UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE CIÊNCIAS BIOLÓGICAS E DA SAÚDE
DEPARTAMENTO DE GENÉTICA E EVOLUÇÃO
PROGRAMA DE PÓS GRADUAÇÃO EM GENÉTICA EVOLUTIVA E BIOLOGIA MOLECULAR

MILENA CARDOSO TELHE

FILOGENIA, BIOGEOGRAFIA, ECOLOGIA E EVOLUÇÃO DE CARACTERES DOS GÊNEROS *MELOCACTUS* LINK & OTTO E *DISCOCACTUS* PFEIFF. (CACTACEAE).

SÃO CARLOS -SP
2025

MILENA CARDOSO TELHE

FILOGENIA, BIOGEOGRAFIA, ECOLOGIA E EVOLUÇÃO DE CARACTERES DOS GÊNEROS *MELOCACTUS* LINK & OTTO E *DISCOCACTUS* PFEIFF. (CACTACEAE).

Tese apresentada ao Programa de Pós-Graduação em Genética Evolutiva e Biologia Molecular, Universidade Federal de São Carlos, *campus* São Carlos, como parte dos requisitos para obtenção do título de doutor em Ciências.

Orientador: Dr. Evandro Marsola de Moraes

Financiamento: Fapesp (2019/11233-2)

SÃO CARLOS -SP
2025

Este trabalho foi realizado com auxílio financeiro da Fundação de Amparo a Pesquisa do estado de São Paulo (FAPESP) através dos processos 2019/11233-2 e 2022/11683-0.

Dedico a minha mãe, pai, irmã e companheiro, meus maiores incentivadores e exemplos de resiliência.

“Understand life is not a conquest, but a slow lesson” - The plausibility of life: Resolving Darwin's Dilemma by Marc W. Kirschner et al.

Agradecimentos

Agradeço,

A Universidade Federal de São Carlos (UFSCar) e o Programa de Pós Graduação em Genética Evolutiva e Biologia Molecular (PPGGEV), e todo seu corpo docente, técnicos, e secretarias, pela infraestrutura, ensinamentos e apoio, durante todo período do meu doutorado.

À Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) pela concessão da bolsa que permitiu a realização desse projeto de doutorado.

Ao Prof. Dr. Evandro Marsola de Moraes pela confiança, orientação, paciência, e ensinamentos que mudaram para sempre a profissional que sou e busco ser.

Ao Prof. Dr. Fernando Faria Franco, que mesmo não sendo meu orientador, sempre incentiva a mim e a todos os alunos do laboratório, além de sempre estar disposto a ajudar e nos ensinar. Obrigada por todas as discussões enriquecedoras.

Ao Dr. Nigel P. Taylor e à Profa. Dra. Daniela C. Zappi pelo auxílio e colaboração como taxonomistas especialistas em Cactaceae, trazendo insights importantíssimos e discussões enriquecedoras. A paixão de vocês pelos cactos e pela ciência me inspira, e me faz relembrar sempre o porquê escolhi o caminho da ciência.

A Profa. Dra. Thais N. C. Vasconcelos por todo ensinamento, apoio, incentivo, e acolhimento. O período em que trabalhamos juntas mudou pra sempre minha vida profissional, voltei com a bagagem cheia de novos conhecimentos acadêmicos, mas também com um novo horizonte sobre a possibilidade de uma academia mais igualitária, justa, muitas vezes intensa, mas sempre respeitosa.

Ao Gerardus Olsthoorn por gentilmente ter cedido inúmeras amostras de seu acervo pessoal, o que foi crucial para o desenvolvimento do presente trabalho. Além de compartilhar seus conhecimentos, trazendo questões e discussões importantíssimas sobre os cactos.

Ao Dr. Marlon Machado e Dr. Diego Gonzaga por ceder amostras de seu acervo pessoal, e do Jardim Botânico do Rio de Janeiro, respectivamente. Essas amostras foram de extrema importância para completar a amostragem de espécies extra-brasileiras do presente trabalho.

Ao Prof. Dr. Luciano A. Digiampietri por compartilhar seus conhecimentos e auxiliar nas abordagens de aprendizado de máquina apresentadas no presente trabalho.

À Dra. Juliana Martinez, e a todo o corpo técnico que passou pelo laboratório por seu apoio prático, teórico, e também pela amizade.

Aos meus queridos amigos e amigas que fizeram e fazem parte do Laboratório de Genética Evolutiva (LaGEvol), por todo conhecimento compartilhado, incentivo, risadas, cafés, e também apoio nos momentos desafiadores.

Aos meus queridos amigos Dra. Monique Romeiro-Brito, Profa. Dra. Isabel A. S. Bonatelli, e Prof. Dr. Danilo T. Amaral pela amizade, colaboração, incentivo, scripts e conhecimento compartilhados. Vocês são pra mim exemplos de vida pessoal e profissional. É imensa a gratidão que tenho por vocês.

A minha família de Sorocaba que o LaGEvol me presentou, a Msc. Juliana R, Bombonato, Msc. Fernanda M. Leite, Msc. Heidi S. M. Utsonomyia, Dr. Antonio M. C. Neto, Wesley Matheus, Jefferson. Vocês, juntamente com a Dra. Monique Romeiro-Brito, Profa. Dra. Isabel A. S. Bonatelli, e Prof. Dr. Danilo T. Amaral, se tornaram minha família, e todo o apoio e incentivo foram essenciais durante o doutorado. Minha gratidão e amor por vocês é enorme.

Ao meu querido amigo Dr. Antonio M. C. Neto, por toda amizade, incentivo, discussões científicas e filosóficas, e também pelos puxões de orelha quando eu precisei. Você será pra sempre um exemplo de amigo e cientista.

Ao meu melhor amigo e companheiro, Leonardo R. Pereira, por escolher dividir a vida comigo, compartilhando as alegrias, comemorando as conquistas, e sendo apoio e incentivo nos momentos difíceis. Sou também eternamente grata pelas comidinhas que apareciam magicamente na minha mesa durante as longas horas de trabalho.

A minha sobrinha, que mesmo tão pequena tem o poder de escancarar a realidade, e me lembrar diariamente o que realmente importa nessa vida. Sendo fonte de força e amor para que eu chegasse até aqui.

À minha mãe Sueli Cardoso, pai Luiz Telhe, e irmã Alana Telhe, pelo apoio e incentivo incondicionais. Desde a minha infância, vocês são meus exemplos e grandes incentivadores por essa busca pelo conhecimento.

Foram muitas as pessoas e instituições que colaboraram direta ou indiretamente para o desenvolvimento e conclusão deste doutorado. Peço desculpas àqueles que porventura eu não tenha mencionado. Faltam páginas e palavras para agradecer cada um de vocês.

À todos, o meu sincero e eterno,
Obrigada!

RESUMO GERAL

Apesar de bastante conhecida devido a sua riqueza de espécies, uma questão ainda pouco discutida é o papel dos intercâmbios bióticos para a manutenção da diversidade na região Neotropical. Eventos de intercâmbios de biota entre habitats distintos foram identificados em grande número nas regiões Neotropicais, indicando um papel fundamental para o estabelecimento das biotas regionais. No entanto, estudos envolvendo o papel dos intercâmbios bióticos sobre as alterações nas taxas de diversificação são escassos na região Neotropical, sobretudo entre ambientes secos e abertos nessa região. Embora algumas hipóteses foram propostas para explicar o favorecimento e as consequências evolutivas da conexão entre esses ambientes, poucos estudos foram dedicados a testá-las. Dessa forma, mais estudos são necessários para melhor compreender a dispersão e conexão entre os biomas secos e abertos da região Neotropical. O presente estudo tem como objetivo investigar eventos biogeográficos e demográficos dos gêneros *Melocactus* e *Discocactus*, contribuindo para o entendimento dos eventos que impulsionaram a diversificação nos biomas secos e abertos da região Neotropical. Para tanto, foi utilizado o painel de sondas Cactaceae591, para a obtenção de inúmeras regiões ortólogas utilizando o método de captura de sequência. Essas sequências foram utilizadas para realizar inferências filogenéticas de ambos os gêneros, estimar datas de divergencia, inferir eventos biogeograficos, inferir a reconstrução de caracteres ancestrais e de nicho filogenético. Além disso, foram estimadas taxas de diversificação e investigada sua associação com variáveis bióticas e abióticas. De forma resumida, as centenas de regiões sequenciadas a partir do método de captura utilizando o painel Cactaceae591 foram capazes de inferir de forma robusta as relações entre e dentro de ambos os gêneros, permitindo rearranjos taxonômicos, como demonstrado nos capitulos 1 e 2. O período de diversificação de ambos os gêneros coincide com um periodo de mudanças climáticas (Quaternário), alternando entre períodos de maior aridez e períodos mais húmidos. Como consequencia das mudanças climáticas, nesse período também ocorreram mudanças na distribuição dos ambientes, permitindo, por exemplo, a expansão dos ambientes áridos durante os períodos climáticos favoraveis, e promovendo maior heterogeneidade edáfica ao longo da distribuição desses ambientes. A linhagem ancestral de ambos os gêneros tiverem origem em um bioma composto por Florestas Sazonalmente Secas, posteriormente experienciando eventos de dispersão para novas áreas, como por exemplo as Savanas. Não foram observadas alterações significativas nas taxas de diversificação em nenhum dos gêneros. Os eventos de mudança de bioma, e a presença de características adaptativas não estão associados com diferenças entre as taxas de diversificação das linhagens,

como demonstrado no capitulo 3. O presente estudo sugere que as mudanças climáticas do passado, e características ecológicas em uma escala mais fina, como por exemplo a heterogeneidade edáfica, desempenharam um papel fundamental para a diversificação de *Melocactus* e *Discocactus*.

Palavras chave: Cactaceae591, especialização edáfica, trait adaptativo, mudança de bioma, filogenômica, aprendizagem de máquina.

GENERAL ABSTRACT

Despite being well known due to its species richness, a question that has not been deeply discussed is the role of biotic exchanges in maintaining diversity in the Neotropical region. Biota exchange events between distinct habitats have been identified in large numbers in the Neotropical regions, indicating a fundamental role in establishing regional biotas. However, studies involving the role of biotic exchanges on changes in diversification rates are scarce in the Neotropical region, especially between dry and open environments in this region. Although some hypotheses have been proposed to explain the evolutionary consequences of the connection between these environments, few studies have been dedicated to testing them. Thus, more studies are needed to better understand the dispersal and connection between the dry and open biomes of the Neotropical region. The present study aims to investigate the systematic and ecological relationships of *Melocactus* and *Discocactus*. Additionally, we investigated biogeographic events within the *Melocactus* and *Discocactus* genera, contributing to the understanding of how these events drove diversification in the dry and open biomes of the Neotropical region. For this purpose, the Cactaceae591 probe panel was used to obtain numerous orthologous regions using the sequence capture method. These sequences were used to perform phylogenetic inferences of both genera, estimate diversification ages, infer biogeographic events, and infer the reconstruction of ancestral characters and phylogenetic climatic niches. In addition, diversification rates were estimated to investigate their association with biotic and abiotic variables. In summary, the hundreds of regions sequenced using the Cactaceae591 probe set were able to robustly infer the relationships between and within both genera, allowing taxonomic rearrangements, as demonstrated in chapters 1 and 2. The diversification period of both genera coincides with a period of climate change (Quaternary), alternating between periods of greater aridity and more humid periods. As a consequence of climate change, changes in the distribution of environments also occurred during this period, allowing, for example, the expansion of arid environments during favorable climatic periods, and promoting greater soil heterogeneity throughout the distribution of these environments. The ancestral lineage of both genera originated in a biome composed of Seasonally Dry Forests, later experiencing dispersal events to new areas, such as Savannas. No significant changes in diversification rates were observed in either genus. Biome change events and the presence of adaptive traits are not associated with differences in the diversification rates of the lineages, as demonstrated in Chapter 3. The present study suggests that past climate change and ecological

characteristics on a finer scale, such as soil heterogeneity, may have played an important role in the *Melocactus* and *Discocactus* diversification.

Keywords: Cactaceae591, soil specialization, adaptive trait, biome shifts, phylogenomics, machine learning.

FIGURE LIST

General Introduction

Chapter 1

Chapter 2

Chapter 3

TABLE LIST

Chapter 1

Chapter 2

Chapter 3

General introduction

1.1 The role of biotic interchange between Neotropical dry and open habitats

The Neotropical region is well known for its high species diversity and high levels of endemism, housing approximately 37% of all species in the world (Antonelli & Sanmartín, 2011). This region ranges from southern South America to central Mexico, encompassing diverse biomes, from moist to dry (Antonelli & Sanmartín, 2011). In this region, climatic and geological events have been proposed as primary agents of taxon diversification (Rull, 2011). Moreover, intrinsic characteristics of taxon life histories also seem to play a fundamental role in diversification rates in this region (Antonelli et al., 2018). More recently, biotic interchanges have also been postulated as important to the diversity found in this region (Antonelli et al., 2018). Several lineage dispersal events between biomes (biotic interchange) have been identified in the Neotropical region (Ledo & Colli, 2017), with lineage age in the original biome being the most consistent predictor of dispersal and colonization events (Antonelli et al., 2018). These events are suggested to be linked to changes in diversification rates and to taxon diversity within a region, as a result of a greater ecological opportunity to colonize, adapt, and diversify (Pichardo-Marcano et al., 2019; Freitas et al., 2019; Koenen et al., 2013). Therefore, studying diversification and connectivity among Neotropical biomes is crucial for understanding the origin and maintenance of the diversity found in this region. However, studies investigating the role of biotic interchanges explicitly testing shifts in diversification rates in the Neotropical region are scarce (Antonelli et al., 2018; Souza-Neto et al., 2016; Simon et al., 2009; Becerra & Venable, 2008; Nieto-Blázquez et al., 2017; Serrano et al., 2024), especially between dry and open environments (Santos et al., 2011).

In the Neotropical region, biomes composed of dry and open vegetation are characterized by highly seasonal climates with low precipitation (Hughes et al., 2013). Even though they are all seasonal, they significantly differ in time and intensity of this seasonality and temperature gradient, ranging from arid and semi-arid to subhumid, sometimes all three occurring in a contiguous area (e.g., *Dry Diagonal of South America*; Ab'saber, 1977; Luebert, 2021). Moreover, they differ in more intrinsic aspects (see Hughes et al., 2013 and references therein), such as soil composition (e.g., Seasonally Dry Tropical Forests - fertile soil; Savannas - poorer soil), and other environmental filters, such as the presence of natural fire events (e.g., Savannas).

Seasonal Dry Tropical Forest (SDTF) – forests with fertile soils and usually a more or less continuous canopy, mostly composed of deciduous woody vegetation and succulent species, with a marked dry season.
Savannas – found on poorer soils, under similar or slightly wetter climatic conditions, with sclerophyllous trees, and having a xeromorphic, fire-tolerant grass as an important component.

Studies within Neotropical dry and open biomes identified dispersal events connecting them, in which all biomes have acted as both source and sink of lineages (Antonelli et al., 2018; Souza-Neto et al., 2016; Simon et al., 2009; Amaral et al, 2021). The biotic interchange between these regions, with different environmental filters, is evidenced by shared characteristics among sister lineages, such as fire adaptations in lineages from Seasonally Dry Tropical Forests (SDTFs) and Savannas (Souza-Neto et al., 2016; Simon et al., 2009). These characteristics could be pre-adaptations present in individuals that were able to establish in new environmental conditions or physiological adaptations due to new selective pressures faced in this new environment (Wiens et al., 2010; Simon et al., 2009). In that way, incorporating biological characteristics that may have allowed species to be established in different habitats, can help to better understand the role of the biotic interchange between biomes in species diversification (Wiens et al., 2010; Simon et al., 2009).

A potential biological model for studies of the Neotropical dry and open habitats is the Cactaceae family. Adapted to thrive in low-water environments, this family is a species-rich and notable element of Neotropical dry and open habitats (Arakaki et al., 2011). Cacti genera from the Cereeae tribe (Romeiro-Brito et al, 2023) are mostly endemic to South America, with the majority of the diversity of this tribe found in a large contiguous area known as eastern South America Dry Diagonal (*sensu* Luebert, 2021) and dry enclaves within Atlantic Forest (Taylor & Zappi, 2004), and Amazon (Devecchi et al., 2020). Within this tribe we can mention the sister genera (Silva et al., 2018; Barthlott et al., 2015; Hernandéz-Hernandéz et al., 2011; RITZ et al., 2007) *Melocactus* Link & Otto (38 species; *sensu* Hunt, 2006; and Flora e Funga do Brasil, 2025) and *Discocactus* Pfeiff. (15 species; *sensu* Hunt, 2006; Flora e Funga do Brasil, 2025). Both genera exhibit a globose-discoid shape, a habit suggested to provide greater resistance to fire (Taylor & Zappi, 2004). Most of the taxonomic diversity of *Melocactus* and *Discocactus* occurs in the Seasonal Dry Tropical Forest (SDTF) and Savanna, respectively, with some species of both genera also occurring in transitional areas (Figure 1; Taylor & Zappi, 2004). Considering *Melocactus* and *Discocactus* distribution, the presence of a shared characteristic related to the colonization of a fire-prone habitats such as the Savanna, and their sister relationship, these genera are good models for studying the processes of dispersion and connection between dry and open habitats, such as SDTFs and Savannas biomes (sensu Dinerstein et al., 2017).

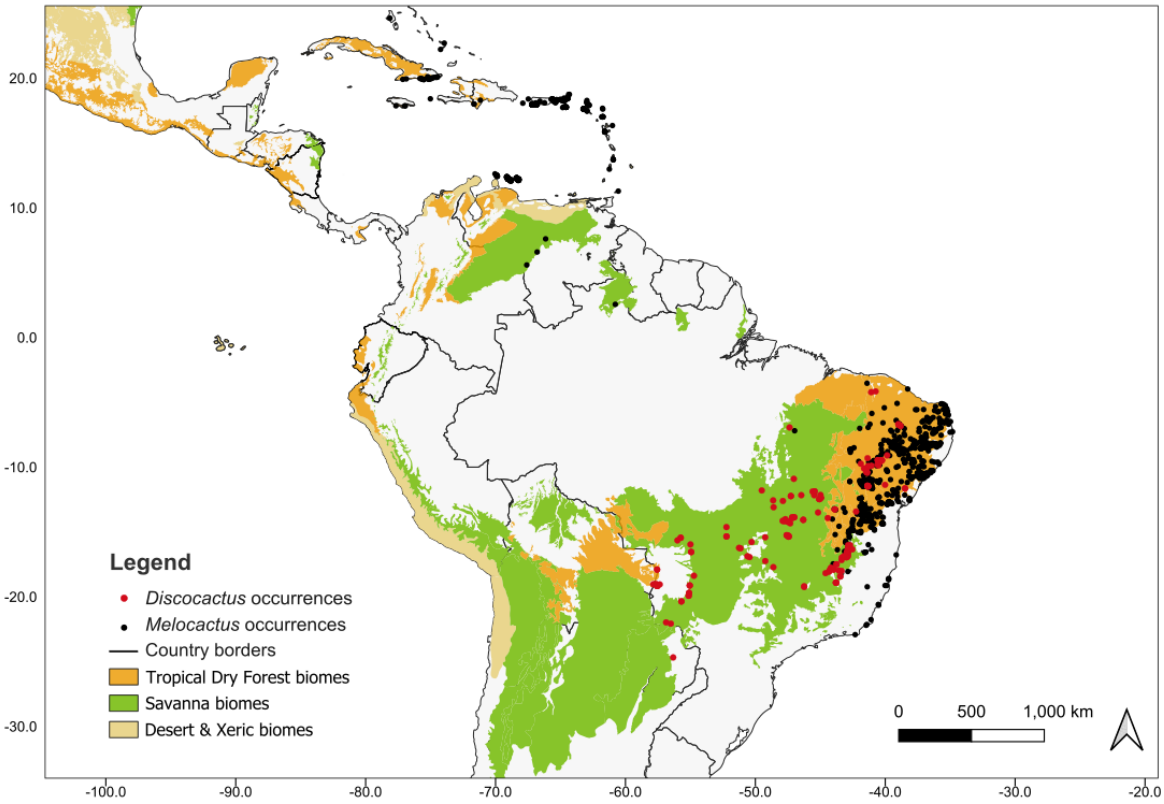


Figure 1. Distribution of *Melocactus* and *Discocactus* species across the Neotropical biomes delimited by Dinerstein et al. (2017). The map was produced on QGIS v.3.32 by plotting the geographic coordinates of species extracted from GBIF and specieslink.

Although they may represent informative models, molecular studies addressing the diversity and phylogenetic relationship of *Melocactus* and *Discocactus* lineages are scarce (Santos et al., 2013; Machado, 2005; Majure et al, 2022). Until now, neither genus has a robust phylogenetic hypothesis, especially with a comprehensive taxon sampling. This is mainly due to the challenge of obtaining a robust phylogeny for taxa with rapid and recent diversification, like the Cactaceae family (Copetti et al., 2017; Franco et al, 2022). In this scenario, a few genetic differences may have accumulated, leading to low or even absence of phylogenetic signals in the molecular markers (Small et al., 1998), and consequently, unresolved phylogenies (Calvente et al., 2017; Townsend, 2007). In addition to the rapid and recent diversification, the frequent hybridization found in Cactaceae (Taylor & Zappi, 2004; Tel-Zur et al., 2004; Colaço et al., 2006; Nyffeler & Eglli, 2010) and incomplete lineage sorting (Copetti et al., 2017) can produce misleading phylogenies due to incongruent evolutionary histories among gene trees, and between gene trees and species tree (Maddison, 1997). One way to overcome these challenges is to use multiple loci associated with methods based on the multispecies coalescent

model (MSC; Mccormack et al., 2013). The use of multiple loci allows for greater sampling of the genome and incorporating the expected genealogical heterogeneity among loci (Brito, & Edwards, 2009). The MSC method also accounts for gene tree heterogeneity and gene tree/species tree discrepancies enabling the inference of phylogenetic trees with greater accuracy and support among branches (Degnan & Rosenberg 2009).

