Package*. R package version 2.6-4. <https://CRAN.R-project.org/package=vegan> (accessed May 2026)
- Onstein RE, Kissling WD, Chatrou LW, Couvreur TLP, Morlon H, Sauquet H. 2019. Which frugivory-related traits facilitated historical long-distance dispersal in the custard apple family (Annonaceae)? *Journal of Biogeography* **46**: 1874-1888. DOI: <https://doi.org/10.1111/jbi.13552>
- Ornelas JF, Sosa V, Soltis DE, Daza JM, González C, Soltis PS, Gutiérrez-Rodríguez C, Espinosa de los Monteros A, Castoe TA, Bell C, Ruiz-Sanchez E. 2013. Comparative phylogeographic analyses illustrate the complex evolutionary history of threatened cloud forests of northern Mesoamerica. *Plos One* **8**: e56283. DOI: <https://doi.org/10.1371/journal.pone.0056283>

- Ortiz-Rodriguez AE, Licona-Vera Y, Vásquez-Aguilar AA, Hernández-Soto M, López-Huicochea EA, Ornelas JF. 2020. Genetic differentiation among *Psittacanthus rhynchanthus* (Loranthaceae) populations: Novel phylogeographic patterns in the Mesoamerican tropical lowlands. *Plant Systematics and Evolution* **306**: 1-20. DOI: <https://doi.org/10.1007/s00606-020-01638-y>
- Ortiz-Rodriguez AE, Nge FJ, Rodrigues-Vaz C, Soulé V, Schatz GE, Martínez-Velarde MF, Hurtado-Revels L, Rangel-Olguin M, Couvreur TLP. 2024. Taxonomy, systematics and conservation of the highly threatened and endemic Mexican genus *Tridimeris* (Annonaceae). *Plant Systematics and Evolution* **310**: 1-32. DOI: <https://doi.org/10.1007/s00606-024-01929-8>
- Ortiz-Rodriguez AE, Ornelas JF, Ruiz-Sanchez E. 2018. A jungle tale: Molecular phylogeny and divergence time estimates of the *Desmopsis*–*Stenanona* clade (Annonaceae) in Mesoamerica. *Molecular Phylogenetics and Evolution* **122**: 80-94. DOI: <https://doi.org/10.1016/j.ympev.2018.01.021>
- Paradis E, Schliep K. 2019. ape 5.0: An environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* **35**: 526-528. DOI: <https://doi.org/10.1093/bioinformatics/bty633>
- Pebesma E, Bivand R. 2005. Classes and methods for spatial data in R. *R News* **5**: 9-13. <https://CRAN.R-project.org/doc/Rnews/> (accessed May 2026)
- Pebesma E, Bivand R. 2023. *Spatial data science: With applications in R*. New York: Chapman and Hall/CRC, eBook ISBN9780429459016. DOI: <https://doi.org/10.1201/9780429459016>
- Pedersen T. 2024. ggraph: An implementation of grammar of graphics for graphs and networks. R package version 2.2.1. <https://CRAN.R-project.org/package=ggraph> (accessed May 2026)
- Pérez-Escobar OA, Lucas E, Jaramillo C, Monro A, Morris SK, Bogarín D, Greer D, Dodsworth S, Aguilar-Cano J, Sanchez Meseguer A, Antonelli A. 2019. The origin and diversification of the hyperdiverse flora in the Chocó biogeographic region. *Frontiers in Plant Science* **10**: 1328. DOI: <https://doi.org/10.3389/fpls.2019.01328>
- Prieto-Torres DA, Rojas-Soto OR, Bonaccorso E, Santiago-Alarcon D, Navarro-Sigüenza AG. 2018. Distributional patterns of Neotropical seasonally dry forest birds: A biogeographical regionalization. *Cladistics* **35**: 446-460. DOI: <https://doi.org/10.1111/cla.12366>
- R Core Team. 2022. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/> (accessed May 2026)
- Rainer H. 2026. *Annonaceae*. Herbarium, Department of Botany and Biodiversity Research, University of Vienna. <https://herbarium.botanik.univie.ac.at/annonaceae/listTax.php> (accessed March, 2026)
- Rambaut A, Drummond AJ, Xie D, Baele G and Suchard MA. 2018. Posterior summarisation in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology* **67**: 901-904. DOI: <https://doi.org/10.1093/sysbio/syy032>
- Ramírez-Barahona S, Eguiarte LE. 2013. The role of glacial cycles in promoting genetic diversity in the Neotropics: The case of cloud forests during the Last Glacial Maximum. *Ecology and Evolution* **3**: 725-738. DOI: <https://doi.org/10.1002/ece3.483>
- Revell LJ. 2024. phytools 2.0: An updated R ecosystem for phylogenetic comparative methods. *PeerJ* **12**: e16505. DOI: <https://doi.org/10.7717/peerj.16505>
- Richardson JE, Pennington RT, Pennington TD, Hollingsworth PM. 2001. Rapid diversification of a species-rich genus of Neotropical rain forest trees. *Science* **293**: 2242-2245. DOI: <https://www.science.org/doi/10.1126/science.1061421>
- Rzedowski J. 1991. Diversidad y orígenes de la flora fanerogámica de México. *Acta Botanica Mexicana* **14**: 3-21. DOI: <https://doi.org/10.21829/abm14.1991.611>
- Sánchez-García EA, López-Rosas H, Infante-Mata D, Moreno-Casasola P. 2026. *Annona glabra* L., a key swamp tree: Ecological, functional, and ethnobotanical aspects for conservation and restoration. *Botanical Sciences*. DOI: <https://doi.org/10.17129/botsoci.3740>
- Schatz GE. 2025. Annonaceae. In: Ulloa Ulloa C, Hernández HM, Barrie FR, Davidse G, Knapp S, eds. *Flora Mesoamericana* (Vol. 2, Part 1, pp. [I]-XXV, 1-583). St. Louis: Missouri Botanical Garden.
- Scrucca L, Fraley C, Murphy TB, Raftery AE. 2023. *Model-based clustering, classification, and density estimation using mclust in R*. Broken Sound Pkwy NW: Chapman and Hall/CRC. ISBN 978-1032234953. DOI: <https://doi.org/10.1201/9781003277965>