In the past few years, there has been an increase in genomic-scale studies within the Cactaceae family, as next-generation sequencing techniques have become more affordable for non-model organisms (Franco et al., 2022). In this context, different sequencing strategies have been explored for phylogenomic studies in cacti (e.g. reduced genome representation - Bombonato et al., 2020; whole-genome - Amaral et al., 2021a; transcriptome Walker et al., 2018). Recently, the target sequencing by hybridization capture method has been successfully applied to plant groups with a challenging taxonomical and evolutionary history (e.g., radiations; Carter et al., 2019; Loiseau et al., 2019; Nicol et al., 2024), like Cactaceae, as a cost-effective enrichment method for sequencing hundreds of putatively single-copy nuclear regions (Mamanova et al., 2010). Among cacti, the newly developed probe set, Cactaceae591, which targets both the coding and noncoding sequences of 591 putative orthologous regions within the subfamily Cactoideae (Romeiro-Brito et al., 2022), was able to resolve multiple taxonomy levels of phylogenetic relationships (e.g. deep level - Romeiro-Brito et al., 2023; and shallow level - Taylor et al., 2023), proven to be an effective tool for phylogenetic inferences within this subfamily.

Considering the challenges faced for the inference of *Melocactus* and *Discocactus* phylogenies, obtaining multiple loci through Next Generation Sequencing (NGS), such as target sequence capture, should allow for robust phylogenetic inferences, an essential step to biogeographical and comparative phylogenetics approaches. Hence, the use of the Cactaceae591 probe set, may not only enable us to unravel *Melocactus* and *Discocactus* phylogenetic relationships, but also to answer questions regarding their origin and diversification processes, such as: To what extent does the data generated contribute to taxonomic discussions and rearrangements within these genera? When and where did these genera originate? Did the dispersion events of these genera mainly occur from SDTFs (e.g., Caatinga domain) to Savannas (e.g., Cerrado domain), or did they occur in the opposite

direction? Which biogeographical events, biotic and abiotic factors may have favored the dispersion and diversification of these genera in Neotropical dry and open habitats?

1.2 The main of this thesis

The main purpose of this study is to investigate the role of biotic interchanges between Neotropical arid and semi-arid biomes on species diversification, based on phylogenetic comparative analyses between *Melocactus* and *Discocactus* genera. Seeking to contribute to the understanding of the role of these events in the diversification of lineages from the Neotropical dry and open biomes.

To achieve this aim, specific goals were set:

(i) Obtaining large-scale molecular data through target sequencing capture methodology for reliably inferring the phylogenetic relationships within *Melocactus* and *Discocactus* (Chapters 1 and 2);
(ii) Phylogenetic inference of *Melocactus* and *Discocactus*, comparing the results of conventional probabilistic methods and methods based on coalescence (Chapters 1 and 2);
(iii) Estimating the diversification time of *Melocactus* and *Discocactus* genera (Chapters 1, 2, and 3);
(iv) Estimating the ancestral distribution area and changes in the distribution of the main lineages of each genus (Chapter 3);
(v) Investigating the association between shifts in diversification rates within these genera and biotic and abiotic factors using phylogenetic comparative methods (Chapter 3).

1.3 References

Ab'Saber, A. N. (1977). Os domínios morfo-climáticos na América do Sul: primeira aproximação Geomorfológica. *São Paulo: IG/USP*.

Amaral, D. T., Bombonato, J. R., da Silva Andrade, S. C., Moraes, E. M., & Franco, F. F. (2021a). The genome of a thorny species: comparative genomic analysis among South and North American Cactaceae. *Planta*, *254*(3), 44.

Amaral, D. T., Minhós-Yano, I., Oliveira, J. V. M., Romeiro-Brito, M., Bonatelli, I. A. S., Taylor, N. P., ... & Franco, F. F. (2021). Tracking the xeric biomes of South America: The spatiotemporal diversification of Mandacaru cactus. *Journal of Biogeography*, *48*(12), 3085-3103.

Antonelli, A., & Sanmartín, I. (2011). Why are there so many plant species in the Neotropics? *Taxon*, *60*(2), 403-414.

Antonelli, A., Zizka, A., Carvalho, F. A., Scharn, R., Bacon, C. D., Silvestro, D., & Condamine, F. L. (2018). Amazonia is the primary source of Neotropical biodiversity. *Proceedings of the National Academy of Sciences*, *115*(23), 6034-6039.

Arakaki, M., Christin, P. A., Nyffeler, R., Lendel, A., Eggli, U., Ogburn, R. M., ... & Edwards, E. J. (2011). Contemporaneous and recent radiations of the world's major succulent plant lineages. *Proceedings of the National Academy of Sciences*, *108*(20), 8379-8384.

Barthlott, W. (2015). *Biogeography and biodiversity of cacti*. Univ.-Verlag Isensee.

Becerra, J. X., & Venable, D. L. (2008). Sources and sinks of diversification and conservation priorities for the Mexican tropical dry forest. *PLos one*, *3*(10), e3436.

Brito, P. H., & Edwards, S. V. (2009). Multilocus phylogeography and phylogenetics using sequence-based markers. *Genetica*, *135*, 439-455.

Calvente, A., Moraes, E. M., Lavor, P., Bonatelli, I. A., Nacaguma, P., Versieux, L. M., ... & Zappi, D. C. (2017). Phylogenetic analyses of *Pilosocereus* (Cactaceae) inferred from plastid and nuclear sequences. *Botanical Journal of the Linnean Society*, *183*(1), 25-38.

Carter, K. A., Liston, A., Bassil, N. V., Alice, L. A., Bushakra, J. M., Sutherland, B. L., ... & Hummer, K. E. (2019). Target capture sequencing unravels *Rubus* evolution. Frontiers in plant science, 10, 1615.

Colaço, M. A., Fonseca, R., Lambert, S. M., Costa, C. B., Machado, C. G., & Borba, E. L. (2006). Biologia reprodutiva de *Melocactus glaucescens* Buining & Brederoo e *M. paucispinus* G. Heimen & R. Paul (Cactaceae), na Chapada Diamantina, Nordeste do Brasil. *Brazilian Journal of Botany*, *29*, 239-249.

Copetti, D., Búrquez, A., Bustamante, E., Charboneau, J. L., Childs, K. L., Eguiarte, L. E., ... & Sanderson, M. J. (2017). Extensive gene tree discordance and hemiplasy shaped the genomes of North American columnar cacti. *Proceedings of the National Academy of Sciences*, *114*(45), 12003-12008.

Degnan, J. H., & Rosenberg, N. A. (2009). Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends in ecology & evolution*, *24*(6), 332-340.

Devecchi, M. F., Lovo, J., Moro, M. F., Andrino, C. O., Barbosa-Silva, R. G., Viana, P. L., ... & Zappi, D. C. (2020). Beyond forests in the Amazon: biogeography and floristic relationships of the Amazonian savannas. *Botanical Journal of the Linnean Society*, *193*(4), 478-503.

Flora do Brasil. (2020). Flora do Brasil. Jardim Botânico do Rio de Janeiro. Available online: http://floradobrasil.jbrj.gov.br/ (Accessed, June 20th, 2024).

Franco, F. F., Amaral, D. T., Bonatelli, I. A., Romeiro-Brito, M., Telhe, M. C., & Moraes, E. M. (2022). Evolutionary genetics of cacti: research biases, advances and prospects. *Genes*, *13*(3), 452.

Freitas, C. G., Bacon, C. D., Souza-Neto, A. C., & Collevatti, R. G. (2019). Adjacency and area explain species bioregional shifts in neotropical palms. *Frontiers in Plant Science*, *10*, 55.

Hernández-Hernández, T., Hernández, H. M., De-Nova, J. A., Puente, R., Eguiarte, L. E., & Magallón, S. (2011). Phylogenetic relationships and evolution of growth form in Cactaceae (Caryophyllales, Eudicotyledoneae). *American journal of botany*, *98*(1), 44-61.

Hughes, C. E., Pennington, R. T., & Antonelli, A. (2013). Neotropical plant evolution: assembling the big picture. *Botanical Journal of the Linnean Society*, *171*(1), 1-18.

Hunt, D. (2006). The new cactus lexicon. descriptions & illustrations of the cactus family. Compiles and edited by members or the International Cactaceae Systematic Group. England.(Ed.). 373 p.

Koenen, E. J. M., De Vos, J. M., Atchison, G. W., Simon, M. F., Schrire, B. D., De Souza, E. R., ... & Hughes, C. E. (2013). Exploring the tempo of species diversification in legumes. *South African Journal of Botany*, *89*, 19-30.

Ledo, R. M. D., & Colli, G. R. (2017). The historical connections between the Amazon and the Atlantic Forest revisited. *Journal of biogeography*, *44*(11), 2551-2563.

Loiseau, O., Olivares, I., Paris, M., de La Harpe, M., Weigand, A., Koubínová, D., ... & Salamin, N. (2019). Targeted capture of hundreds of nuclear genes unravels phylogenetic relationships of the diverse Neotropical palm tribe Geonomateae. Frontiers in plant science, 10, 864.

Luebert, F. (2021). The two South American dry diagonals. *Frontiers of Biogeography*, *13*(4).

Machado, M. C. (2005). O gênero *Discocactus* Pfeiff.(Cactaceae) no estado da Bahia, Brasil: variabilidade morfológica, variabilidade genética, taxonomia e conservação. *Unpublished MSc Thesis, Universidade Estadual de Feira de Santana, Bahia*.

Maddison, W. P. (1997). Gene trees in species trees. *Systematic biology*, *46*(3), 523-536.

Majure, L. C., Barrios, D., Díaz, E., Bacci, L. F., & Piñeyro, Y. E. (2022). Phylogenomics of the Caribbean melocacti: Cryptic species and multiple invasions. *Taxon*, *71*(5), 993-1012.

Mamanova, L., Coffey, A. J., Scott, C. E., Kozarewa, I., Turner, E. H., Kumar, A., ... & Turner, D. J. (2010). Target-enrichment strategies for next-generation sequencing. *Nature methods*, *7*(2), 111-118.

McCormack, J. E., Hird, S. M., Zellmer, A. J., Carstens, B. C., & Brumfield, R. T. (2013). Applications of next-generation sequencing to phylogeography and phylogenetics. *Molecular phylogenetics and evolution*, *66*(2), 526-538.

Neves, D. M., Dexter, K. G., Pennington, R. T., Bueno, M. L., & Oliveira Filho, A. T. (2015). Environmental and historical controls of floristic composition across the South American Dry Diagonal. *Journal of Biogeography*, *42*(8), 1566-1576.

Nicol, D. A., Saldivia, P., Summerfield, T. C., Heads, M., Lord, J. M., Khaing, E. P., & Larcombe, M. J. (2024). Phylogenomics and morphology of Celmisiinae (Asteraceae: Astereae): Taxonomic and evolutionary implications. Molecular Phylogenetics and Evolution, 195, 108064.

Nieto-Blázquez, M. E., Antonelli, A., & Roncal, J. (2017). Historical Biogeography of endemic seed plant genera in the Caribbean: Did GAAR landia play a role?. *Ecology and Evolution*, *7*(23), 10158-10174.

Nyffeler, R., & Eggli, U. (2010). A farewell to dated ideas and concepts: molecular phylogenetics and a revised suprageneric classification of the family Cactaceae. *Schumannia*, *6*, 109-149.

Pichardo-Marcano, F. J., Nieto-Blázquez, M. E., MacDonald, A. N., Galeano, G., & Roncal, J. (2019). Phylogeny, historical biogeography and diversification rates in an economically important group of Neotropical palms: Tribe Euterpeae. *Molecular phylogenetics and evolution*, *133*, 67-81.

Porzecanski, A. L., & Cracraft, J. (2005). Cladistic analysis of distributions and endemism (CADE): using raw distributions of birds to unravel the biogeography of the South American aridlands. *Journal of Biogeography*, *32*(2), 261-275.

Ritz, C. M., Martins, L., Mecklenburg, R., Goremykin, V., & Hellwig, F. H. (2007). The molecular phylogeny of *Rebutia* (Cactaceae) and its allies demonstrates the influence of paleogeography on the evolution of South American mountain cacti. *American journal of Botany*, *94*(8), 1321-1332.

Romeiro-Brito, M., Taylor, N. P., Zappi, D. C., Telhe, M. C., Franco, F. F., & Moraes, E. M. (2023). Unravelling phylogenetic relationships of the tribe Cereeae using target enrichment sequencing. *Annals of Botany*, *132*(5), 989-1006.

Romeiro-Brito, M., Telhe, M. C., Amaral, D. T., Franco, F. F., & Moraes, E. M. (2022). A target capture probe set useful for deep-and shallow-level phylogenetic studies in Cactaceae. *Genes*, *13*(4), 707.

Rull, V. (2011). Neotropical biodiversity: timing and potential drivers. *Trends in ecology & evolution*, *26*(10), 508-513.

Santos, J. C., Leal, I. R., Almeida-Cortez, J. S., Fernandes, G. W., & Tabarelli, M. (2011). Caatinga: the scientific negligence experienced by a dry tropical forest. *Tropical Conservation Science*, *4*(3), 276-286.

Santos, M. R., Garcia, F. C. P., Machado, M. C., Barbosa, A. R., Taylor, N. P., & Van Den Berg, C. (2013). A phylogenetic study of *Discocactus* Pfeiff.(Cactaceae) based on nuclear and plastid DNA sequences. *Filogenia molecular, taxonomia, biogeografia e conservação de Discocactus Pfeiff. PhD thesis–Universidade Federal de Viçosa, Viçosa, Brazil*, 13-36.

Serrano, F. C., Pontes-Nogueira, M., Sawaya, R. J., Alencar, L. R., Nogueira, C. C., & Grazziotin, F. G. (2024). There and back again: when and how the world's richest snake family (Dipsadidae) dispersed and speciated across the Neotropical region. *Journal of Biogeography*, *51*(5), 878-893

Silva, G. A. R., Antonelli, A., Lendel, A., Moraes, E. D. M., & Manfrin, M. H. (2018). The impact of early Quaternary climate change on the diversification and population dynamics of a South American cactus species. *Journal of Biogeography*, *45*(1), 76-88.

Simon, M. F., Grether, R., de Queiroz, L. P., Skema, C., Pennington, R. T., & Hughes, C. E. (2009). Recent assembly of the Cerrado, a neotropical plant diversity hotspot, by in situ evolution of adaptations to fire. *Proceedings of the National Academy of Sciences*, *106*(48), 20359-20364.

Small, R. L., Ryburn, J. A., Cronn, R. C., Seelanan, T., & Wendel, J. F. (1998). The tortoise and the hare: choosing between noncoding plastome and nuclear Adh sequences for phylogeny reconstruction in a recently diverged plant group. *American journal of botany*, *85*(9), 1301-1315.

Souza-Neto, A. C., Cianciaruso, M. V., & Collevatti, R. G. (2016). Habitat shifts shaping the diversity of a biodiversity hotspot through time: insights from the phylogenetic

structure of Caesalpinioideae in the Brazilian Cerrado. *Journal of Biogeography*, *43*(2), 340-350.

Taylor, N. P., & Zappi, D. (2004). Cacti of eastern Brazil Royal Botanic Gardens. *Kew. UK*.

Taylor, N. P., Zappi, D. C., Romeiro-Brito, M., Telhe, M. C., Franco, F. F., & Moraes, E. M. (2023). A phylogeny of *Cereus* (Cactaceae) and the placement and description of two new species. *Taxon*, *72*(6), 1321-1333.

Tel-Zur, N., Abbo, S., Bar-Zvi, D., & Mizrahi, Y. (2004). Genetic relationships among *Hylocereus* and *Selenicereus* vine cacti (Cactaceae): Evidence from hybridization and cytological studies. *Annals of Botany*, *94*(4), 527-534.

Townsend, J. P. (2007). Profiling phylogenetic informativeness. *Systematic biology*, *56*(2), 222-231.

Walker, J. F., Yang, Y., Feng, T., Timoneda, A., Mikenas, J., Hutchison, V., ... & Smith, S. A. (2018). From cacti to carnivores: Improved phylotranscriptomic sampling and hierarchical homology inference provide further insight into the evolution of Caryophyllales. *American Journal of Botany*, *105*(3), 446-462.

Werneck, F. P., Gamble, T., Colli, G. R., Rodrigues, M. T., & Sites, Jr, J. W. (2012). Deep diversification and long-term persistence in the South American 'dry diagonal': integrating continent-wide phylogeography and distribution modeling of geckos. *Evolution*, *66*(10), 3014-3034.

Wiens, J. J., & Graham, C. H. (2005). Niche conservatism: integrating evolution, ecology, and conservation biology. *Annu. Rev. Ecol. Evol. Syst.*, *36*(1), 519-539.

Zanella, F. C. (2002). Systematics and biogeography of the bee genus *Caenonomada* Ashmead, 1899 (Hymenoptera: Apidae: Tapinotaspidini). *Studies on Neotropical Fauna and Environment*, *37*(3), 249-261.

CHAPTER 1: Phylogenomic and ecological systematics of *Melocactus* (Cactaceae)

Milena C. Telhe[1], Nigel P. Taylor[1], Monique Romeiro-Brito[2], Daniela C. Zappi[3], Gerardus Olsthoorn[4], Fernando F. Franco[1], Evandro M. Moraes[1].

1. Departamento de Biologia. Universidade Federal de São Carlos, Sorocaba, São Paulo 18052-780, Brazil.
2. Natural History Department, Florida Museum of Natural History, Gainesville, FL 32611, USA.
3. Programa de Pós-Graduação em Botânica, Instituto de Ciências Biológicas, Universidade de Brasília (UNB), Brasília 70919-970, Brazil.
4. Community scientist, Caixa Postal 169, Holambra, São Paulo, 13825-000, Brazil

Accepted for publication at Systematic Botany 2025.

Abstract

Melocactus (L.) Link & Otto (Cactaceae) is a widely distributed genus in the Neotropical region and has fascinated the botanical community since the 15th century due to its unique appearance. Even though this genus has been studied for centuries, its diversification, phylogenetic relationships, and species delimitation have not been evaluated using comprehensive taxon sampling and genetic data. Here, we inferred maximum likelihood and coalescent phylogenies using the Cactaceae591 genomic dataset and investigated potential environmental variables associated with *Melocactus* diversification. Ancestral character reconstructions were performed using edaphic and climatic data. A well-resolved *Melocactus* phylogeny was estimated, allowing the redefinition of informal infrageneric groups and the taxonomic rearrangement of some taxa. The ancestral character reconstructions suggest that the observed species relationships and distribution patterns in *Melocactus* result from the interplay of climatic and edaphic factors, highlighting a complex evolutionary history for the genus.

Keywords—Cactaceae591, Caribbean radiations, cryptic species, edaphic specialization, genomics.

2.1 Introduction

Melocactus (L.) Link & Otto is a genus of Cactaceae comprising c. 38 species (sensu Hunt, 2006; Flora e Funga do Brasil, 2025) widely distributed along the Neotropical region. This genus ranges from eastern Brazil up to the Amazonian Llanos, Central America (including coastal Mexico), central and northern Andes, and the Caribbean region (Taylor & Zappi 2004), with two centers of diversity along its distribution (Caribbean region - c. 7 species, and Brazil - c. 24 species; sensu Hunt et al. 2006; Flora do Brasil 2020; but see Barrios et al. 2023). *Melocactus* belongs to the South American Cactoideae clade (tribe Cereeae), and it is the second most diverse genus in the subtribe Cereinae (Romeiro-Brito et al. 2023b). The species of this genus do not go unnoticed in the field due to their peculiar adult appearance. *Melocactus* is one of the two globose cactus genera in the Cactaceae with a terminal cephalium - a wolly and bristly columnar fertile extension developed on top of the vegetative stem (Fig. 1). These characteristics and its diversity have long fascinated the botanical community, making *Melocactus* one of the earliest cactus genera documented by naturalists in the 15th century.

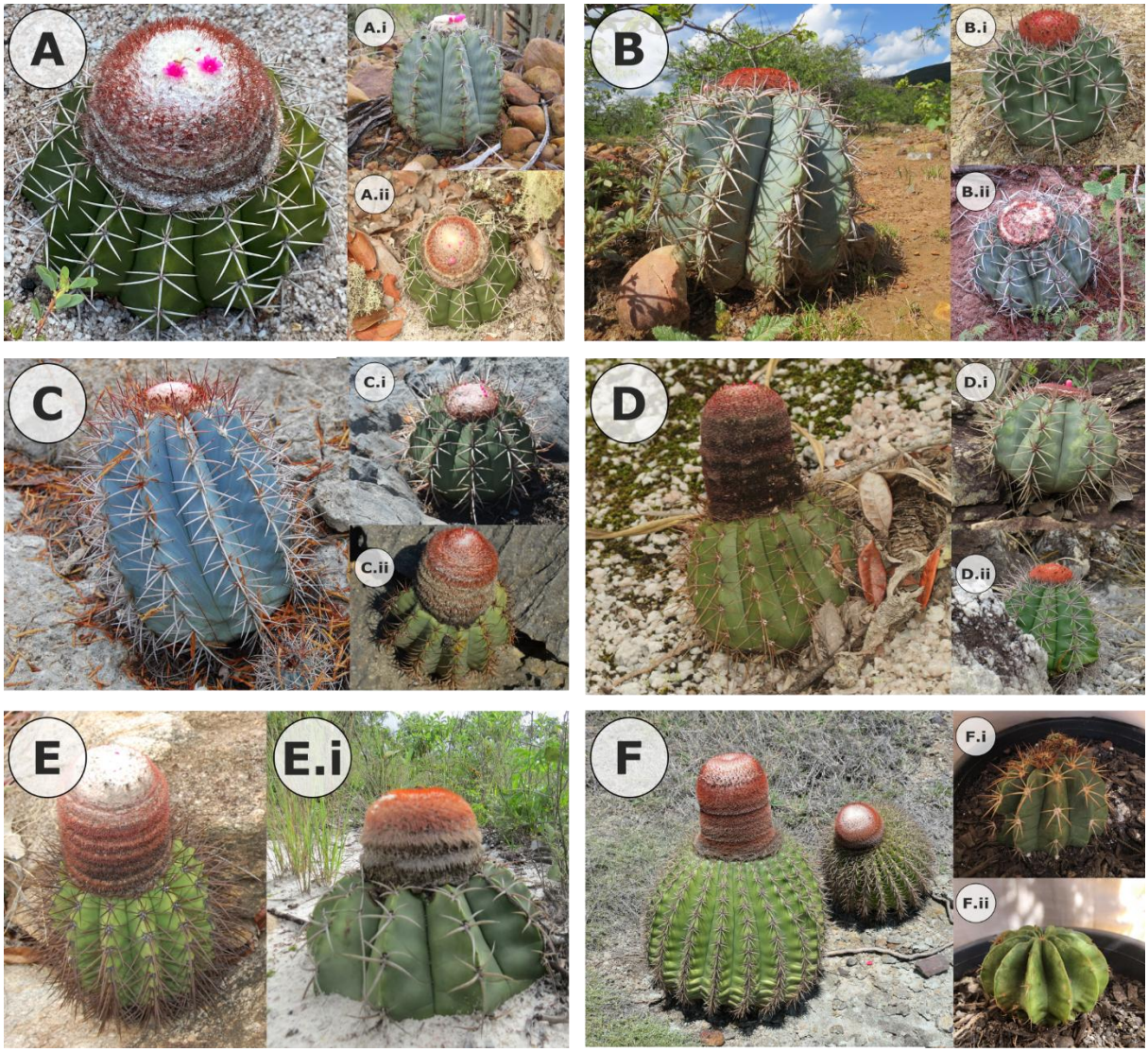


Fig. 1 Representatives of the genus *Melocactus* and its clades. (a) *M. violaceus* illustrating the M. VIOLACEUS clade (a.i) *M. glaucescens*, and (a.ii) *M. depressus*. (b) *M. zehntneri* illustrating M. ZEHNTNERI clade (b.i) *M. salvadorensis*, and (b.ii) *M. concinnus*. (c) *M. azureus* illustrating M. AZUREUS clade (c.i) *M. sergipensis*, and (c.ii) *M. levitestatus*. (d) *M. oreas* illustrating M. OREAS clade (d.i) *M. amethystinus*, and (d.ii) *M. bahiensis*. (e) *M. deinacanthus* illustrating M. DEINACANTHUS clade (e.i) *M. paucispinus*. (f) *M. intortus* illustrating M. INTORTUS clade (f.i) a juvenile *M. curvispinus*, and (f.ii) a juvenile *M. schatzlii*. Photo credits: a – Caetano, A. H.; d.i - Zappi, D.; f – Gdaniec, A.; b.ii - Moraes, E. M.; a.ii and d– Albuquerque-Lima, S.; c, c.i, c.ii, E - Olsthoorn, G.; a.i, b, b.i, d.i, e.i, f.i, and f.ii – Telhe, M. C.

The first naturalists who studied species belonging to the genus attributed suggestive names, such as *Echinomelocactus* (Tabernaemontanusi 1591), and *Cactus melocactus*, the last one described by Linnaeus (1753) in his famous work *Species Plantarum*. A year later, Linnaeus (1754) published his *Genera Plantarum*, in which the genus *Cactus* L. included an unranked infrageneric group called *Melocactus*. Later, Link & Otto (1827) formally described the genus, adopting the name *Melocactus* and circumscribing four named species (one referable to *Discocactus* Pfeiff.*)*. After Link & Otto (1827), several studies regarding *Melocactus* were published (e.g., Miquél 1840; Suringar 1896; Britton & Rose 1922; Werdermann 1934), however, these seminal studies usually lacked clear indications of species relationships. Taylor (1991) proposed the most recent systematic treatment of *Melocactus*, which included an informal infrageneric classification of species groups and evolutionary hypotheses for the genus. In his study, Taylor recognized 31 species, arranging them into six species groups delimited by a few gross-morphological and seed characters. Moreover, Taylor (1991) noted that the species from each group usually share a similar ecology, occurring in similar habitats or substrates (e.g., species endemic to geologically distinct rocky outcrops). Although *Melocactus* is a conspicuous cactus, the last in-depth formal inquiry into its systematics did not include phylogenetic hypotheses to explain species limits and relationships. Consequently, the diversification processes within *Melocactus* and its species' relationships remain poorly understood.

The lack of resolution in phylogenies based on few traditional molecular markers hindered earliest attempts to estimate the species relationships within *Melocactus* using genetic data (Marlon Machado unpubl. data), a typical pattern found within the family Cactaceae (see Franco et al. 2022 and references therein). After the advent of genomic sequencing techniques, well-resolved species-level phylogenies have become feasible for cacti lineages (e.g., Bombonato et al. 2020; Merklinger et al. 2021; Romeiro-Brito et al. 2023b). Recently, Majure et al. (2022) proposed a phylogenetic hypothesis for *Melocactus* based on plastome data. However, in their study, the phylogenetic placement of the clade representing the main diversity center, and therefore, most of its taxonomic diversity and geographic range, lacks statistical support. Additionally, their sampling primarily focused on the Caribbean species, including only four of the 24 species from the primary center of diversity of the genus in Brazil. Since two centers of diversity are recognized for the genus, and only one has been extensively sampled, a significant portion of *Melocactus* diversity remains unexamined.

The most robust phylogenetic hypotheses regarding *Melocactus* involve its closely related lineage *Discocactus* Pfeiff., which is consistently recovered as a sister genus in

broadscale phylogenetic studies of Cactaceae (Hernández-Hernández et al. 2011; Silva et al. 2018; Romeiro-Brito et al. 2023b), with a Plio-Pleistocene origin. This timing suggests that climatic oscillations may have contributed to their diversification (Hernández-Hernández et al. 2014; Silva et al. 2018). However, no studies have yet investigated which abiotic factors, such as climate or substrate conditions, may have contributed to their diversification.

In Cactaceae, lineage diversification and distribution are generally associated with local environmental features, particularly temperature and substrate conditions (Hernandez & Barcenas 1995; Godinez-Alvarez, Valverde & Ortega-Baes 2003; Ruedas, Valverde & Zavala-Hurtado 2006; Amaral et al. 2022; Thompson et al. 2024). Therefore, a focused investigation into the role of these abiotic factors in *Melocactus* diversification could provide valuable insights into the evolutionary history of the genus and clarify the influence of environmental conditions on species boundaries within cacti.

Here, we aim to estimate the phylogenetic relationship of species within *Melocactus* using genomic data and improving the taxonomic sampling from eastern Brazil. Moreover, we are interested in testing the monophyly of the informal infrageneric groups proposed by Taylor (1991) based on morphological and ecological characters. Additionally, we aim to estimate the divergence time among species and investigate which environmental variables might be associated with *Melocactus* diversification. To this end, we generated genomic data using the Cactaceae591 target sequence probe set (Romeiro-Brito et al. 2022) and integrated plastid sequence data from Majure et al. (2022) available in public databases to infer the phylogenetic hypothesis. Additionally, using a calibrated tree, we investigate which environmental variables could be associated with the evolutionary history of the genus by carrying out ancestral character state reconstructions based on climatic and substrate datasets.

2.2 Methods

2.2.1 Data Sampling, Sequencing, and Processing

We sampled 51 specimens representing 31 currently recognized species of the genus *Melocactus*, which accounts for 78% of the species as defined by Hunt et al. (2006) and the Flora e Funga do Brasil (2025). This sampling comprises 95% of the diversity found in eastern Brazil (the primary *Melocactus* center of diversity), and approximately 57% of the diversity in the Caribbean region, the secondary center of diversity. Our assessment of the diversity within the Caribbean region relies on Hunt et al. (2006), who employs a more conservative taxonomic approach that aligns with our understanding of *Melocactus* species. However, we acknowledge

that other authors may have different classifications (e.g. González et al. 2016; Majure et al. 2022; Barrios et al., 2023). The outgroup comprises 14 genera from the subfamily Cactoideae and one species from the subfamily Opuntioideae (Table S1). Genomic DNA was extracted from root tissue (see Table S1) using a high-salt CTAB protocol (Romeiro-Brito et al. 2023b). The DNA was sent to RAPiD Genomics (Gainesville, Florida, USA), where enriched libraries with nuclear loci targeted by the Cactaceae591 probe set (Romeiro-Brito et al. 2022) and Illumina sequencing were performed. All new specimens used in this study were collected following Brazilian law through special permits provided to one of the authors of the current study (E.M.M.) by the Chico Mendes Biodiversity Conservation Institute (permit number 12014-1).

We removed the adapters, short reads (reads < 60bp), and low-quality reads (phred < 20) using *fastp* v. 0.23.3 (Chen et al. 2018). Subsequently, the reads were mapped and assembled using HybPiper v. 2.1.6 with the *--run_intronerate* function to recover supercontigs (Johnson et al. 2016). The reference sequences for the 591 nuclear regions available in Romeiro-Brito et al. (2022) were used for the data assembling. We identified and excluded putative paralog regions in our dataset, following the pipeline used by Romeiro-Brito et al. (2022). The multiple alignment was performed using MAFFT v. 7.520 (Katoh & Standley 2013). The poorly aligned sequences were trimmed using TrimAL v. 1.4 (Capella-Gutiérrez et al. 2009) with the function *-strict*, an automatic selection algorithm that combines a gappyout trimming with a subsequent trimming based on similarity threshold (automatically selected). An additional filter for hypervariable regions and outlier sequences was employed using *spruceup* with a lognormal distribution and a cutoff of 0.95 (Borowiec 2019). Aware of the effects of missing data in the subsequent analysis (Xi et al., 2016), our final nuclear dataset allowed only regions with up to 30% of missing samples. All raw reads were deposited in GenBank (Table S1).

To enhance our sampling and phylogenetic inferences, we reused raw genome skimming data available on the NCBI GenBank (BioProject PRJNA811606 and PRJNA715621) for 22 *Melocactus* specimens, along with nine species from the tribe Cereeae as the outgroup (Table S1). To integrate these data with the samples sequenced using the Cactaceae591 probe set, we assembled the plastid regions using the same pipeline described above, but in this case, using the plastid sequences available in Amaral et al. (2021) as reference. We also used the same protocols to sequence alignments and trimming as described previously (except for *spruceup*). The final plastid dataset allowed only regions with up to 20% of missing samples. To evaluate the informativeness of both datasets (plastid and nuclear),

standard indexes (e.g., number of variable sites) were calculated using AMAS v. 0.94. All datasets are available in Dryad (Moraes et al. Forthcoming 2024).

2.2.2 Phylogenetic inference and Divergence time estimation

For the plastid dataset, Maximum likelihood (ML) phylogenetic inference was performed using a concatenated matrix in IQ-TREE v. 2.2 (Minh et al. 2020) with the ModelFinder Plus option and ultrafast bootstrap (Hoang et al. 2018) with 10,000 replicates. For the nuclear dataset, phylogenetic inferences were carried out using two approaches: the ML on concatenated and single locus datasets, and the multispecies coalescent (MSC) on gene tree datasets. These maximum likelihood inferences were carried out following the same program and parameters described in the plastid inference above. The MSC approach was inferred using Weighted ASTRAL (wASTRAL; Zhang & Mirarab 2022) with maximum likelihood gene trees as input and considering bootstrap values under 70 as low node support. To reduce the impact of inconsistent quartets in the species tree recovery, we used the function *astral-hybrid*, which weights each gene tree quartet based on branch lengths and node supports (Zhang & Mirarab 2022). Additionally, branch support was calculated using the local posterior probability (LPP).

Divergence time was estimated using the relative rate framework implemented on the *RelTime* method (Tamura et al. 2012) available in MEGA11 (Tamura et al. 2021). Due to the lack of fossil records for the family Cactaceae, we performed a secondary calibration assuming the age interval previously estimated by Hernandez-Hernandez et al. (2014) to the subfamily Cactoideae (crown age: 10.94 – 21.85 Ma) with a normal distribution (mean: 16.13; SD: 2.90). We used all loci recovered for the final nuclear dataset and the phylogeny estimated by wASTRAL as input to the divergence time estimates.

2.2.3 Environmental data collection and phylogenetic signal

To characterize the environmental preferences of *Melocactus* lineages, we remotely collected climatic and substrate variables. The geographic occurrences for each species sampled in our phylogeny were retrieved from SpeciesLink (https://specieslink.net/) and GBIF (https://www.gbif.org) databases. To eliminate spatial errors, occurrence data were filtered using the CoordinateCleaner package (Zizka et al. 2019) in R v. 4.3.1 (R Core Team 2023). Then, we excluded any record outside of the known distribution of each species, plotting the occurrences and manually inspecting them in a Neotropical region map. Using the retrieved occurrences, we extracted 30s resolution climatic variables from the WorldClim2.1 (Fick &

Hijmans 2017) and Envirem (Title & Bemmels 2018) databases for each taxon. Next, the mean value of these variables for each taxon was used to test for a phylogenetic signal. For this purpose, we used the package phytools (Revell, 2012), with the function *phylosig* and Pagel's lambda estimations (Pagel 1999; Blomberg, Garland & Ives 2003), retaining in our dataset only the variables with a high phylogenetic signal (lambda > 0.7). Pagel's lambda statistic was applied because it tends to outperform other statistics, such as Blomberg et al.'s (2003) K, even in cases of incomplete phylogeny and suboptimal branch-length information (Molina-Venegas & Rodríguez 2017). We accessed the predominant substrate in which each species occurs (see Table S2) through literature reviews and taxonomic descriptions (e.g., Taylor 1991; Taylor & Zappi 2004; Hunt et al., 2006), as well as more than 35 years of field knowledge of each of the three cactus taxonomist authors (N.P.T; D.C.Z and G.O).

2.2.4 Ancestral character state estimation

To investigate which environmental variables may be associated with *Melocactus* diversification, we first performed a Spearman's correlation test between the climatic variables with a phylogenetic signal and retained those with a correlation coefficient lower than 0.7. To visualize the climatic variation along the *Melocactus* phylogeny, we reconstructed and mapped the non-correlated continuous climatic variables along the phylogeny using the *contmap* function in the phytools package (Revell 2012). We first selected the best model for categorical ancestral reconstruction to evaluate the discrete variation of the substrate type along the *Melocactus* phylogeny. We tested the *all different rate*s (ARD) and *equal rates* (ER) models, considering the substrate character as unordered, using the function *fitpolyMk* in the phytools package. We used log-likelihood to select the best-fit model. Next, we estimated the categorical evolution using the package corHMM (Boyko & Beaulieu 2022), allowing only one rate category (rate.cat = 1), independent estimation of transition rates (model = ARD), and unrestricted movement between the different types of substrates (allowed transitions from all states to any other state).

2.3 Results

2.3.1 Data sampling

Our final nuclear dataset (Cactaceae591: on-target + flanking regions) accounts for 51 *Melocactus* specimens comprising 482 loci, with 682,059 bp. The plastid dataset (54 plastid regions and 44,699 bp) accounts for only 32 *Melocactus* specimens due to the uneven sequence

recovery across our taxonomic sampling (Fig. S1). Both datasets also differ in the content of informativeness, with nuclear showing 142,296 variable sites, of which 58,426 are parsimoniously informative, while the plastid dataset shows 4,215 variable sites, of which 1,487 are parsimoniously informative.

2.3.2 Phylogenetic relationships and diversification ages

Both ML and MSC methods applied to the nuclear dataset recovered *Melocactus* as a monophyletic genus, sister to *Discocactus*, with well-supported relationships (Fig. 2). Both inferences recovered the Brazilian genus *Coleocephalocereus* Backeb. as a sister to the Melocactus-Discocactus clade. However, there were some inconsistencies between the methods. The main discrepancies involved the position of the clade containing *Melocactus depressus* Hook. and *Melocactus glaucescens* Buining & Brederoo. While the ML approach on the concatenated dataset recovered it as a sister to the M. ZEHNTNERI clade, though with low node support (UFBoot = 68), the MSC approach recovered it as a sister to the *Melocactus violaceus* species complex (LPP = 0.9). Another incongruence concerns the relationship between *Melocactus paucispinu*s Heimen & R.J. Paul and *Melocactus deinacanthus* Buining & Brederoo. The ML approach recovered *M. paucispinus* as a sister to *M. deinacanthus*-M. INTORTUS clade (UFBoot = 100). In contrast, the MSC approach recovered *M. paucispinus* and *M. deinacanthus* as sister lineages (LPP = 0.5), closely related to the M. INTORTUS clade (LPP = 0.88). Additionally, the position of *Melocactus bahiensis* (Britton & Rose) Luetzelb. differs depending on the method. While the ML approach recovered *M. bahiensis* nested within the subclade (a) M. bahiensis (M. OREAS clade; see Fig. 3 and Table S3), the MSC approach recovered *M. bahiensis* as a sister lineage of the same subclade. The names of clades and subclades proposed in this study (e.g. M. OREAS, M. DEINACANTHUS, M. INTORTUS, M. AZUREUS, M. ZEHNTNERI, and M. VIOLACEUS clades) are derived from the earliest properly named species within each group, reflecting the principle of priority in taxonomic nomenclature.

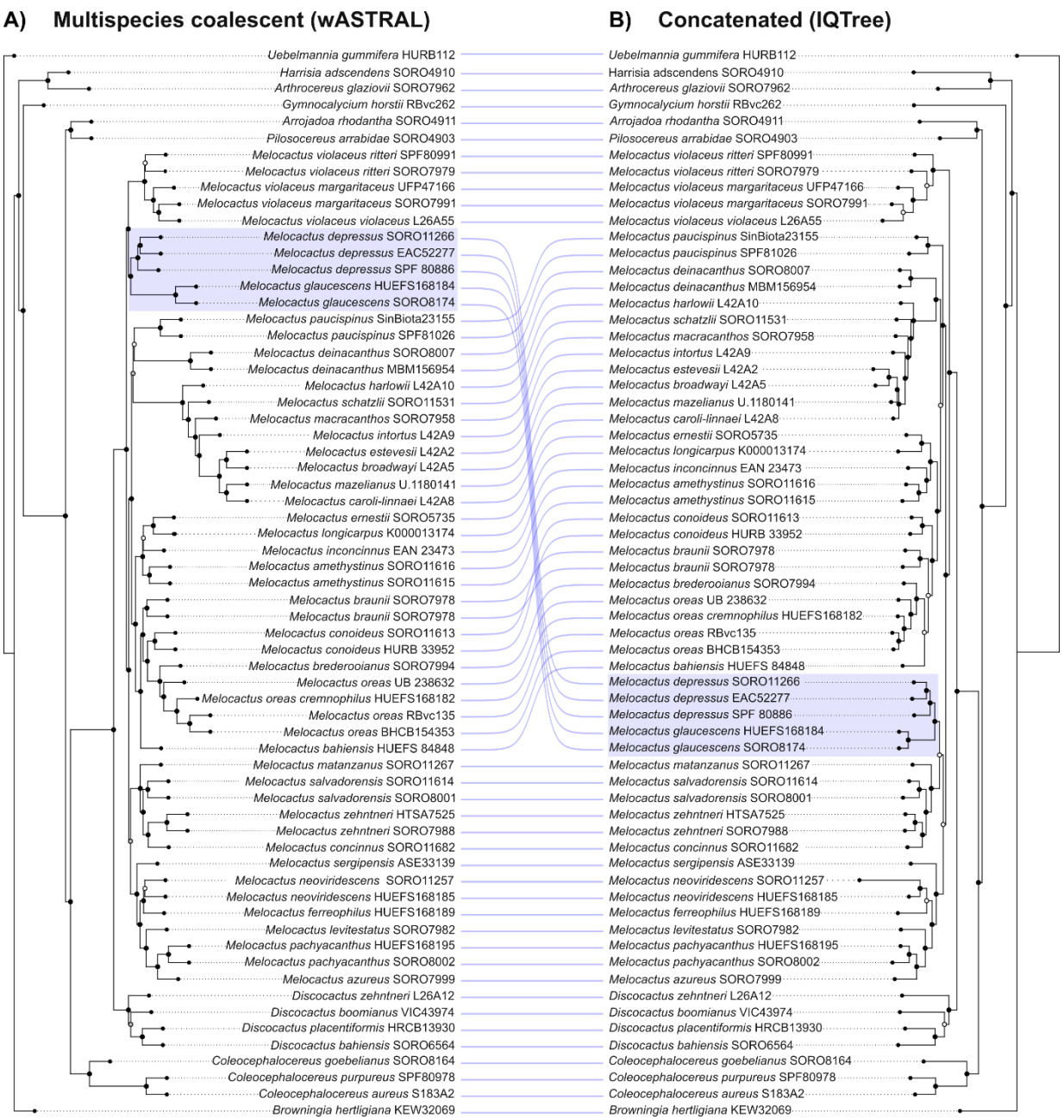


Fig. 2 Comparing *Melocactus* phylogenies inferred using **(a)** multispecies coalescent and **(b)** concatenated (maximum likelihood) approaches. The nodes were classified as highly-supported branches (LPP ≥ 0.95; UFBoot ≥ 95), moderately-supported (LPP 0.94 – 0.75), or low-supported (LPP < 0.75; UFBoot < 95), and are depicted as black, gray, and white circles, respectively. Incongruent portions between the trees are highlighted in light blue. The voucher number of each accession is provided after the species names.