- Smith SA, O'Meara BC. 2012. treePL: Divergence time estimation using penalized likelihood for large phylogenies. *Bioinformatics* **28**: 2689-2690. DOI: <https://doi.org/10.1093/bioinformatics/bts492>
- Smith SA, Brown JW, Walker JF. 2018. So many genes, so little time: A practical approach to divergence-time estimation in the genomic era. *Plos One* **13**: e0197433. DOI: <https://doi.org/10.1371/journal.pone.0197433>
- Sosa V, De-Nova JA, Vásquez-Cruz M. 2018. Evolutionary history of the flora of Mexico: Dry forests cradles and museums of endemism. *Journal of Systematics and Evolution* **56**: 523-536. DOI: <https://doi.org/10.1111/jse.12416>
- Soulé VRC, Couvreur TLP, Mariac C. 2023. Annonaceae DNA extraction protocol from silicagel dried and herbarium preserved leaves, v1. *Protocols.io*. <https://doi.org/10.17504/protocols.io.5qpvorqx9v4o/v1> (accessed July 28, 2024).
- Soulé VRC, Mariac C, Delaigue T, Couvreur TLP. 2024. Illumina library preparation and dual hybridization protocol of ERC GLOBAL V.3, v3. *Protocols.io*. <https://doi.org/10.17504/protocols.io.n92ldp78715b/v3>
- Stamatakis A. 2014. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**: 131-1313. DOI: <https://doi.org/10.1093/bioinformatics/btu033>
- Tong Y, Binford G, Rheims CA, Kuntner M, Liu J, Agnarsson I. 2019. Huntsmen of the Caribbean: Multiple tests of the GAARlandia hypothesis. *Molecular Phylogenetics and Evolution* **130**: 259-268. DOI: <https://doi.org/10.1016/j.ympev.2018.09.017>
- Upham NS, Esselstyn JA, Jetz W. 2019. Inferring the mammal tree: Species-level sets of phylogenies for questions in ecology, evolution, and conservation. *Plos Biology* **17**: e3000494. DOI: <https://doi.org/10.1371/journal.pbio.3000494>
- Vázquez-López M, Córtes-Rodríguez N, Robles-Bello SM, Bueno-Hernández A, Zamudio-Beltrán LE, Ruegg K, Hernández-Baños BE. 2021. Phylogeography and morphometric variation in the Cinnamon Hummingbird complex: *Amazilia rutila* (Aves: Trochilidae). *Avian Research* **12**: 61. DOI: <https://doi.org/10.1186/s40657-021-00295-0>
- Webb SD. 2006. The Great American Biotic Interchange: Patterns and processes. *Annals of the Missouri Botanical Garden* **93**: 245-257.
- Wendt T. 1993. Composition, floristic affinities, and origins of the canopy tree flora of the Mexican Atlantic slope rain forests. In: Ramamoorthy TP, Bye R, Lot A, Fa J, eds. *Biological Diversity of Mexico* (pp. 595-680). New York: Oxford University Press.
- Wickham H. 2016. *ggplot2: Elegant graphics for data analysis*. Springer-Verlag, New York.
- Wickham H, Henry L. 2023. *purrr: Functional programming tools*. R package version 1.0.2. <https://CRAN.R-project.org/package=purrr> (accessed May 2026)
- Wickham H, François R, Henry L, Müller K, Vaughan D. 2023. *dplyr: A grammar of data manipulation*. R package version 1.1.4. <https://CRAN.R-project.org/package=dplyr> (accessed May 2026)
- Wickham H, Vaughan D, Girlich M. 2024. *tidyr: Tidy messy data*. R package version 1.3.1. <https://CRAN.R-project.org/package=tidyr> (accessed May 2026)
- Zhang C., Rabiee M, Sayyari E, Mirarab S. 2018. ASTRAL-III: Polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinformatics* **19**: 16-30. <https://doi.org/10.1186/s12859-018-2129-y>

Associate editor: German Carnevali Fernández Concha

Author contributions: AEOR conceptualization, investigation, database construction, formal analyses, draft writing, revision and final version; LH database construction, formal analyses, revision and final version; LHR database construction, revision and final version; FJN database construction, formal analyses, revision and final version; YLV formal analyses, revision and final version; TLPC conceptualization, formal analyses, revision and final version.

Supporting Agencies: T.L.P. Couvreur acknowledges funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (Grant Agreement No. 865787).

Conflict of interests: The authors declare that there is no conflict of interest, financial or personal, in the information, presentation of data and results of this article.