Both methods consistently recovered the *Melocactus violaceus* complex as the sister group of the remaining lineages of the genus. While in the ML analysis, the *M. violaceus* complex exclusively forms a clade sister to all other lineages, the MSC method clusters the *M.*

violaceus complex with *M. glaucescens* and *M. depressus* in a clade sister to the rest of the genus. Additionally, all extra-Brazilian species, except for *Melocactus matanzanus* León, formed a well-supported and exclusive clade (LPP = 1.0; UFBoot = 100). The position of *M. matanzanus* was also congruent between methods, always recovering it as closely allied to the Brazilian species within the M. ZEHNTNERI clade (LPP = 0.99; UFBoot = 100). Regardless of the method, *M. depressus* was inferred as an independent evolving lineage, distinct from *Melocatus violaceus* Pfeiff. (LPP = 0.98; UFBoot = 100). Similarly, *Melocactus amethystinus* Buining & Brederoo was consistently inferred as an independent evolving lineage from *M. bahiensis sensu stricto* (LPP = 1.0; UFBoot = 100).

Our phylogenetic inferences partially supported the informal species groups proposed by Taylor (1991). Three of the six groups proposed (M. OREAS, M. DEINACANTHUS, and M. CURVISPINUS) were recovered as monophyletic, while the remaining groups were either partially or not supported as monophyletic. The M. AZUREUS Group was recovered as paraphyletic with the monotypic M. LEVITESTATUS Group nested within it. The M. VIOLACEUS Group was recovered as polyphyletic, as the species attributed to this group were placed in distinct lineages in the phylogeny *(M. paucispinu*s; M. ZEHNTERI clade; and M. VIOLACEUS clade).

The phylogenetic inferences based on the plastid dataset agree that the *M. violaceus* complex is the sister clade of all other *Melocactus* species (UFBoot = 100; Fig. S2). Additionally, *M. matanzanus* is still inferred as closely related to the Brazilian lineages, recovered within a clade composed of Brazilian species (UFBoot = 99). The clade comprising all extra-Brazilian species, except for *M. matanzanus*, is well-supported (UFBoot = 100). Most relationships within this clade were poorly supported with the plastid dataset (low support: UFBoot < 95), except for the subclade with *M. harlowii* Vaupel and its synonymous taxa. In this clade, *M. harlowii* and its conspecific taxa (e.g. *M. acunae* León, *M. lagunaensis* (Z.Mészáros) D.Barrios & Majure) are all gathered in a well-supported and exclusive subclade (UFBoot = 99). The *M. curvispinus* Pfeiff. and *M. curvispinus* subsp. *caesius* (H.L.Wendl.) N.P.Taylor were not recovered as closely related. However, the phylogenetic position of *M. curvispinus* lacked support (UFBoot = 57). Even though the plastid phylogeny recovered a backbone similar to the nuclear dataset inference, it was not feasible to test if the informal groups proposed for the genus would be recovered using this dataset due to the absence of most Brazilian taxa (c. 56% missing).

Divergence time estimation recovered the late Pliocene (mean: 2.90 million years ago [Ma], Fig. 3) as the divergence period between the genera *Melocactus* and *Discocactus*, with

the confidence interval extending up to the Pleistocene (CI 95%: 5.72 Ma–1.47 Ma; Table S4). The diversification within *Melocactus* began during the Pleistocene (mean: 2.39 Ma), with most of the lineages originating during the late Calabrian Age (Fig. 3).

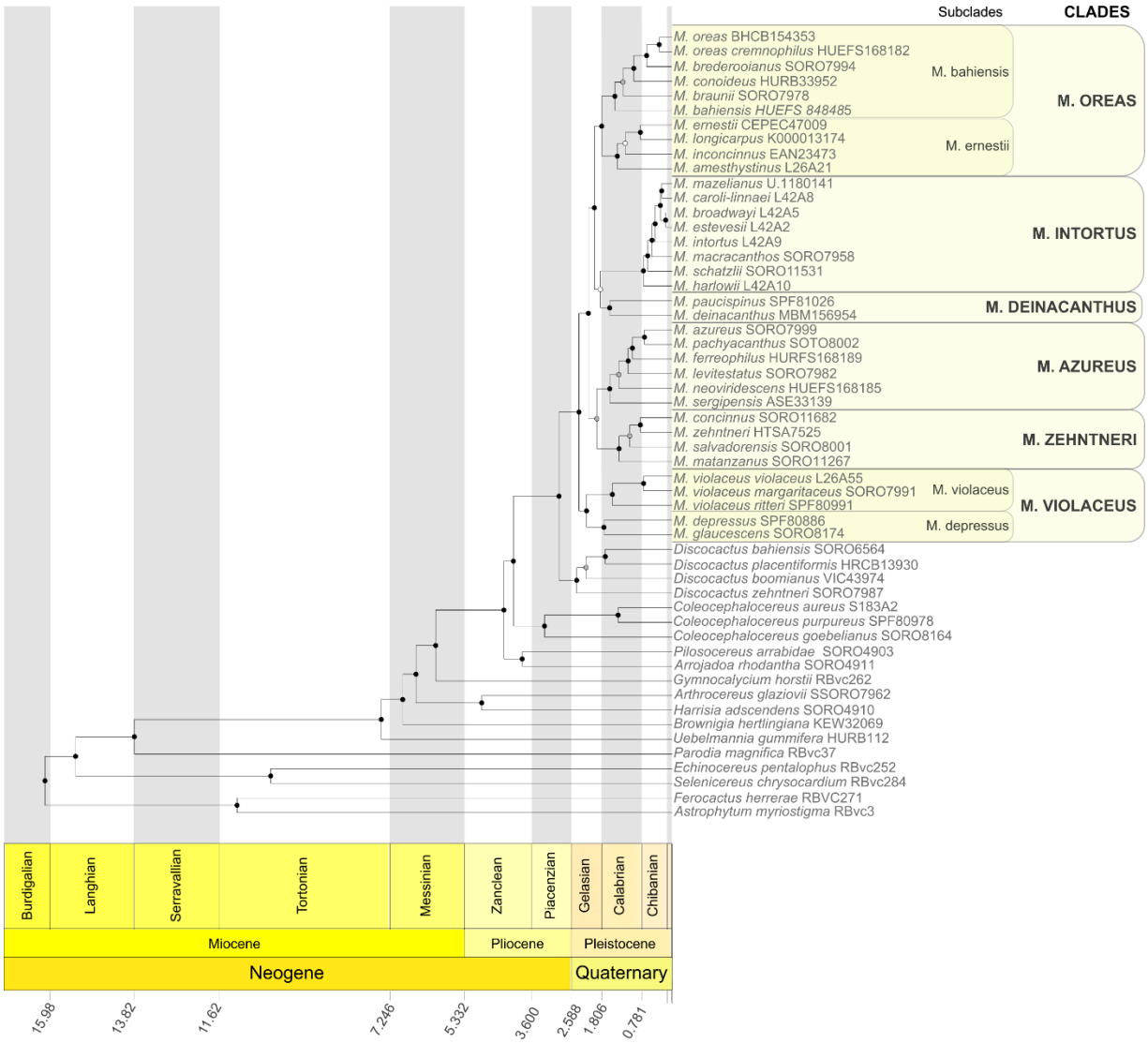


Fig. 3 Time-calibrated tree of *Melocactus* using the RelTime method in MEGA11. The highly supported branches (LPP ≥ 0.95) are depicted as black circles, moderately supported branches (LPP 0.94–0.75) are represented as gray circles, and low-supported branches (LPP < 0.75) are represented as white circles.

2.3.3 Ancestral environment reconstruction

We estimated statistically significant phylogenetic signals for five climatic variables (lambda > 0.7 and p < 0.05; Table S5). After the correlation test, only two variables were retained: maximum temperature of the coldest month (maxTempColdest; Table S6) and

potential evapotranspiration in the wettest quarter (PETwettestQuarter; Table S6). The continuous ancestral reconstruction for the retained variables is shown in Figure 4, in which most *Melocactus* lineages showed medium values of the maxTempColdest (Fig. 4.A). Nonetheless, the M. INTORTUS clade, along with *M. deinacanthus*, and *M. depressus* showed medium to higher values, while the M. OREAS clade, along with *M. pauscispinus*, *M. glaucescens*, and *M. violaceus* subsp. *ritteri* N.P.Taylor showed the lowest values. For the PETwettestQuarter, the ancestral reconstruction recovered medium values as the ancestral state for the genus (Fig. 4.B). However, shifts to higher values occurred during the diversification of most of the Brazilian species, with some clades showing the highest values. The first shift occurred during the diversification of the M. AZUREUS clade, except for *Melocactus sergipensis* N.P.Taylor & Meiado. The second shift was observed in the M. OREAS clade, and the third in the M. DEINACANTHUS clade (Fig. 4.B).

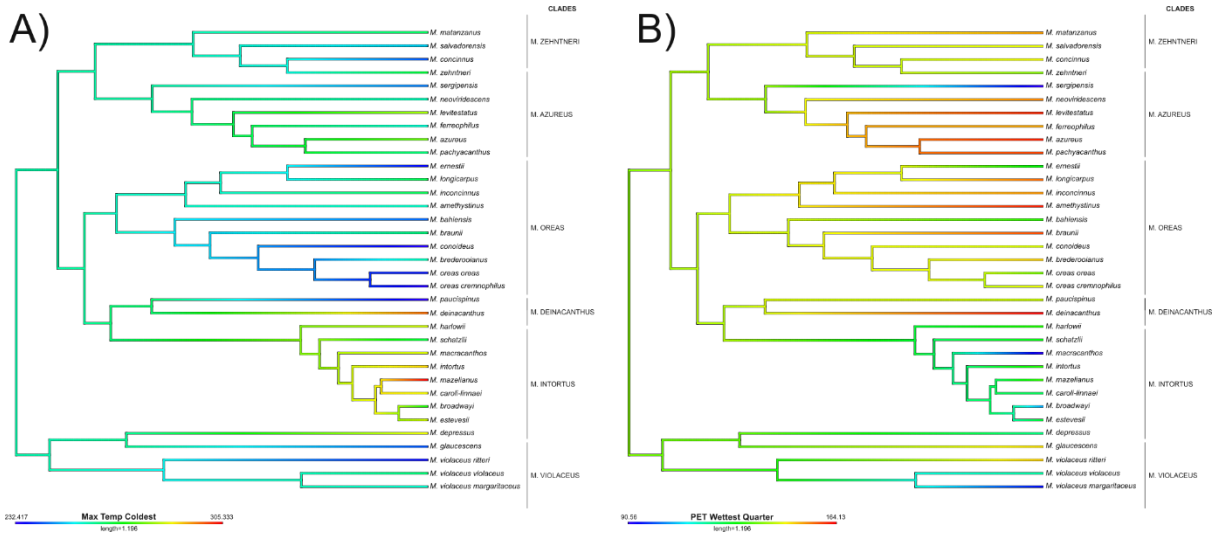


Fig. 4 Continuous ancestral character state reconstruction for two low-correlated climatic variables with significant phylogenetic signal. a) maximum temperature of the coldest month (maxTempColdest). b) potential evapotranspiration in the wettest quarter (PETwettestQuarter).

The discrete ancestral reconstruction for the substrate recovered sand as the ancestral edaphic state for the genus. The M. VIOLACEUS clade retained sand as the most probable ancestral state, while a change to quartz occurred in the ancestor of the remaining species (Fig. 5 and Table S7). Changes from quartz to granitic and from quartz to sandstone were detected in the clades M. INTORTUS and M. OREAS, respectively. The M. ZEHTNERI clade retained quartz as the ancestral state in its terminal lineages. In contrast, a state change from quartz to

limestone occurred in the M. AZUREUS clade, except for *M. sergipensis*, which recovered an inconclusive state. *M. sergipensis* node recovered uncertainties regarding the ancestral substrate, with mix-rocky being the likeliest ancestral state, albeit with a marginal probability of 0.32 (see Fig. 5; Table S7; and Fig. S3).

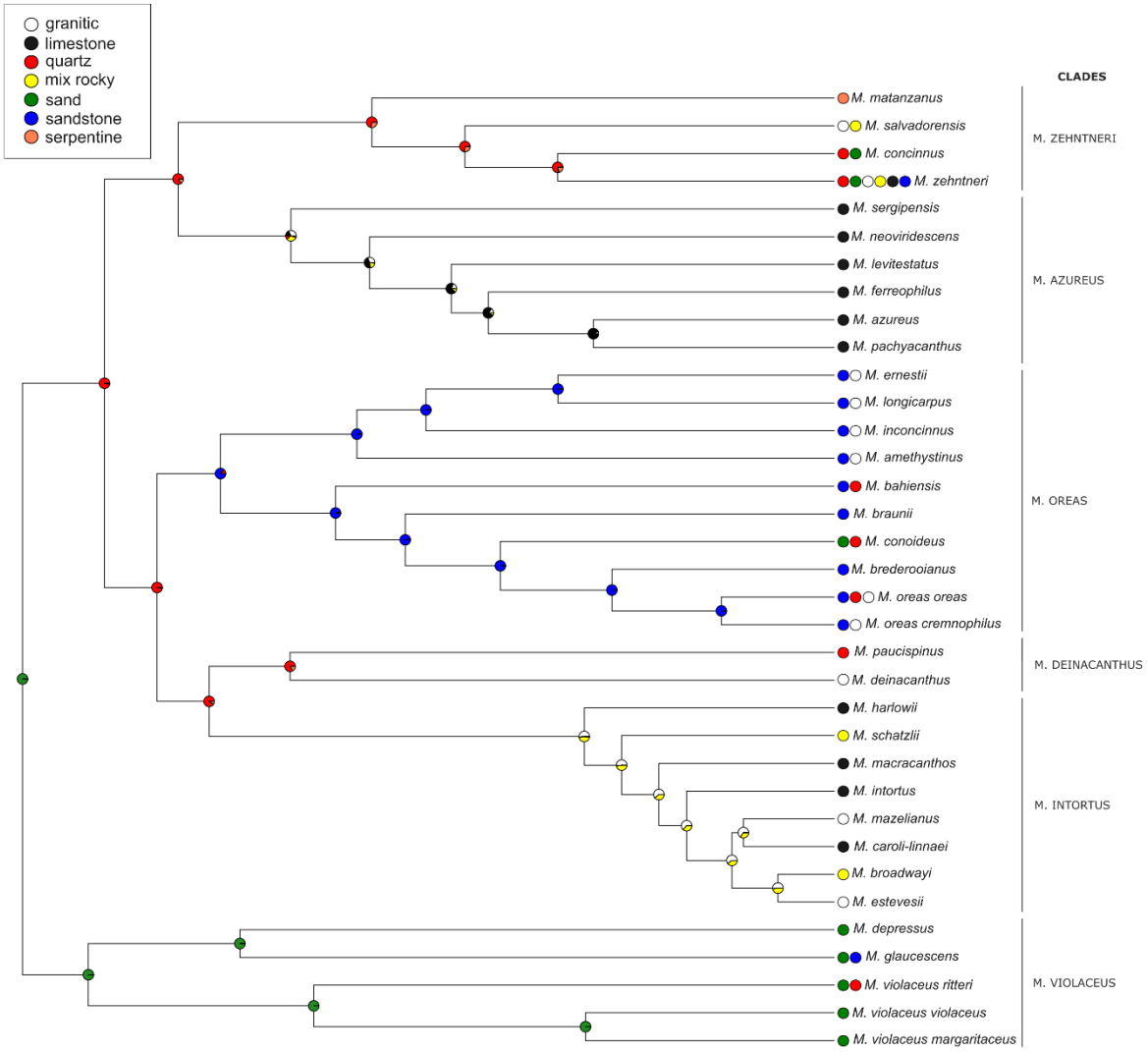


Fig. 5 Discrete ancestral character state reconstruction for substrate types in the genus *Melocactus.*

2.4 Discussion

2.4.1 Nuclear and plastid phylogenomic inferences of Melocactus

Both concatenated (maximum likelihood), and multispecies coalescent approaches applied to the nuclear dataset clarified the relationships within the genus *Melocactus* with high confidence despite some topological incongruences. Using both methods in genomic analyses

often results in incongruent topologies (e.g., Wu et al. 2018; Cai et al. 2021; Romeiro-Brito et al. 2022) due to different method assumptions and the complex evolutionary history of the genus. The concatenation method combines all genes into a supermatrix, assuming they share the same evolutionary history, which oversimplifies complex species diversification and disregards biological events that cause gene tree heterogeneity, such as hybridization, incomplete lineage sorting (ILS), and recombination (Maddison 1997; Edwards 2009; Edwards et al. 2016). As noted by Taylor (1991) and confirmed by Khan et al. (2020), hybridization has played a significant role in modern mixed *Melocactus* populations and should be considered when choosing phylogenetic inference methods. The conflicting phylogenetic position of *M. glaucescens* and *M. bahiensis* between concatenation and MSC inferences may reflect the challenges of the concatenation method in accounting for introgression events. Natural hybrids of *Melocactus glaucescens* (e.g. *M. glaucescens* × *M. ernestii*) in sympatric populations have been reported (Fonseca et al., 2008; Khan et al., 2020). Additionally, some populations of *M. bahiensis* are known to be tetraploids (Aguilera et al., 2014), which could point to ancient introgression (Stuessy & Weiss-Schneeweiss 2019). Aware of the hybridization occurrence in sympatric populations of *M. glaucescens,* we sampled only allopatric populations for this study. Future investigations are essential to better understand the impact of hybridization on the phylogenetic relationships of these species.

The MSC method infers the species phylogeny while accounting for gene tree heterogeneity and gene tree/species tree discrepancies (Edwards 2009; Edwards et al. 2016; Rannala et al. 2020). The debate over the most appropriate method for phylogenetic inferences, concatenated or coalescent, remains controversial (e.g., Kubatko & Degnan 2007; Gatesy & Springer 2014; Edwards et al. 2016; Scornavacca & Galtier 2017; Jiang, Edwards & Liu 2020). However, it has been shown that in scenarios with extensive ILS generating gene tree heterogeneity, such as in North American cactus species from the tribe Pachycereeae (Copetti et al. 2017), the MSC method and genomic dataset frequently outperforms the concatenation inference method (Jiang, Edwards & Liu 2020).

Our plastid dataset yielded valuable insights into *Melocactus* systematics, despite limited resolution within some clades. Notably, a well-supported clade clustering *M. harlowii* and *M. acunae* aligns with the taxonomic hypothesis proposed by Hunt et al. (2006), which synonymizes *M. acunae* with *M. harlowii*, rather than assuming they are distinct species. Besides, *M. curvispinus* and *M. curvispinus subsp. caesius* were not recovered as sister lineages (see Fig. S2), a pattern also obtained by Majure et al. (2022). These results could indicate that both lineages have independent evolutionary histories, and therefore, are two unrelated taxa.

Due to limited support in our results and those of Majure et al. (2022), along with the sampling of only two out of four subspecies, a further investigation accounting for all subspecies and their geographic distribution is essential for better understanding the systematics of the *M. curvispinus* complex. Despite the low recovery of Brazilian taxa, our plastid dataset yielded a phylogenetic backbone similar to those obtained from the nuclear dataset. In both datasets, *M. violaceus* was recovered as a phylogenetically isolated group, sister to the remaining diversity of the genus. At the same time, *M. matanzanus* was positioned alongside other Brazilian species. These results differ from those of Majure et al. (2022) using plastome data. While we recovered a Brazilian group of species as a sister to the remaining lineages of the genus with high support, Majure et al. (2022) recovered a clade comprising only endemic Cuban species as a sister to the rest of the genus. Although Majure et al. (2022) provide valuable information about these lineage relationships with extensive extra-Brazilian taxon sampling, their phylogeny contrasts with the clear picture of a South American origin of *Melocactus*, as well as most of the genera within the tribe Cereeae (Hernández-Hernández et al. 2014). This controversial position of the Brazilian *Melocactus* species in the phylogeny of Majure et al. (2022) may result from inadequate taxonomic sampling and insufficient phylogenetic resolution in plastid regions. The absence of most Brazilian species in Majure et al. (2022) study likely led to an imprecise placement of these species, which comprises at least 50% of the genus diversity (sensu Hunt et al. 2006; Flora e Funga do Brasil 2025).

Despite the widespread recognition of the importance of taxon sampling for reliable phylogenetic inference (Nabhan & Sarka 2012 and references herein), the nature of the molecular data used cannot be ignored. Studies have shown that inadequate taxon sampling combined with a lack of phylogenetic signals can lead to unresolved and misleading phylogenies, even with multiple loci (Baurain, Brinkmann & Philippe 2007; Reddy et al. 2017). Although plastid data have been valuable for plant phylogenetics, providing outstanding progress in clarifying deep phylogenetic relationships within plants (Gitzendanner et al. 2018 and references therein), the variation in this genome is often insufficient to resolve shallower levels, especially in groups that experienced rapid and recent radiation (e.g. Schneider et al. 2021; Chen et al. 2022), like the Cactaceae family (Franco et al. 2022). This issue is reflected in our plastid phylogeny, in which the deeper relationships are well-established and highly supported (major clades), but the shallowest relationships (within clades) lack resolution. Because our main interest is understanding the species relationships, we focused the following discussion on the phylogenetic hypothesis that recovered mostly well-supported relationships in the genus (nuclear dataset), even though the plastid dataset recovered a similar backbone.

2.4.2 Phylogenomic inferences and systematics

The nuclear dataset combined with MSC inference clarified the closest lineages, monophyly, and relationships in *Melocactus*. As expected, *Melocactus* was recovered as sister to the genus *Discocactus*, agreeing with previous molecular studies (e.g., Hernández-Hernández et al. 2011; Silva et al. 2018; Romeiro-Brito et al. 2023b). This result refutes the taxonomic hypothesis suggesting these genera are distantly related and that their morphological similarities, such as globose habit, basic flower structure, and the presence of a terminal cephalium, arose from evolutionary convergence (Taylor & Zappi 2004). Furthermore, our results are consistent with many taxonomists who have long proposed a close relationship between *Coleocephalocereus* and *Melocactus* (e.g., Ritter 1968, 1979; Barthlott 1979; Braun 1988; Taylor 1991). Both genera share similarities in their flowers, fruits (including their emergence from the cephalium), seeds, and apically hooked spines in seedlings (Taylor 1991). Notably, three of the six informal species groups proposed by Taylor (1991) were supported as clades, even though these groups were originally based on limited morphological and seed micro-morphological traits, some of which have evolved independently elsewhere in the genus (Taylor 1991).

The M. OREAS clade comprises all species belonging to the M. OREAS Group defined by Taylor (1991), confirming its monophyly. This clade is distinguished by synapomorphies in seed, fruit, and spine morphology (details Table S3). Even though some traits have evolved independently in other *Melocactus* groups (e.g., the spine characteristics also seen in the distantly related *M. violaceus*), the combination of these traits with geographical distribution, ecological preferences (e.g., species restricted to gneiss, sandstone, or other crystalline rocks; Taylor 1991), and monophyly corroborates the M. OREAS clade as a well-established group. Although recovered as monophyletic, this is a taxonomically challenging group in which the genetic data highlighted the need for some reclassification of terminal taxa. For example, *M. bahiensis* subsp. *bahiensis* (*M. bahiensis*) and *M. bahiensis* subsp. *amethystinus* (Buining & Brederoo) N.P.Taylor (*M. amethystinus* Buining & Brederoo) appear as distinct and unrelated taxa. Our results indicated that *M. bahiensis sensu stricto* is closely related to *Melocactus oreas* Miq. and *Melocactus conoideus* Buining & Brederoo. Conversely, its former subspecies, *M. bahiensis* subsp. *amethystinus* (Buining & Brederoo) N.P.Taylor, is grouped with *M. ernestii* Vaupel and *M. longicarpus* Buining & Brederoo, supporting its recognition as a separate species. Diagnostic traits distinguishing *M. bahiensis* from *M. amethystinus* include the presence of very acute, often triangular ribs in cross-section, these swollen subtending the

areoles, and longer spines in *M. amethystinus*. Additionally, *M. amethystinus* is exclusively found on solid rocks (e.g., granite, gneiss; Fig. 1) and never in gravelly substrates (e.g., quartz gravel; Fig. 1) like *M. bahiensis* sensu *stricto*.

The M. DEINACANTHUS clade suggests a sister relationship between *M. deinacanthus* and *M. paucispinus*, though this relationship requires further investigation. *Melocactus deinacanthus* is an isolated species endemic to exposed gneiss/granite, differing significantly from *M. paucispinus*, a *campo rupestre* species that grows in sandy or occasionally quartz gravel substrates (Taylor & Zappi 2004). Additionally, *M. deinacanthus* shows distinctive fruit and seed morphology, which readily distinguishes it from *M. paucispinus* and other *Melocactus* species (Taylor 1991; Taylor & Zappi 2004). Despite these marked differences, the close phylogenetic relationship between the two species was weakly supported in our MSC tree and was not recovered on the concatenated tree. Although our results suggest that both species form a monophyletic group – the M. DEINACANTHUS clade – this relationship requires further investigation.

The M. AZUREUS clade includes all species from the M. AZUREUS Group (including *M. sergipensis*) and the monotypic M. LEVITESTATUS Group, rejecting the monophyly of these two groups proposed by Taylor (1991). *Melocactus sergipensis*, described by Taylor & Meiado (2014), was previously suggested as closely related to the M. VIOLACEUS Group due to its morphology, resembling especially *Melocactus concinnus* Buining & Brederoo and *M. violaceus* (e.g., grey-green to glaucous epidermis, dwarf stature, small pinkish fruits). However, it differs ecologically from M. VIOLACEUS Group by being restricted to limestone substrate. All species in the M. AZUREUS clade are restricted to limestone outcrops, varying only in geographic distribution (Taylor 1991; Taylor & Zappi 2004). Species that once belonged to the M. AZUREUS Group occur on limestone in northern Bahia, while *Melocactus levitestatus* Buining & Brederoo (previously the sole member of the M. LEVITESTATUS Group) is found on limestone in central-northern Minas Gerais, western Bahia, and eastern Goiás, and *M. sergipensis* is restricted to limestone in Sergipe. In this clade, as evidenced by the inclusion of *M. sergipensis*, the edaphic substrate appears to better explain species relationships than morphology. Although the species in this clade share the presence of white or pale pink fruits, and seeds often large with testa-cells almost smooth (details Table S3), morphology alone does not precisely define this group. The shared traits, consistent limestone ecology, and confirmed monophyly suggest that this clade represents a well-defined assemblage adapted exclusively to limestone substrates.

The M. INTORTUS clade comprises exclusively Caribbean and northern South American/Andean species. As previously suggested by Taylor (1991) and confirmed by Majure et al. (2022) the Cuban species *M. matanzanus* is not a member of this clade. The M. INTORTUS clade corresponds to the M. CURVISPINUS Group and is confirmed here as a monophyletic group. Despite the absence of synapomorphies, Taylor (1991) described this clade as a natural group based on shared floral traits (e.g., flowers often well-exserted from cephalium). Regardless of the confirmed monophyly, the species within this clade exhibit significant morphological diversity, making species relationships challenging to resolve based on characteristics like spines and fruits (see Table S3). Additional studies with broader sampling are needed to better understand the relationships within this clade. While *M. curvispinus* Pfeiff. was omitted from the final phylogenetic inferences with the nuclear dataset due to the unsuccessful sequencing (93% missing data), it clustered within the M. INTORTUS clade in preliminary nuclear phylogenies (data not shown), and this relationship received high support in plastid phylogeny. Based on morphology, preliminary nuclear findings, and plastid phylogenetic inference, *M. curvispinus* is confidently a member of the M. INTORTUS clade.

The species in the M. ZEHNTERI and M. VIOLACEUS clades largely align with the M. VIOLACEUS Group as circumscribed by Taylor (1991), revealing the non-monophyly of Taylor (1991) group. Together with *M. matanzanus*, the M. ZEHNTNERI clade comprises *M. zehntneri*, *M. concinnus*, and *M. salvadorensis*, previously classified within the M. VIOLACEUS Group by Taylor (1991) (see details Table S3). A close relationship between *Melocactus zehntneri* (Britton & Rose) Luetzelb and *Melocactus salvadorensis* Werderm. has long been suggested due to their morphological similarities, particularly in the vegetative state, which can make it difficult to distinguish between species in the absence of fruit or seed characters (Rizzini 1982; Taylor 1991). Similarly, the close relationship between *M. concinnus* and *M. zehntneri* is not unexpected. Specimens of *M. concinnus* were once described as a subspecies of *M. zehntneri* (*M. zehntneri* subsp. *robustispinus*; Braun, 1988) but later recognized as a synonym of *M. concinnus* by Taylor (1991). *Melocactus matanzanus* has also been suggested as closely related to Brazilian species based on morphology (Taylor 1991; Hunt et al. 2006), resembling a dwarf form of the Amazonian *M. neryi* (not included in this study) and *M. concinnus*.

The M. VIOLACEUS clade includes species formerly assigned to the M. VIOLACEUS Group, specifically *M. violaceus* and *M. glaucescens*. Within this clade, two subclades are identified: one for the species and subspecies of *M. violaceus* (subclade M. violaceus *s.s.*; see Table S3) and another for *M. depressus* and *M. glaucescens* (subclade M. depressus). Notably,

M. depressus was recovered as distinct from *M. violaceus*, revealing it to be a cryptic species. Originally described by William Hooker in 1838, *M. depressus* was distinguished by its sharply acute ribs, and shorter and slightly curved spines (Miquel 1843). However, Taylor (1980) later synonymized it with *M. violaceus* due to its extreme morphological similarity. Our phylogenetic inference, together with the morphological differences identified by Hooker (1838) and morphological differences, such as rib number and plant size, identified and assumed as cryptic by Taylor (1991), supports the recognition of *M. depressus* as a distinct species. Further field studies are necessary to clarify its geographic distribution, particularly along the coastal areas where its range overlaps with *M. violaceus*.

Even though some relationships still require further investigation due to topological incongruence (e.g., *M. paucispinus*) or due to its absence in our sampling (e.g. *M. smithii* (Alexander) Buining; *M. neryi* K.Schum.), the Cactaceae591 probe set associated with the MSC approach provided high support for most of the lineage relationships within *Melocactus*. In the present study, this probe set proved to be an effective tool for unraveling relationships in this species-rich and recently diverged group, confirming its resolution power as suggested by previous studies within Cactaceae (Romeiro-Brito et al. 2022; Romeiro-Brito et al. 2023b; Taylor et al. 2023; Zappi et al. 2024). Target sequencing capture has been efficiently used to clarify the phylogenetic relationships in plants, particularly in taxonomically and evolutionary challenging groups (e.g., radiation), becoming a valuable tool for taxonomic revisions and the discovery of cryptic taxa, regardless of the availability of fresh living materials (e.g., Carter et al. 2019; Loiseau et al. 2019; Nicol et al. 2024).

2.4.3 Climate and substrate as speciation drivers

We estimated younger divergence times than those by Silva et al. (2018), though the confidence intervals overlap (Silva et al. 2018: 95% HPD 6.77–1.07 Ma). We found that the divergence between *Melocactus* and *Discocactus* began in the late Pliocene (2.90 Ma), with *Melocactus* diversification initiating in the Pleistocene. Notably, speciation intensified in the M. INTORTUS clade from the Gelasian to the Chibanian Ages (0.774 – 0.129 Ma). The estimated ages coincide with periods of atmospheric CO_2 reduction and glaciation cycles followed by interglacial periods (Tripati, Roberts & Eagle 2009; Guillermic et al. 2022). Several Cactaceae genera, including *Cereus* Mill. (Franco et al. 2017), *Eulychnia* Phil. (Merklinger et al. 2021), *Mammillaria* Haw. (Cornejo-Romero et al. 2014), and *Pilosocereus* Byles & G.D.Rowley (Bonatelli et al. 2014, Lavor et al. 2018, Romeiro-Brito et al. 2023a),

show similar diversification timing, suggesting Quaternary climatic changes as key diversification drivers.

Research has highlighted the impact of climate on Cactaceae evolution, especially the variables related to temperature and precipitation (Alvarado-Sizzo et al. 2018; Mosco 2019; Aquino et al. 2021; Amaral et al. 2022; Freilij et al. 2023; Thompson et al. 2024). In arid environments, climate changes impose significant selective pressures, affecting the distribution and reproductive success of cactus lineages (Gudiño et al. 2011; Freilij et al. 2023). Our results provide empirical support for the role of climate in driving cactus evolution, as our ancestral reconstructions revealed climatic divergence among *Melocactus* lineages (Fig. 4), a finding supported by observations of temperature preferences among *Melocactus* lineages in greenhouse conditions. For instance, species from the M. INTORTUS clade struggle at low minimum temperatures (<20°C), while most Brazilian lineages tolerate them (Taylor, N. and Olsthoorn, G. pers. obs.).

Edaphic factors also play a significant influence on plant species distribution and evolution (Hulshof & Spasojevic, 2020; Azevedo et al. 2023). Many South American Cactaceae, including *Pilosocereus* Byles & G.D.Rowley, *Cereus* Mill., *Melocactus* (L.) Link & Otto, *Discocactus* Pfeiff., and *Arrojadoa* Britton & Rose, are restricted to rocky outcrops. Edaphic specialization seems to be fundamental to *Melocactus* evolution, as evidenced by multiple shifts in substrate preference. The diversification of major clades often occurred subsequent to these shifts, as observed in the M. AZUREUS clade, whose species exclusively occur on limestones.

Rocky outcrops and spatial edaphic heterogeneity, prevalent in core areas of *Melocactus* distribution, such as the Brazilian Seasonally Dry Tropical Forest (Taylor & Zappi 2004; Fernandes et al. 2022), provide colonization opportunities that may promote speciation (Kruckeberg 1986; Rajakaruna 2004). This heterogeneity leads to a fragmented species range, restricting gene flow among populations and facilitating genetic differentiation (Rajakaruna 2018). During Pleistocene interglacial periods, rocky outcrops likely acted as ecological refuges for dry-adapted lineages, fragmenting and isolating widespread populations (e.g., Bonatelli et al. 2014; Maswanganye et al. 2017; de Aguiar-Campos et al. 2020).

Despite the challenges of bare rocky and sandy substrates, *Melocactus* species are well-adapted to these harsh habitats exhibiting traits such as a water storing stem, waxy epidermis, and superficial roots with root hair buds/spurs (Boke 1979; Zappi & Taylor 2024). These adaptations likely facilitate the colonization of hostile and low-competition environments, suggesting a potential tradeoff in competitive-colonization ability (Tilman, 1994). While cacti

are generally poor competitors due to their slow growth rate (Godinez-Alvarez et al. 2003), their resilience to harsh conditions found in bare rocky and sand environments allows them to establish in these challenging environments.

Adaptive plasticity may also play a role in *Melocactus* diversification. Although its role in edaphic specialization remains unclear, ancestral adaptive plasticity may have facilitated lineage expansion into new edaphic habitats. This plasticity enables lineages to initially tolerate and subsequently adapt to the new environment (e.g. Robinson & Dukas 1999; Sommer, 2020 and references therein). Rather than acting as an evolutionary mechanism, plasticity acts as an initial step to adaptation (Ghalambor et al. 2007). For example, *M. violaceus*, a sand specialist, can germinate and grow in clay soils, although exhibiting much less vigorous growth (Fernandes & Maciel 2023).

The potential for adaptive plasticity, combined with past climatic oscillations, facilitated range expansion and species diversification in *Melocactus*. The interplay among dispersal facilitated by plasticity, isolation in habitat refuges, and local adaptations likely exercised a crucial role in driving their diversity and geographic distribution (Rajakaruna 2004; Rajakaruna 2018). Our phylogenetic ancestral reconstruction aligns with the complex evolutionary history of *Melocactus*, providing a clear understanding of its phylogenetic relationships and distribution patterns.

In addition to spatial edaphic heterogeneity, other factors intrinsic to the ecology may have further contributed to population genetic differentiation. *Melocactus* pollinators, involving hummingbirds with localized territories (Taylor 1991; Taylor & Zappi 2004), and seed dispersals occurring locally by lizards and ants (Nassar et al. 2001), limit gene flow between distant populations. It is worth mentioning that, even though principal pollinators and dispersals act more locally leading to a more limited gene exchange among populations, eventual long-distance dispersal cannot be disregarded. Long-distance dispersal could be mediated by birds (Taylor, 1991) or by the movement of detached cephalia. *Melocactus* cephalia can contain hundreds of viable seeds, and due to their lightness and shape, they can be transported by water and heavy rain, wind, or ocean currents, colonizing new and more distant areas (Barrios et al. 2021).

Understanding the complex evolutionary history of *Melocactus* is crucial for developing effective conservation strategies. Approximately 40% of *Melocactus* species are listed as Vulnerable, Endangered, or Critically Endangered (IUCN 2023), particularly the edaphic specialists. Beyond illegal *Melocactus* collection for sale, mining activities like limestone extraction for cement and fertilizer production pose a significant threat (Ribeiro-Silva et al.

2011). Conservation efforts should prioritize preserving these habitats, especially for littoral species like *M. violaceus* and *M. depressus*, where tourism activities have already caused population declines near many coastal resorts (Taylor pers. obs.).

2.4.4 Taxonomic adjustments

Melocactus comprises taxonomically challenging lineages, in which the phylogenomic inferences highlighted the need for some taxon rearrangement. Based on our phylogenomic inferences, two species were reinstated as follows (see keys to the clades for each species below):

Melocactus amethystinus Buining & Bredero in Kakteen, Lfg 50-51: CVI d. 1972. Type: Brazil, Bahia, Caitité, Brejinho das Ametistas, *Horst 270* (U531291, holo.). = *Melocactus bahiensis* subsp. *amethystinus* (Buining & Bredero) N.P.Taylor in Bradleya 9:30. 1991.

Diagnosis/Notes: The reinstated species, *M. amethystinus*, stands out as an independent evolutionary lineage in our phylogenetic inferences. It is geographically isolated from *M. bahiensis* and is morphologically distinct due to its very acute ribs, usually triangular in cross-section, swollen subtending the areoles, and longer spines. Ecologically, *M. amethystinus* is found exclusively on rock massifs in the Cadeia do Espinhaço mountain range, in stark contrast to *M. bahiensis*, which generally occurs on gravelly quartzitic substrates at the eastern edges of the Chapada Diamantina and on hills further east. *Melocactus amethystinus* is distinguished with difficulty from the similarly lithophytic *M. inconcinnus* Buining & Brederoo, which ranges further to the east and northeast. *Melocactus amethystinus* exhibits a more depressed-globose habit and shorter radial spines, while *M. inconcinnus* tends to be slightly more elongated with longer spines, however, these differences are subtle.

Key To M. Ernestii Subclade

1. Ribs +/- rounded in cross-section; central spines 4–8 per areole, lowermost radial spine 45–160 mm ...…...…..…. 2
1. Ribs acutely triangular in cross-section; central spines 1(-4) per areole, lowermost radial spine 40–60 mm ……..…..... 3
2. Central spines 4 per areole, lowermost radial spine to 160 mm; stigma-lobes not exserted, white ……...……..……. *M. ernestii*

2. Central spines 4–8 per areole, lowermost radial spine 45–90 mm; stigma-lobes exserted, pink ……..…….…….…. *M. longicarpus*
3. Stem (excl. cephalium) subglobose to somewhat elongate; ribs 8–12, smooth-sided; lowermost radial spine to 60 mm …….……..... *M. inconcinnus*
3. Stem (excl. cephalium) depressed globose; ribs 9–14, swollen subtending the areoles; lowermost radial spine to 40 mm …….……...…… *M. amethystinus*

Melocactus depressus Hook. in Curtis's Botanical Magazine 65: tab. 3691. 1838. Type: Pernambuco. Gardner, not preserved; Lectotype (Taylor, 1991): loc. cit., tab. 3691. = *Melocactus violaceus* Pfeiff. (subsp. *violaceus*) in Allg. Gartenz. 3: 313 (1835); Enum. Cact. 45-46 (1837), pro parte (Taylor, 1991).

Diagnosis/Notes: The reinstated species, *M. depressus*, is an independent evolutionary lineage in our phylogenetic inferences; however, it is vegetatively and ecologically very similar to *M. violaceus* subsp. *violaceus*. Both taxa are specialized in sand-dune habitats mainly distributed on the Brazilian littoral coast. However, there are also key differences, with *M. depressus* showing more numerous spines, number of ribs, and larger plant size. The range of *M. depressus* is markedly disjunct, from high altitude sandy cerrado in northeastern Minas Gerais at c. 1150 m asl (SE Brazil) to littoral habitats more than 1000 km distant in Pernambuco to Ceará (NE Brazil).

Key To M. Violaceus Clade

1. Stem strongly blue-waxy glaucous; fruit red ……...….. *M. glaucescens*
1. Stem not glaucous; fruit pink to white ……...…….. 2
2. Principal spines 7–8 per stem areole, 1 often obviously central; stem 10-16 cm in diameter; ribs 10-15 ……..……..……..…….……. *M. depressus*
2. Principal spines 5–6 per stem areole, central lacking; stem 6–10 cm in diameter; ribs 8–10 ……...…… *M. violaceus*

2.5 Acknowledgments

We are especially grateful to Dr. Marlon Machado, Dr. Diego Gonzaga, and the *Coleção de Cactos e Suculentas* of the Instituto de Pesquisas Jardim Botânico do Rio de Janeiro for contributing with the cactus samples used in this study.

2.6 Funding

This study was financially supported by FAPESP, which provided fellowships for doctoral students (Processes 2019/11233-2 to MCT and 2018/06937-8 to MRB) and research grants (2019/03211-9 to EMM, 2020/15161-3 to FFF and 2018/03428-5 to EMM and FFF). DCZ and EMM hold productivity grants from CNPq (304178/2021-7 and 310963/2022-6, respectively).

2.7 References

Aguilera, P. M., H. J. Debat, Y. S. García, D. A. Martí, & M. Grabiele. (2014). IAPT/IOPB chromosome data 18. *Taxon*, 63(6), 1387-1393.

Alvarado-Sizzo, H., A. Casas, F. Parra, H. J. Arreola-Nava, T. Terrazas, & C. Sánchez. (2018). Species delimitation in the *Stenocereus griseus* (Cactaceae) species complex reveals a new species, *S. huastecorum*. *Plos One*, 13(1), e0190385.

Amaral, D. T., J. R. Bombonato, S. C. da Silva Andrade, E. M. Moraes, & F. F. Franco. (2021). The genome of a thorny species: comparative genomic analysis among South and North American Cactaceae. *Planta*, 254(3), 44.

Amaral, D. T., I. A. S. Bonatelli, M. Romeiro-Brito, E. M. Moraes, & F. F. Franco. (2022). Spatial patterns of evolutionary diversity in Cactaceae show low ecological representation within protected areas. *Biological Conservation*, 273, 109677.

Amaral, D. T., I. Minhós-Yano, J. V. M. Oliveira, M. Romeiro-Brito, I. A. S. Bonatelli, N. P. Taylor, N. P., D. C. Zappi, E. M. Moraes, D. A. R. Eaton, F. F. Franco. (2021). Tracking the xeric biomes of South America: The spatiotemporal diversification of Mandacaru cactus. *Journal of Biogeography*, 48(12), 3085–3103.

Aquino, D., A. Moreno-Letelier, M. A. González-Botello, & S. Arias. (2021). The importance of environmental conditions in maintaining lineage identity in *Epithelantha* (Cactaceae). *Ecology and Evolution*, 11(9), 4520–4531.

Azevedo, L., D. C. Zappi, D. M. G. de Oliveira, L. Meyer, E. Nic Lughadha, R. Clegg, L. D. Meireles, P. H. A. de Melo, R. T. Pennington, D. M. Neves. (2024). On the rocks: Biogeography and floristic identity of rocky ecosystems in eastern South America. *Journal of Systematics and Evolution*, *62*(2), 305-320.

Barrios, D., S. Arias, L. R. González-Torres, & L. C. Majure. (2023). Lista anotada de cactus nativos y naturalizados de Cuba. *Botanical Sciences*, *101*(4), 1249-1300.

Barrios, D., S. Toledo, J. A. Sánchez, & L. R. González-Torres. (2021). Serotiny in *Melocactus matanzanus* (Cactaceae) and role of cephalium in dispersal of seeds after the individual's death. *Seed Science Research, 31*(4), 326-332.

Barthlott, W. (1979). *Cacti:* Botanical aspects, Descriptions, and Cultivation. Cheltenham, UK. Stanley Thornes.

Baurain, D., H. Brinkmann, & H. Philippe. (2007). Lack of resolution in the animal phylogeny: closely spaced cladogeneses or undetected systematic errors? *Molecular Biology and Evolution*, 24(1), 6–9.

Blomberg, S. P., Jr. T. Garland, & A. R. Ives. (2003). Testing for phylogenetic signal in comparative data: behavioral traits are more labile. *Evolution*, 57(4), 717–745.

Boke, N. H. (1979). Root glochids and root spurs of *Opuntia arenaria* (Cactaceae). *American Journal of Botany*, 66(9), 1085–1092.

Bombonato, J. R., D. T. do Amaral, G. A. R. Silva, G. Khan, E. M. Moraes, S. C. da Silva Andrade, D. A. R. Eaton,D. P. Alonso, P. E. M. Ribolla, N. P. Taylor, D. C. Zappi, & F. F. Franco. (2020). The potential of genome-wide RAD sequences for resolving rapid radiations: a case study in Cactaceae. *Molecular Phylogenetics and Evolution, 151*, 106896.

Bonatelli, I. A., M. F. Perez, A. T. Peterson, N. P. Taylor, D. C. Zappi, M. C. Machado, I. Koch, A. H. C. Pires, & E. M. Moraes. (2014). Interglacial microrefugia and diversification of a cactus species complex: phylogeography and palaeodistributional reconstructions for *Pilosocereus aurisetus* and allies. *Molecular Ecology, 23*(12), 3044–3063.

Boyko, J. D., & J. M. Beaulieu, J. M. (2022). corHMM 2.1: Generalized hidden Markov models. *R package version 2.8*.

Cai, L., Z. Xi, E. M. Lemmon, A. R. Lemmon, A. Mast, C. E. Buddenhagen, L. Liu, & C. C. Davis. (2021). The perfect storm: gene tree estimation error, incomplete lineage sorting, and ancient gene flow explain the most recalcitrant ancient angiosperm clade, Malpighiales. *Systematic Biology, 70*(3), 491–507.

Capella-Gutiérrez, S., J. M. Silla-Martínez, & T. Gabaldón. (2009). trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics, 25*(15), 1972–1973.

Carter, K. A., A. Liston, N. V. Bassil, L. A. Alice, J. M. Bushakra, B. L. Sutherland, T. C. Mockler, D. W. Bryant, & K. E. Hummer. (2019). Target capture sequencing unravels Rubus evolution. *Frontiers in Plant Science, 10*, 1615.

Chen, S., Y. Zhou, Y. Chen, & J. Gu. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics, 34*(17), i884–i890.

Chen, Y. P., F. Zhao, A. J. Paton, P. Sunojkumar, L. M. Gao, & C. L. Xiang. (2022). Plastome sequences fail to resolve shallow level relationships within the rapidly radiated genus *Isodon* (Lamiaceae). *Frontiers in Plant Science*, *13*, 985488.

Copetti, D., A. Búrquez, E. Bustamante, J. L. Charboneau, K. L. Childs, L. E. Eguiarte, S. Lee, T. L. Liu, M. M. McMahon, N. K. Whiteman, R. A. Wing, M. F. Wojciechowski, & M. J. Sanderson. (2017). Extensive gene tree discordance and hemiplasy shaped the genomes of North American columnar cacti. *Proceedings of the National Academy of Sciences*, *114*(45), 12003–12008.

Cornejo-Romero, A., J. Medina-Sánchez, T. Hernández-Hernández, B. Rendón-Aguilar, P. L. Valverde, A. Zavala-Hurtado, S. P. Rivas-Arancibia, M. A. Pérez-Hernández, G. López-Ortega, C. Jiménez-Sierra, & C. F. Vargas-Mendoza. (2014). Quaternary origin and genetic divergence of the endemic cactus *Mammillaria pectinifera* in a changing landscape in the Tehuacán Valley, Mexico. *Genetics and Molecular Research*, *13*(1), 73–88.

de Aguiar-Campos, N., V. A. Maia, W. B. da Silva, C. R. de Souza, & R. M. dos Santos. (2020). Can fine-scale habitats of limestone outcrops be considered litho-refugia for dry forest tree lineages? *Biodiversity and Conservation*, **29**(3), 1009–1026.

Edwards, S. V., Z. Xi, A. Janke, B. C. Faircloth, J. E. McCormack, T. C. Glenn, B. Zhong, S. Wu, E. M. Lemmon, A. R. Lemmon, A. D. Leaché, L. Liu, & C. C. Davis. (2016). Implementing and testing the multispecies coalescent model: a valuable paradigm for phylogenomics. *Molecular Phylogenetics and Evolution*, **94**, 447–462.

Edwards, S. V. (2009). Is a new and general theory of molecular systematics emerging? *Evolution*, 63(1), 1–19.

Fernandes, M. F., D. Cardoso, R. T. Pennington, & L. P. de Queiroz. (2022). The origins and historical assembly of the Brazilian Caatinga seasonally dry tropical forests. *Frontiers in Ecology and Evolution*, **10**, 723286.

Fernandes, M. M., & J. R. Maciel. (2023). Potencial adaptativo de *Melocactus violaceus* Pfeiff (Cactaceae) para solos argilosos. *Hoehnea*, 50, e582021.

Fick, S. E., & R. J. Hijmans. (2017). WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, 37(12), 4302–4315.

Flora do Brasil. (2020). Flora do Brasil. Jardim Botânico do Rio de Janeiro. Available online: http://floradobrasil.jbrj.gov.br/ (Accessed, May 13th, 2024)

Fonseca, R. B. S., L. S. Funch, & E. L. Borba. (2008). Reproductive phenology of *Melocactus* (Cactaceae) species from Chapada Diamantina, Bahia, Brazil. *Brazilian Journal of Botany*, *31*, 237–244.

Franco, F. F., D. T. Amaral, I. A. Bonatelli, M. Romeiro-Brito, M. C. Telhe, & E. M. Moraes. (2022). Evolutionary genetics of cacti: research biases, advances and prospects. *Genes*, **13**(3), 452.

Franco, F. F., G. A. R. Silva, E. M. Moraes, N. Taylor, D. C. Zappi, C. L. Jojima, & M. C. Machado. (2017). Plio-Pleistocene diversification of *Cereus* (Cactaceae, Cereeae) and closely allied genera. *Botanical Journal of the Linnean Society*, **183**(2), 199–210.

Freilij, D., D. Larrea-Alcazar, R. P. Lopez, F. Velarde Simonini, K. Naoki, & C. Bessega. (2023). Environmental gradualism explains variation in pollination systems of columnar cacti: Phylogenetic and trait evolution analyses. *Global Ecology and Biogeography*, **32**(6), 1015–1029.

Gatesy, J., & M. S. Springer. (2014). Phylogenetic analysis at deep timescales: unreliable gene trees, bypassed hidden support, and the coalescence/concatalescence conundrum. *Molecular Phylogenetics and Evolution*, 80, 231–266.

Ghalambor, C. K., J. K. McKay, S. P. Carroll, & D. N. Reznick. (2007). Adaptive versus non-adaptive phenotypic plasticity and the potential for contemporary adaptation in new environments. *Functional Ecology*, **21**(3), 394–407.

Godinez-Alvarez, H., T. Valverde, & P. Ortega-Baes. (2003). Demographic trends in the Cactaceae. *The Botanical Review*, **69**(2), 173–201.

Gitzendanner, M. A., P. S. Soltis, T. S. Yi, D. Z. Li, & D. E. Soltis. (2018). Plastome phylogenetics: 30 years of inferences into plant evolution. *Advances in Botanical Research* (Vol. 85, pp. 293–313). Academic Press.

González, L. R., A. Palmarola, D. Barrios, L. González, E. Testé, & E. R. Bécquer. (2016). Lista roja de la flora de Cuba. *Bissea*, 10, Número Especial 1, 352 p.

Gudiño, W., A. Casas, A. Valiente–Banuet, R. Orozco–Martínez, & E. de la Barrera. (2011). Climate and microenvironmental parameters affecting anthesis and nectar secretion for *Polaskia chende* and *P. chichipe*, endemic columnar cacti from the Tehuacán Valley, Puebla. *Journal of the Professional Association for Cactus Development*, 13, 88–101.

Guillermic, M., S. Misra, R. Eagle, & A. Tripati. (2021). Atmospheric CO_2 estimates for the Miocene to Pleistocene based on foraminiferal $\delta^{11}B$ at Ocean Drilling Program Sites 806 and 807 in the Western Equatorial Pacific. *Climate of the Past Discussions*, 2021, 1–40.

Hernandez, H. M., & R. T. Barcenas. (1995). Endangered cacti in the Chihuahuan Desert: I. Distribution patterns. *Conservation Biology*, 1176–1188.

Hernández-Hernández, T., J. W. Brown, B. O. Schlumpberger, L. E. Eguiarte, & S. Magallón. (2014). Beyond aridification: multiple explanations for the elevated diversification of cacti in the New World Succulent Biome. *New Phytologist*, 202(4), 1382–1397.

Hoang, D. T., O. Chernomor, A. Von Haeseler, B. Q. Minh, & L. S. Vinh. (2018). UFBoot2: improving the ultrafast bootstrap approximation. *Molecular Biology and Evolution*, 35(2), 518–522.

Hulshof, C. M., & Spasojevic, M. J. (2020). The edaphic control of plant diversity. *Global Ecology and Biogeography*, 29(10), 1634–1650.

Hunt, D. R., N. P. Taylor, & G. Charles. (2006). *New Cactus Lexicon*.

IUCN. 2023. *The IUCN Red List of Threatened Species*. Version 2023–1. https://www.iucnredlist.org. Accessed on May 2024.

Jiang, X., S. V. Edwards, S. V., & L. Liu. (2020). The multispecies coalescent model outperforms concatenation across diverse phylogenomic data sets. *Systematic Biology*, 69(4), 795–812.

Johnson, M. G., E. M. Gardner, Y. Liu, R. Medina, B. Goffinet, A. J. Shaw, N. J. C. Zerega, & N. J. Wickett. (2016). HybPiper: Extracting coding sequence and introns for phylogenetics from high-throughput sequencing reads using target enrichment. *Applications in Plant Sciences*, **4**(7), 1600016.

Katoh, K., & D. M. Standley. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution*, 30(4), 772–780.

Khan, G., F. F. Franco, G. A. Silva, J. R. Bombonato, M. Machado, D. P. Alonso, P. E.M. Ribolla, D. C. Albach, & E. M. Moraes. (2020). Maintaining genetic integrity with high promiscuity: Frequent hybridization with low introgression in multiple hybrid zones of *Melocactus* (Cactaceae). *Molecular Phylogenetics and Evolution*, **142**, 106642.

Kubatko, L. S., & J. H. Degnan. (2007). Inconsistency of phylogenetic estimates from concatenated data under coalescence. *Systematic Biology*, 56(1), 17–24.

Linnaeus, C. 1753. *Species Plantarum.* Stockholm (Sweden): impensis Laurentii Salvii.

Loiseau, O., I. Olivares, M. Paris, M. de La Harpe, A. Weigand, D. Koubínová, ... & N. Salamin. (2019). Targeted capture of hundreds of nuclear genes unravels phylogenetic relationships of the diverse Neotropical palm tribe Geonomateae. *Frontiers in Plant Science*, **10**, 864.

Maddison, W. P. (1997). Gene trees in species trees. *Systematic Biology*, 46(3), 523–536.

Majure, L. C., D. Barrios, E. Díaz, L. F. Bacci, & Y. E. Piñeyro. (2022). Phylogenomics of the Caribbean melocacti: Cryptic species and multiple invasions. *Taxon*, **71**(5), 993–1012.

Maswanganye, K. A., M. J. Cunningham, N. C. Bennett, C. T. Chimimba, & P. Bloomer. (2017). Life on the rocks: Multilocus phylogeography of rock hyrax (*Procavia capensis*) from southern Africa. *Molecular Phylogenetics and Evolution*, **114**, 49–62.

Merklinger, F. F., T. Böhnert, M. Arakaki, M. Weigend, D. Quandt, & F. Luebert. (2021). Quaternary diversification of a columnar cactus in the driest place on Earth. *American Journal of Botany*, **108**(2), 184–199.

Minh, B. Q., H. A. Schmidt, O. Chernomor, D. Schrempf, M. D. Woodhams, A. Von Haeseler, & R. Lanfear. (2020). IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Molecular Biology and Evolution*, 37(5), 1530–1534.

Miquel, F. A. W. (1843). *Monography Generis Melocacti*. 1–120.

Molina-Venegas, R., & M. Á. Rodríguez. (2017). Revisiting phylogenetic signal; strong or negligible impacts of polytomies and branch length information? *BMC Evolutionary Biology*, 17, 1–10.

Moraes, Evandro; M. C. Telhe, N. P. Taylor, M. Romeiro-Brito, D. C. Zappi, G. Olsthoorn, & F. F. Franco. (Forthcoming 2024). Data from: Phylogenomic and ecological systematics of *Melocactus* (Cactaceae) [Dataset]. Dryad. https://doi.org/10.5061/dryad.1rn8pk14k

Mosco, A. (2019). Specific habitat requirements and niche conservatism for nine species of the Mexican genus *Thelocactus* (Cactaceae). *Revista Mexicana de Biodiversidad*, 90, e902246.

Nabhan, A. R., & I. N. Sarkar. (2012). The impact of taxon sampling on phylogenetic inference: a review of two decades of controversy. *Briefings in Bioinformatics*, 13(1), 122–134.

Nassar, J. M., J. L. Hamrick, & T. H. Fleming. (2001). Genetic variation and population structure of the mixed-mating cactus, *Melocactus curvispinus* (Cactaceae). *Heredity*, *87*(1), 69–79.

Nicol, D. A., P. Saldivia, T. C. Summerfield, M. Heads, J. M. Lord, E. P. Khaing, & M. J. Larcombe. (2024). Phylogenomics and morphology of Celmisiinae (Asteraceae: Astereae): Taxonomic and evolutionary implications. *Molecular Phylogenetics and Evolution*, 195, 108064.

Pagel, M. (1999). Inferring the historical patterns of biological evolution. *Nature*, 401(6756), 877–884.

R Core Team (2023). R: A language and environment for statistical computing. *R Foundation for Statistical Computing*, Vienna, Austria. https://www.R-project.org/.

Rajakaruna, N. (2004). The edaphic factor in the origin of plant species. *International Geology Review*, 46(5), 471–478.

Rajakaruna, N. (2018). Lessons on evolution from the study of edaphic specialization. *The Botanical Review*, 84, 39–78.

Rannala, B., A. Leache, S. Edwards, & Z. Yang. (2020). The multispecies coalescent model and species tree inference. *Self Published*.

Reddy, S., R. T. Kimball, A. Pandey, P. A. Hosner, M. J. Braun, S. J. Hackett, K. Han, J. Harshman, C. J. Huddleston, S. Kingston, B. D. Marks, K. J. Miglia, W. S. Moore, F. H. Sheldon, C. C. Witt, T. Yuri, & E. L. Braun. (2017). Why do phylogenomic data sets yield conflicting trees? Data type influences the avian tree of life more than taxon sampling. *Systematic Biology*, *66*(5), 857–879.

Revell, L. J. (2012). phytools: an R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution*, (2), 217–223.

Ribeiro-Silva S, D. C. Zappi, N. P. Taylor, M. Machado. (2011). *Plano de Ação para Conservação das Cactáceas*. Série Espécies Ameaçadas nº 24. Instituto Chico Mendes de Conservação da Biodiversidade, ICMBio, Brasília.

Ritter, F. (1968). Die Cephalien tragenden Kakteen Brasiliens. *In Kakt. and. Sukk.*, 19, 119–122.

Ritter, F. (1979). *Kakteen in Südamerika*, vol. 1, Brasilien, Paraguay, Uruguay. Selbstverlag, Spangenberg.

Rizzini, C. T (1982). *Melocactus* no Brasil. 142p. IBDF, Jardim Botânico do Rio de Janeiro.

Robinson, B. W., & R. Dukas. (1999). The influence of phenotypic modifications on evolution: the Baldwin effect and modern perspectives. *Oikos*, *85*(3), 582-589.

Romeiro-Brito, M., G. Khan, M. F. Perez, D. C. Zappi, N. P. Taylor, G. Olsthoorn, F. F. Franco, & E. M. Moraes. (2023)a. Revisiting phylogeny, systematics, and biogeography of a Pleistocene radiation. *American Journal of Botany*, 110(3), e16134.

Romeiro-Brito, M., N. P. Taylor, D. C. Zappi, M. C. Telhe, F. F. Franco, & E. M. Moraes. (2023)b. Unravelling phylogenetic relationships of the tribe Cereeae using target enrichment sequencing. *Annals of Botany*, 132(5), 989–1006.

Romeiro-Brito, M., M. C. Telhe, D. T. Amaral, F. F. Franco, & E. M. Moraes. (2022). A target capture probe set useful for deep-and shallow-level phylogenetic studies in Cactaceae. *Genes*, 13(4), 707.

Ruedas, M., T. Valverde, & J. A. Zavala-Hurtado. (2006). Analysis of the factors that affect the distribution and abundance of three *Neobuxbaumia* species (Cactaceae) that differ in their degree of rarity. *Acta Oecologica*, 29(2), 155–164.

Schneider, J. V., J. Paule, T. Jungcurt, D. Cardoso, A. M. Amorim, T. Berberich, & G. Zizka. (2021). Resolving recalcitrant clades in the Pantropical Ochnaceae: insights from comparative phylogenomics of plastome and nuclear genomic data derived from targeted sequencing. *Frontiers in Plant Science, 12,* 638650.

Scornavacca, C., & N. Galtier. (2017). Incomplete lineage sorting in mammalian phylogenomics. *Systematic Biology, 66*(1), 112–120.

Silva, G. A. R., A. Antonelli, A. Lendel, E. D. M. Moraes, & M. H. Manfrin. (2018). The impact of early Quaternary climate change on the diversification and population dynamics of a South American cactus species. *Journal of Biogeography, 45*(1), 76–88.

Sommer, R. J. (2020). Phenotypic plasticity: from theory and genetics to current and future challenges. *Genetics, 215*(1), 1–13.

Stuessy, T., & H. Weiss-Schneeweiss. (2019). What drives polyploidization in plants? *The New Phytologist, 223*(4), 1690–1692.

Suringar, W. F. R. (1896). Vierde bijdrage tot de kennis der Melocacti (No. 3). Müller.

Tabernaemontanus, J. T. (1591). *New vollkommen Kräuter-Buch.* 2nd ed. Frankfurt a. M.: Bassaeus.

Tamura, K., F. U. Battistuzzi, P. Billing-Ross, O. Murillo, A. Filipski, & S. Kumar. (2012). Estimating divergence times in large molecular phylogenies. *Proceedings of the National Academy of Sciences, 109*(47), 19333–19338.

Tamura, K., G. Stecher, & S. Kumar. (2021). MEGA11: molecular evolutionary genetics analysis version 11. *Molecular Biology and Evolution, 38*(7), 3022–3027.

Taylor, N. P. (1980). Notes on the genus *Melocactus* (1): E. Brazil. *The Cactus and Succulent Journal of Great Britain,* 63–70.

Taylor, N. P. (1991). The genus *Melocactus* (Cactaceae) in Central and South America. *Bradleya, 1991*(9), 1–80.

Taylor, N., M. V. Meiado, E. Bravo Filho, & D. Zappi. (2014). A new *Melocactus* from the Brazilian state of Sergipe. *Bradleya, 2014*(32), 99–104.

Taylor, N. P., & D. C. Zappi. (2004). *Cacti of Eastern Brazil.* Richmond: Royal Botanic Gardens, Kew.

Thompson, J. B., T. Hernández-Hernández, G. Keeling, M. Vásquez-Cruz, & N. K. Priest. (2024). Identifying the multiple drivers of Cactus diversification. *Nature Communications, 15*(1), 7282.

Tilman, D. (1994). Competition and biodiversity in spatially structured habitats. *Ecology, 75*(1), 2–16.

Title, P. O., & Bemmels, J. B. (2018). ENVIREM: an expanded set of bioclimatic and topographic variables increases flexibility and improves performance of ecological niche modeling. *Ecography, 41*(2), 291-307.

Tripati, A. K., C. D. Roberts, & R. A. Eagle. (2009). Coupling of CO2 and ice sheet stability over major climate transitions of the last 20 million years. *Science, 326*(5958), 1394–1397.

Werdermann, E. (1934). Cactaceae novae. *Notizblatt des Königl. botanischen Gartens und Museums zu Berlin, (112),* 223–229.

Wu, M., J. L. Kostyun, M. W. Hahn, & L. C. Moyle. (2018). Dissecting the basis of novel trait evolution in a radiation with widespread phylogenetic discordance. *Molecular Ecology, 27,* 3301–3316.

Xi, Z., L. Liu, & C. C. Davis. (2016). The impact of missing data on species tree estimation. *Molecular Biology and Evolution, 33*(3), 838–860.

Zappi, D., & N. P. Taylor. (2024). A note on the roots of *Maihuenia* (Cactaceae). *Bradleya, (42),* 43–45.

Zappi, D. C., N. P. Taylor, F. N. Costa, S. N. Fonseca, P. L. Ferreira, M. Romeiro-Brito, M. C. Telhe, M. Köhler, F. F. Franco, & E. M. Moraes. (2024). A microendemic and enigmatic new cactus species from the campo rupestre of Minas Gerais, Brazil: *Uebelmannia nuda* (Cactaceae, Cactoideae). *Taxon,* 73(4), 992-1000.

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

Zizka, A., D. Silvestro, T. Andermann, J. Azevedo, C. Duarte Ritter, D. Edler, H. Farooq, A. Herdean, M. Ariza, R. Scharn, S. Svantesson, N. Wengström, V. Zizka, & A. Antonelli. (2019). CoordinateCleaner: Standardized cleaning of occurrence records from biological collection databases. *Methods in Ecology and Evolution, 10*(5), 744–751.

.8 Supplementary Material

Figure S1. plastid gene recovery success per *Melocactus* sample sequenced using Cactaceae591. Scale color indicates success rate.

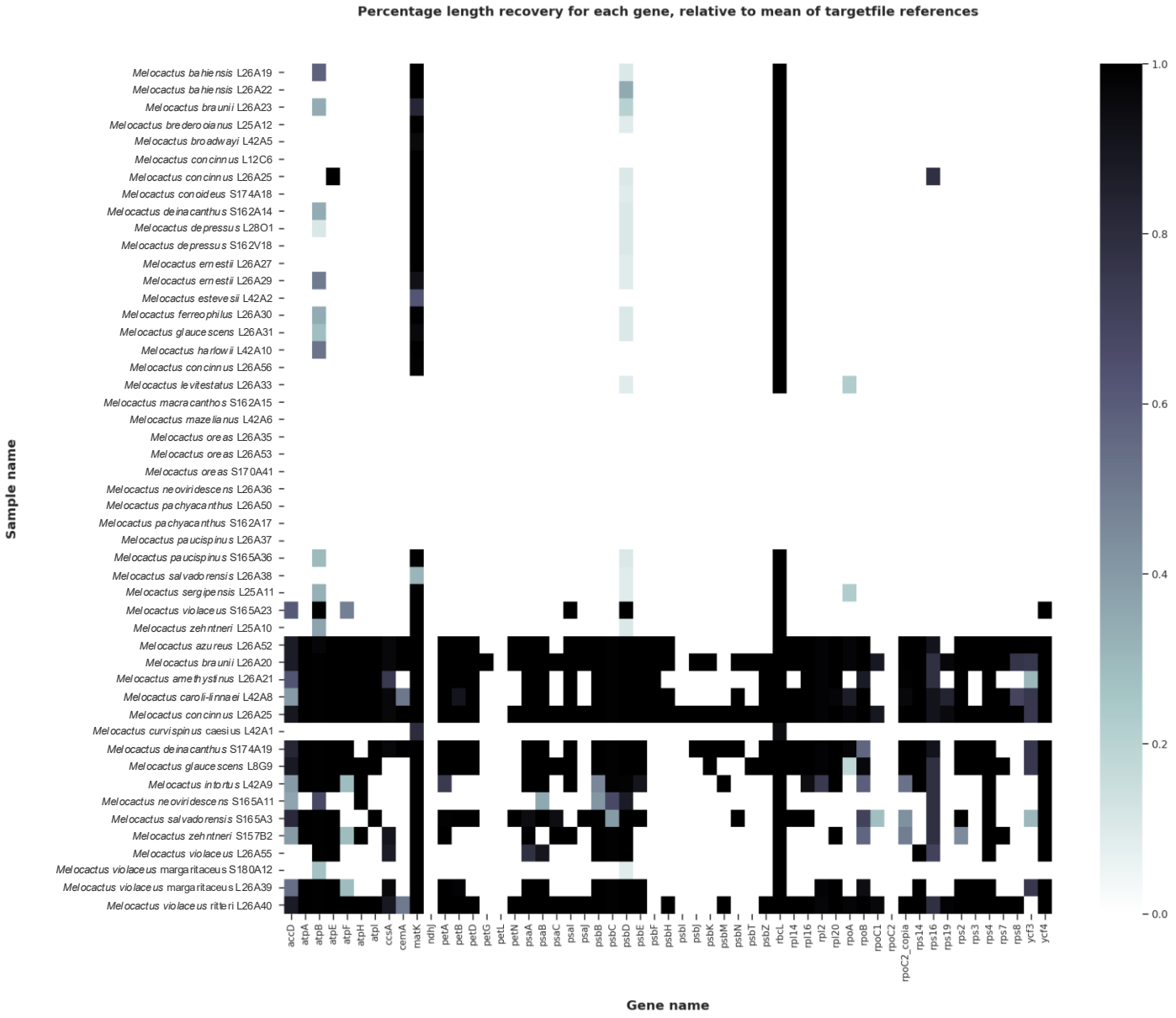


Figure S2. Concatenated phylogenetic tree of *Melocactus* using plastid dataset and IQ-TREE 2.

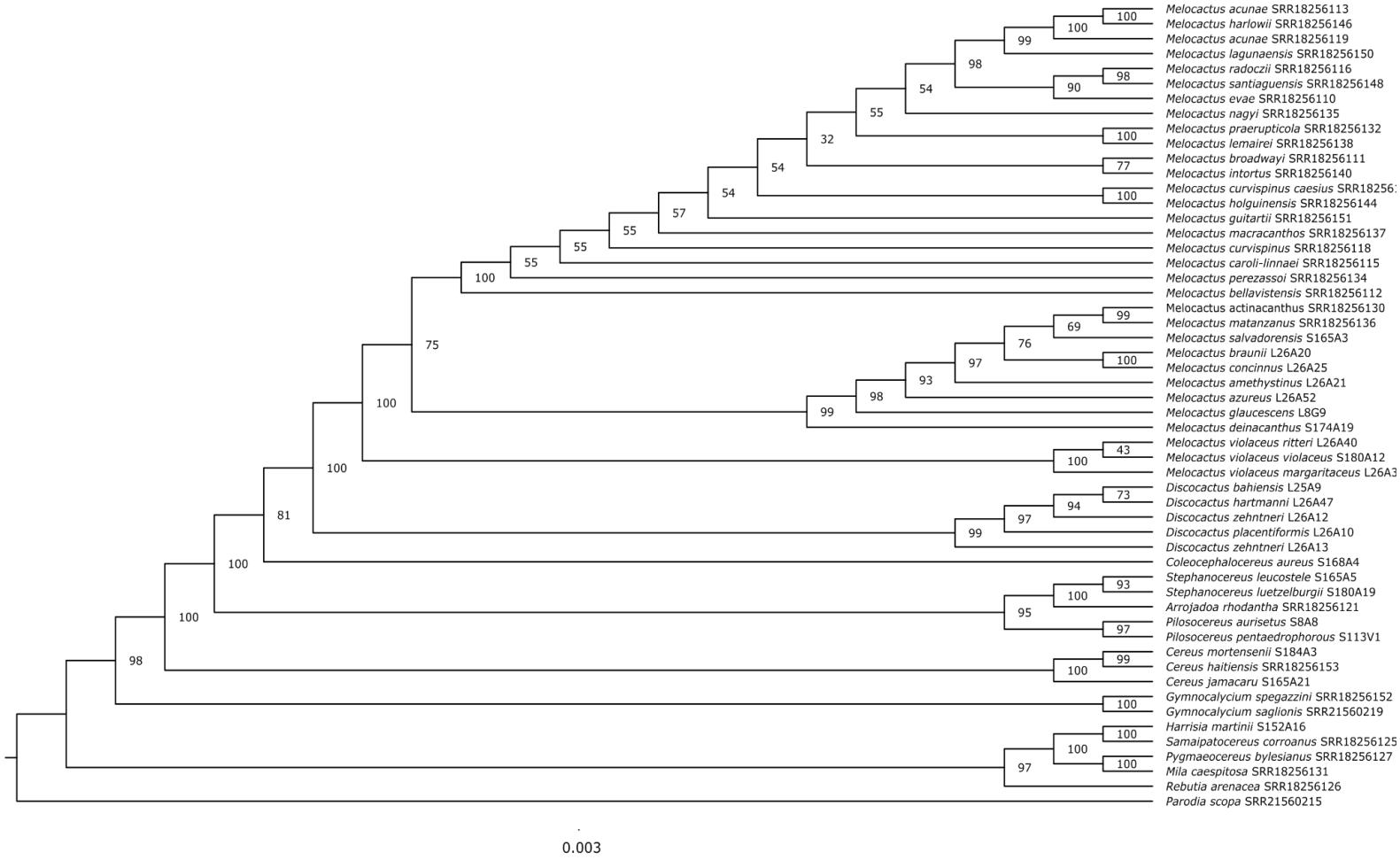


Figure S3. Phylogeny with node number used for the discrete ancestral character state reconstruction for substrate types in the genus *Melocactus*. Node numbers in this figure are the same as in Table S7.

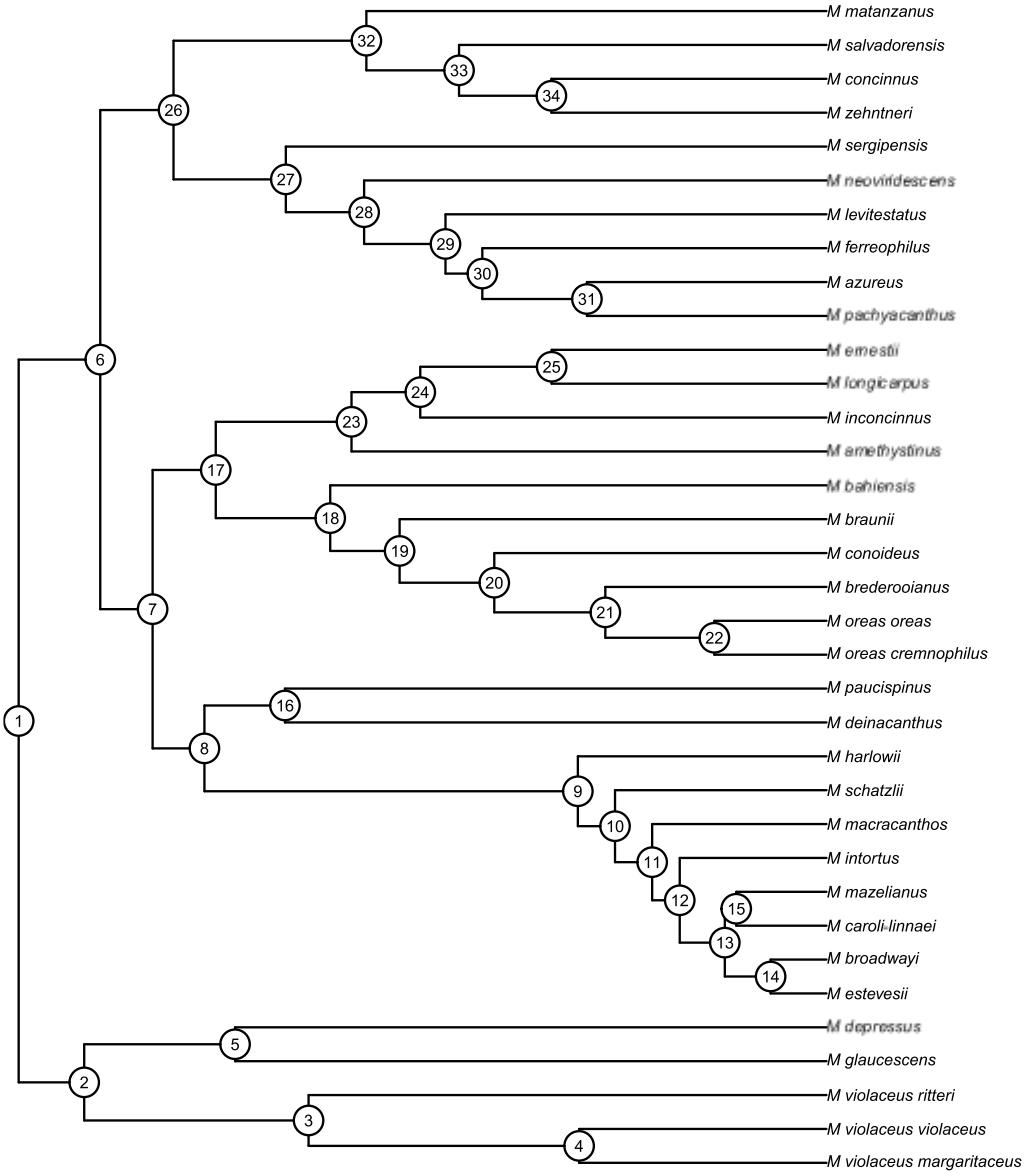


Table S1. Sampling information of *Melocactus* and outgroup species used for phylogenetic analysis. Most species included in this study are considered vulnerable or endangered, for this reason, no coordinates are included.

scientificName	scientificNameAuthorship	recordNumber	collectionID	country	county	locality	identifiedBy
Melocactus Link & Otto.							
Melocactus azureus	Buining & Brederoo	L26A52	SORO7999	Brazil	Bahia	Itaguaçu da Bahia	Olsthoorn, G
Melocactus bahiensis	(Britton & Rose) Luetzelb.	L86D4	Rose&Russell 19935	Brazil	Bahia	Marcionílio Sousa	Olsthoorn, G
Melocactus bahiensis ssp. amethystinus	(Buining & Brederoo) N.P.Taylor	L26A21	L26A21	Brazil	Minas Gerais	Itacambira	Olsthoorn, G
Melocactus bahiensis ssp. amethystinus	(Buining & Brederoo) N.P.Taylor	L26A22	L26A22	Brazil	Minas Gerais	Senador Modestino Gonçalves	Olsthoorn, G
Melocactus braunii	Esteves	L26A20	SORO7978	Brazil	Bahia	Jacobina	Olsthoorn, G
Melocactus braunii	Esteves	L26A23	SORO7978	Brazil	Bahia	Jacobina	Olsthoorn, G

Melocactus brederooianus	Buining	L25A12	SORO7994	Brazil	Bahia	Andorinha	Olsthoorn, G
Melocactus broadwayi	(Britton & Rose) A.Berger	L42A5	Anterberger, 64	Trinidad e Tobago	Tobago	Antesberger	Taylor, N.
Melocactus caroli-linnaei	N.P.Taylor	L42A8	L42A8	Jamaica	-	Black River	Taylor, N.
Melocactus concinnus	Buining & Brederoo	L26A25	L26A25	Brazil	Minas Gerais	Janauba	Olsthoorn, G
Melocactus conoideus	Buining & Brederoo	S174A18	S174A18	Brazil	Minas Gerais	Itinga	Olsthoorn, G
Melocactus conoideus	Buining & Brederoo	L86D1	L86D1	Brazil	Minas Gerais		Taylor, N.
Melocactus curvispinus ssp caesius	(H.L.Wendl.) N.P.Taylor	L42A1	L42A1	Venezue la	Aragua	Ocumare de la Costa	Taylor, N.
Melocactus deinacanthus	Buining & Brederoo	S162A14	SORO 8007	Brazil	Bahia	Bom Jesus da Lapa	Olsthoorn, G

Melocactus deinacanthus	Buining & Brederoo	S174A19	MBM 156954	Brazil	Bahia	Juá	Olsthoorn, G
Melocactus depressus	(Aiton) Hook.	L86D2	SPF 80886	Brazil	Minas Gerais	Jequitinhonha	Taylor, N.
Melocactus depressus	(Aiton) Hook.	L28O1	EAC 52277	Brazil	Ceará	Aquiraz	Menezes, M.
Melocactus depressus	(Aiton) Hook.	S162V18	-	Brazil	Rio Grande do Norte	Touros	Olsthoorn, G
Melocactus ernestii	Vaupel	L26A27	CEPEC 47009	Brazil	Bahia	Morro do Chapéu	Olsthoorn, G
Melocactus ernestii ssp longicarpus	(Buining & Brederoo) N.P.Taylor	L26A29	K000013174	Brazil	Bahia	Guanambi	Olsthoorn, G
Melocactus estevesii	P.J.Braun	L42A2	L42A2	Brazil	Roraima	-	Taylor, N.
Melocactus ferreophilus	Buining & Brederoo	L26A30	HUEFS 168189	Brazil	Bahia	Morro do Chapéu	Olsthoorn, G
Melocactus glaucescens	Buining & Brederoo	L26A31	HUEFS 168184	Brazil	Bahia	Morro do Chapéu	Olsthoorn, G

Melocactus glaucescens	Buining & Brederoo	L8G9	SORO 8174	Brazil	Bahia	Morro do Chapéu	M. Machado
Melocactus harlowii	Britton & Rose) Vaupel	L42A10	L42A10	Cuba	Santiago	Santa Maria de Loreto	Taylor, N.
Melocactus inconcinnus	Buining & Brederoo	L26A56	EAN 23473	Brazil	Bahia	Susuarana	Olsthoorn, G
Melocactus intortus	(Mill.) Urb.	L42A9	L42A9	Puerto Rico	-	-	Taylor, N.
Melocactus levitestatus	Buining & Brederoo	L26A33	SORO7982	Brazil	Minas Gerais	Capitão Enéas	Olsthoorn, G
Melocactus macracanthos	(Salm-Dyck) Link & Otto	S162A15	SORO7958	Curaçao	-	Shete Boka National Park	Olsthoorn, G
Melocactus matanzanus	León	L87A2	L87A2	Cuba	-	-	Olsthoorn, G
Melocactus mazelianus	Riha	L42A6	U.1180141	Venezuela	Ilha no Rio Orinoco	-	Taylor, N.

Melocactus neoviridescens	Guigii	L26A36	HUEFS 168185	Brazil	Bahia	Morro do Chapéu	Olsthoorn, G
Melocactus neoviridescens	Guigii	S165A11	SinBiota 23140	Brazil	Bahia	Morro do Chapéu	Taylor, N.
Melocactus oreas	Miq.	S170A41	RBvc 135	Brazil	Bahia	Itatim	Taylor, N.
Melocactus oreas	Miq.	L26A53	BHCB 154353	Brazil	Bahia	Milagres	Olsthoorn, G
Melocactus oreas	Miq.	L86D3	UB 238632	Brazil	Bahia	Ruy Barbosa	Taylor, N.
Melocactus oreas ssp cremnophilus	(Buining & Brederoo) P.J.Braun	L26A35	HUEFS 168182	Brazil	Bahia	Morro do Chapéu	Olsthoorn, G
Melocactus pachyacanthus	Buining & Brederoo	L26A50	HUEFS 168195	Brazil	Bahia	Morro do Chapéu	Olsthoorn, G
Melocactus pachyacanthus	Buining & Brederoo	S162A17	SORO8002	Brazil	Bahia	Umburanas	Olsthoorn, G
Melocactus paucispinus	Heimen & R.J.Paul	L26A37	SPF 81026	Brazil	Bahia	Seabra	Olsthoorn, G

Melocactus paucispinus	Heimen & R.J.Paul	S165A36	SinBiota 23155	Brazil	Bahia	Jacaraci	Taylor, N.
Melocactus salvadorensis	Werderm.	S165A3	SORO 8001	Brazil	Bahia	Jequié	Taylor, N.
Melocactus salvadorensis	Werderm.	L26A38	L26A38	Brazil	Bahia	Capim Grosso	Olsthoorn, G
Melocactus schatlzii	H.Till & R.Gruber	L87A4	L87A4	-	-	-	Olsthoorn, G
Melocactus sergipensis	N.P.Taylor & M.V.Meiado	L25A11	ASE 33139	Brazil	Sergipe	Simão Dias	Olsthoorn, G
Melocactus violaceus	Pfeiff.	L26A55	L26A55	Brazil	Rio de Janeiro	Saquarema	Olsthoorn, G
Melocactus violaceus	Pfeiff.	S165A23	SORO7979	Brazil	Bahia	Jacobina	Taylor, N.
Melocactus violaceus ssp margaritaceus	N. P. Taylor	S180A12	UFP47166	Brazil	Pernambuco	Catimbau	Taylor, N.

Melocactus violaceus ssp margaritaceus	N. P. Taylor	L26A39	SORO7991	Brazil	Bahia	Mata de São João	Olsthoorn, G
Melocactus violaceus ssp ritteri	N. P. Taylor	L26A40	SPF 80991	Brazil	Bahia	Jacobina	Olsthoorn, G
Melocactus zehntneri	(Britton & Rose) Luetzelb.	L25A10	HTSA 7525	Brazil	Ceará	Lavras da Mangabeira	Olsthoorn, G
Melocactus zehntneri	(Britton & Rose) Luetzelb.	S157B2	SORO7988	Brazil	Rio Grande do Norte	Parnamirim	Taylor, N.
Outgroup							
Arrojadoa rhodantha	Britton & Rose	S142A6	SORO 4911	Brazil	Minas Gerais	Matias Cardoso	Olsthoorn, G
Arthrocereus glaziovii	(K.Schum.) N.P.Taylor & Zappi	S162A2	SORO7962	Brazil			
Astrophytum myriostigma	Lem.	S170A23	RBvc 3	Mexico	-	-	D.C. Zappi
Browningia hertlingiana	(Backeb.) Buxb.	S152A7	KEW 32069	Peru	-	-	-

Coleocephalocereus aureus	F.Ritter	S168A4	S168A4	Brazil	Minas Gerais	Jacinto	
Coleocephalocereus goebelianus	(Vaupel) Buining	S162A4	SORO 008164	Brazil	Minas Gerais	Mato Verde	D. B. Mendes, M. Machado
Coleocephalocereus purpureus	(Buining & Brederoo) F.Ritter	S174A10	SPF 80978	Brazil	Minas Gerais	Itinga	Olsthoorn, G
Discocactus bahiensis	Britton & Rose	L25A9	SORO 006564	Brazil	Ceará	Lavras de Mangabeira	Olsthoorn, G
Discocactus boomianus	Buining & Brederoo	L26A13	VIC 43974	Brazil	Bahia	Sento Sé	Olsthoorn, G
Discocactus placentiformis	K.Schum.	L26A10	HRCB 13930	Brazil	Minas Gerais	Salinas	Olsthoorn, G
Discocactus zehntneri	Britton & Rose	L26A12	SORO 007987	Brazil	Bahia	Sento Sé	Olsthoorn, G
Echinocereus pentalophus	Lem.	S170A29	RBvc 252	Mexico	-	-	D.C. Zappi
Ferocactus herrerae	J.G.Ortega	S170A34	RBvc 271	Mexico	-	-	D.C. Zappi

Gymnocalycium horstii	Buining	S170A35	RBvc 262	Brazil	-	-	D.C. Zappi
Harrisia adscendens	(Gürke) Britton & Rose	S142A7	SORO 4910	Brazil	Bahia	Malhada	Perez, MF; Romeiro-Brito, M; Marsola, EM
Opuntia monacantha	(Willd.) Haw.	S170A46	RBvc 183	Brazil	Rio de Janeiro	Itaipu	D.C. Zappi
Parodia magnifica	(F.Ritter) F.H.Brandt	S170A49	RBvc 37	Brazil	-	-	D.C. Zappi
Pilosocereus arrabidae	(Lem.) Byles & G.D.Rowley	S79H2	SORO4903	Brazil	Espirito Santo	São Mateus	na
Selenicereus chrysocardium	(Alexander) Kimnach	S170A55	RBvc 284	Mexico	-	-	D.C. Zappi
Uebelmannia gummifera	(Backeb. & Voll) Buining	S121A5	HURB 112	Brazil	Minas Gerais	Itamarandiba	L.Y.S. Aona

Table S2 Assigned substrates category for each species' natural occurrence.

Species	Substrates
M_violaceus_violaceus	sand
M_violaceus_ritteri	sand; quartz
M_violaceus_margaritaceus	sand
M_glaucescens	sand; sandstone
M_depressus	sand
M_azureus	limestone
M_pachyacanthus_pachyacanthus	limestone
M_levitestatus	limestone
M_ferreophilus	limestone
M_neoviridenscens	limestone
M_sergipensis	limestone
M_zehntneri	limestone; sandstone; quartz; sand; granitic; rocky
M_salvadorensis	granitic; rocky
M_concinnus	quartz; sand
M_broadwayi	rocky
M_estevesii	granitic
M_mazelianus	granitic
M_caroli_linnaei	limestone
M_intortus	limestone
M_macracanthos	limestone
M_harlowii	limestone
M_paucispinus	quartz
M_deinacanthus	granitic
M_ernestii	sandstone; granitic
M_longicarpus	sandstone; granitic
M_amethystinus	sandstone; granitic
M_inconcinnus	sandstone; granitic
M_bahiensis	sandstone; quartz
M_conoideus	quartz; sand
M_braunii	sandstone
M_oreas_oreas	sandstone; granitic; quartz
M_oreas_cremnophilus	sandstone; granitic
M_brederooianus	sandstone
M_schatzlii	rocky
M_matanzanus	serpentine

Table S3. Comparison of *Melocactus* species groups based on morphology (according to Taylor (1991), Taylor & Zappi (2004, 2018), and subsequent updates (Taylor et al. 2022, 2023)) and phylogenetic clades inferred in this study. Supporting morphological characters and the different soil type occurrences are shown.

Taylor (1991); Taylor & Zappi (2004, 2018); Taylor et al. (2022,2023)	**Clades and subclades in this study (Telhe et al. 2024)**	**Supporting morpho-logical characters**	**Substrate types**
M. OREAS Group **a.** *M. oreas* complex – *M. ernestii* (incl. *M. longicarpus*), *M. oreas*. **b.** *M. bahiensis* (incl. *M. amethystinus, M. braunii, M. brederooianus, M. inconcinnus*), *M. conoideus.*	1. M. OREAS clade **a.** M. bahiensis subclade – *M. bahiensis, M. braunii, M. brederooianus, M. conoideus, M. oreas.* **b.** M. ernestii subclade – *M. amethystinus, M. ernestii, M. inconcinnus, M. longicarpus.*	Spines numerous, slender, the centrals much shorter than the lowermost radials; fruits red to magenta at apex; seeds small, periclinal walls of testa cells convex.	Granitic, sandstone, or other crystalline rock.
M. CURVISPINUS Group Species from northern South America, the western Andes, Central America and the Caribbean, except *M. matanzanus, M. neryi, M. smithii*.	2. M. INTORTUS clade Includes most species from northern South America and the Caribbean region, except *M. matanzanus* and its un-sampled but probable allies (*M. neryi, M. smithii*)	Spines few to many, the centrals longer or only slightly shorter than the lowermost radials; fruits red to magenta at apex often large; seeds various.	Vulcanic rocky substrates, and limestone.
M. DEINACANTHUS Group (*M. deinacanthus* only)	3. M. DEINACANTHUS clade *M. deinacanthus* is weakly supported as sister to *M. paucispinus.*	[there are no shared morphological characters supporting this pair of taxa as sister-species].	Granitic; quartz sand or gravel.

M. LEVITESTATUS Group (M. levitestatus only) and M. AZUREUS Group *M. azureus, M. ferreophilus, M. pachyacanthus* (incl. subsp. *viridis* = *M. neoviridescens*).	4. M. LEVITESTATUS clade Includes *M. levitestatus* and all taxa of the M. AZUREUS Group (see left-hand column), plus the more recently described *M. sergipensis.*	Spines few to many, stout, centrals and radials of similar lengths; fruits white or pale pink; seeds often large, testa ± smooth.	Limestone.
[part of M. VIOLACEUS Group, see below]	5. M. ZEHNTNERI clade *M. concinnus, M. matanzanus, M. salvadorensis, M. zehntneri*. (Likely also includes *M. neryi* and *M. smithii* un-sampled for this study.)	Spines short and stout, centrals (when present) and radials of similar lengths; fruits pale pink to magenta; seeds small, periclinal walls of testa cells convex.	Various types (e.g. granitic, limestone, rocky, quartz sand and gravel), including generalist species.
M. VIOLACEUS Group The M. ZEHNTNERI clade (see next column), plus *M. glaucescens, M. paucispinus, M. violaceus* (incl. *M. depressus*).	6. M. VIOLACEUS clade **a.** <u>M. depressus subclade</u> – *M. depressus, M. glaucescens*. **b.** <u>M. violaceus sens. str. subclade</u> (1 sp.).	As above, but fruits small, white, pink or red.	Quartz sand or gravel, and sandstone.

Table S4. Estimated ages for *Melocactus* species and outgroups.

Species Divergence Time (MA)			
Species	**Crown Age**	**CI_Lower**	**CI_Upper**
Melocactus oreas cremnophilus L26A35	0,33212227	0,1009378	1,0928029
Melocactus oreas L26A53	0,33212227	0,1009378	1,0928029
Melocactusbrederooianus L25A12	0,65479147	0,25209744	1,7007386
Melocactus conoideus L86D1	0,98350003	0,42147307	2,2949801
Melocactus braunii L26A20	1,263772	0,57666654	2,7695751
Melocactus bahiensis L86D4	1,4691462	0,69727937	3,0954462
Melocactus ernestii ernestii L26A27	0,81374518	0,33439729	1,9802230
Melocactus ernestii longicarpus L26A29	0,81374518	0,33439729	1,9802230
Melocactus inconcinnus L26A56	1,2026242	0,54376797	2,6597835
Melocactus amethystinus L26A21	1,4061591	0,66299067	2,9823698
Melocactus caroli-linnaei L42A8	0,2672995	0,073070628	0,97780800
Melocactus mazelianus L42A6	0,2672995	0,073070628	0,97780800
Melocactus estevesii L42A2	0,16551233	0,036116969	0,75848921
Melocactus broadwayi L42A5	0,16551233	0,036116969	0,75848921
Melocactus intortus L42A9	0,43451603	0,14826714	1,2734054
Melocactus macracanthos S162A15	0,51620862	0,18843466	1,41413
Melocactus schatlzii L87A4	0,62606412	0,24542154	1,5970736
Melocactus harlowii L42A10	0,73610807	0,30328312	1,7866312
Melocactus paucispinus L26A37	1,6031375	0,75875551	3,3871910
Melocactus deinacanthus S174A19	1,6031375	0,75875551	3,3871910
Melocactus pachyacanthus S162A17	0,70929790	0,27526572	1,8277013
Melocactus azureus L26A52	0,70929790	0,27526572	1,8277013
Melocactus ferreophilus L26A30	1,0188875	0,43520655	2,3853773
Melocactus levitestatus L26A33	1,1279344	0,49931433	2,5479665
Melocactus neoviridescens L26A36	1,3687894	0,63519814	2,9496064
Melocactus sergipensis L25A11	1,6005733	0,77368772	3,3112002
Melocactus zehntneri L25A10	0,81449894	0,33474460	1,9818348
Melocactus concinnus L26A25	0,81449894	0,33474460	1,9818348
Melocactus salvadorensis S165A3	1,0877986	0,4855355	2,4371144
Melocactus matanzanus L87A2a	1,3621506	0,64034024	2,8976071
Melocactus depressus L86D2	1,750852	0,83536818	3,669619
Melocactus glaucescens L8G9	1,750852	0,83536818	3,669619
Melocactus violaceus violaceus L26A55	0,73236586	0,29873479	1,795437
Melocactus violaceus margaritaceus L26A3	0,73236586	0,29873479	1,795437
Melocactus violaceus ritteri L26A40	1,5332450	0,73628926	3,1928219
Discocactus placentiformis L26A10	1,7156092	0,77828012	3,7818198
Discocactus bahiensis L25A9	1,7156092	0,77828012	3,7818198
Discocactus zehntneri L26A13	2,2084729	1,0569886	4,6143853
Discocactus zehntneri L26A12	2,4496190	1,2234794	4,9045643
Coleocephalocereus aureus S183A2	1,3861687	0,64766464	2,9667573

Coleocephalocereus purpureus S174A10	1,3861687	0,64766464	2,9667573
Coleocephalocereus goebelianus S162A4	3,2762193	1,6368598	6,5574418
Pilosocereus arrabidae S79H2	3,8489590	1,9417605	7,6294088
Arrojadoa rhodantha S142A6	3,8489590	1,9417605	7,6294088
Gymnocalycium horstii S170A35	6,0698873	3,3421729	11,023826
Ferocactus herrarea S170A34	11,169582	8,0597019	15,479426
Astrophytum myriostigma S170A23	11,169582	8,0597019	15,479426
Selenicereus chrysocardium S170A55	10,305810	6,9457911	15,291236
Echinocereus pentalophus S170A29	10,305810	6,9457911	15,291236
Parodia magnifica S170A49	13,806540	9,9952322	19,071149
Uebelmannia gummifera S121A5	7,4675029	4,6184137	12,074189
Browningia hertlingiana S152A7	6,9169535	4,1263985	11,594674
Arthrocereus glaziovii S162A2	4,8881952	2,614581	9,1389197
Harrisia adscendens S142A7	4,8881952	2,614581	9,1389197

Table S5. Phylogenetic signal (Pagel's lambda) estimates for twenty-four bioclimatic variables. Asterisk indicates variables with phylogenetic signal > 0.7.

Phylogenetic signal (lambda)		
Variable	**P-value**	**phylogenetic signal (lambda)**
bio_6*	1.77E-05	0.9293332093
bio_11*	2.78E-05	0.8924406665
maxTempColdest*	0.001131826962	0.7829416706
PETWettestQuarter*	0.008520676022	0.7722242029
bio_9*	0.0008393638827	0.7417647031
PETseasonality	0.0002393628578	0.6635973372
continentality	0.0006863078441	0.6590418827
bio_7	0.000272590991	0.6567730093
bio_4	0.0005130781947	0.6456830465
bio_1	0.002348022111	0.6431892705
current_30arcsec_thermicityIndex	0.004591174826	0.5813368443
bio_15	0.04639801988	0.5698450676
current_30arcsec_minTempWarmest	0.008570921512	0.5451552115
current_30arcsec_growingDegDays0	0.008016196445	0.5426683334
current_30arcsec_growingDegDays5	0.008016196445	0.5426683334
bio_2	0.009781777574	0.5376155645
bio_14	0.00728752689	0.4911694818
current_30arcsec_climaticMoistureIndex	0.008014080685	0.4642029609
bio_10	0.02929165683	0.4405435259
current_30arcsec_embergerQ	0.004301154726	0.426031511
current_30arcsec_topoWet	0.01852879211	0.3999220719
bio_17	0.0166979711	0.3923210544

bio_12	0.01346607511	0.3775221664
bio_13	0.02440479653	0.3532210672

Table S6. Spearman correlation estimation among five climatic variables in which phylogenetic signal was significant.

Spearman Correlation Test				
maxTempColdest	**PETWettestQuarter**	**bio_11**	**bio_6**	**bio_9**
1	0.18642134703309823	0.9273519568824656	0.8525013228432753	0.8700473567253105
0.18642134703309823	1	0.025677429876163043	-0,07355565762	-0,06785915549
0.9273519568824656	0.025677429876163043	1	0.9744373112989961	0.9632027994035204
0.8525013228432753	-0,07355565762	0.9744373112989961	1	0.9575060242143294
0.8700473567253105	-0,06785915549	0.9632027994035204	0.9575060242143294	1

Table S7. Marginal probabilities of discrete ancestral state reconstruction for each node. Bold values are the highest marginal probable state for each node.

Node number	Granitic	Limestone	Quartz	Mix Rocky	Sand	Sandstone	Serpentine
1	5.57E-08	8.34E-05	0.004756	8.22E-09	**0.994971**	6.83E-05	0.000121
2	4.34E-05	5.60E-05	0.000555	4.11E-05	**0.996656**	0.002624	2.50E-05
3	1.41E-05	5.31E-05	0.000155	9.04E-06	**0.999237**	0.000515	1.65E-05
4	1.92E-06	1.27E-05	1.44E-06	1.06E-06	**0.999923**	5.93E-05	7.29E-08
5	2.59E-05	6.18E-05	0.000275	1.97E-05	**0.996921**	0.002676	1.99E-05
6	7.31E-10	6.51E-07	**0.968945**	2.79E-10	0.011842	2.33E-07	0.019212
7	2.80E-10	3.84E-11	**0.982683**	1.77E-10	0.000112	9.02E-05	0.017115
8	8.65E-05	1.80E-09	**0.904686**	0.000218	4.18E-05	1.45E-06	0.094966
9	**0.512824**	1.61E-06	0.020644	0.44828	5.97E-08	4.61E-11	0.01825
10	**0.521693**	1.95E-09	0.006847	0.464057	2.98E-08	1.51E-12	0.007403
11	**0.61324**	7.32E-07	0.00055	0.385389	8.59E-11	1.42E-14	0.00082
12	**0.634504**	8.92E-07	7.22E-05	0.365286	3.90E-10	2.12E-15	0.000138
13	**0.603093**	2.35E-11	1.91E-06	0.3969	1.06E-10	2.25E-18	5.02E-06
14	**0.521274**	3.64E-13	2.17E-07	0.478725	1.55E-10	3.85E-18	5.97E-07
15	**0.655697**	5.33E-07	4.61E-07	0.3443	7.19E-12	5.24E-18	1.48E-06
16	0.000376	1.14E-06	**0.833193**	0.000394	3.48E-05	4.33E-06	0.165997
17	5.17E-10	2.50E-11	0.13953	4.46E-10	7.38E-07	**0.858755**	0.001714
18	1.14E-10	6.42E-10	0.041219	3.85E-11	1.73E-06	**0.95766**	0.00112
19	7.04E-11	1.92E-10	0.019942	3.98E-11	2.41E-06	**0.979492**	0.000563

20	1.74E-09	4.71E-09	0.016483	1.15E-09	0.000268	**0.982814**	0.000434
21	1.39E-08	5.24E-12	0.002805	1.07E-08	2.01E-06	**0.997098**	9.47E-05
22	0.000176	5.56E-13	0.001359	0.000269	7.71E-07	**0.998116**	8.04E-05
23	5.35E-05	1.11E-10	0.022702	7.86E-05	4.41E-07	**0.976121**	0.001044
24	0.000133	7.92E-11	0.009415	0.0002	6.14E-07	**0.989426**	0.000826
25	0.000514	8.72E-11	0.00324	0.000655	1.83E-06	**0.99496**	0.000629
26	0.00014	0.000108	**0.849759**	0.000141	0.000679	1.72E-08	0.149172
27	0.323589	0.174823	0.119518	**0.326977**	3.72E-06	1.25E-09	0.055089
28	0.305303	**0.453622**	0.020394	0.206702	4.26E-08	8.71E-11	0.013979
29	0.190447	**0.697461**	0.001979	0.108173	6.30E-09	1.68E-12	0.00194
30	0.145089	**0.773434**	0.000631	0.080122	1.24E-08	9.74E-13	0.000724
31	0.064758	**0.901316**	6.48E-05	0.033779	2.66E-08	2.82E-13	8.15E-05
32	4.25E-05	6.19E-06	**0.714887**	2.86E-05	0.00048	2.85E-06	0.284553
33	0.007591	4.72E-06	**0.716805**	0.009402	0.000543	2.76E-05	0.265626
34	0.009122	0.010079	**0.755303**	0.006584	0.003668	0.008981	0.206263

CHAPTER 2: Phylogenomic systematics of *Discocactus* (Cactaceae): trait evolution and diversification in arid and semi-arid habitats.

Milena C. Telhe[1], Nigel P. Taylor[1], Monique Romeiro-Brito[2], Daniela C. Zappi[3], Gerardus Olsthoorn[4], Fernando F. Franco[1], Evandro M. Moraes[1].

1. Departamento de Biologia. Universidade Federal de São Carlos, Sorocaba, São Paulo 18052-780, Brazil.

2. Natural History Department, Florida Museum of Natural History, Gainesville, FL 32611, USA.

3. Programa de Pós-Graduação em Botânica, Instituto de Ciências Biológicas, Universidade de Brasília (UNB), Brasília 70919-970, Brazil.

4. Community scientist, Caixa Postal 169, Holambra, São Paulo, 13825-000, Brazil

Draft manuscript

Abstract

Discocactus is a genus of Cactaceae known for its controversial taxonomic history. The number of species within this genus ranges from 11 to 52, depending on the classification system proposed by different authors. These discrepancies highlight the challenges of species delimitation based solely on morphological characteristics within the genus, which complicates the definition of conservation status and the development of strategies for threatened species. *Discocactus* species are primarily distributed across the Neotropical Savanna biome, with a few species occurring in the Seasonally Dry Tropical Forest. However, it is still unclear which key traits might have enabled the *Discocactus* species to establish in different biomes and how environmental differences have affected its evolution. Here, we used the Cactaceae591 probe set to capture loci across the genus *Discocactus* and estimate the first well-resolved phylogenomic hypothesis using both concatenated and coalescent methods. Our phylogenetic hypothesis agrees with the taxonomic circumscription of *Discocactus* which recognizes the morphological variation in certain species as infraspecific taxa. By integrating predictive machine learning with phylogenetic comparative methods, we identified that flower length and stem diameter are associated with differences in environmental conditions. This may have enabled *Discocactus* species to thrive in diverse environmental conditions and colonize two different biomes.

Keywords: adaptive traits, Cactaceae591, machine learning, phylogenetic comparative methods.

3.1 Introduction

Discocactus Pfeiff., as suggested by its name, is a genus of discoid-shaped cacti (Fig. 1) comprising around 15 species (*sensu* Flora e Funga do Brasil, 2025), in which at least 11 are classified as Endangered or otherwise threatened (IUCN, 2024). This genus is primarily distributed in Eastern Brazil, with three species extending their distribution from central-western Brazil to north-eastern Paraguay, and eastern Bolivia (Taylor & Zappi, 2004; Hunt et al., 2006). The earliest description of a species from this genus was in 1826 by Lehmann, who described the *Cactus placentiformis* (later incorporated into the *Discocactus* as *D. placentiformis* (Lehm.) K. Schum.). However, the genus *Discocactus* was established a few years later, in 1837, by Louis Pfeiffer. After the publication of the generic name, many species were described, especially from the 1970s (Santos, 2013 and references therein). All these descriptions are primarily based on morphological characters, but the variations described often overlap among species, as well as their geographical distribution (Machado et al., 2005). Depending on the author's classification system, the number of *Discocactus* species ranges from 11 to 52 (Buining, 1980; Hunt et al., 2006; Santos, 2013 and references therein). This controversial classification reflects the challenge of species delimitation in this genus.

Similar to other cactus genera (see Guerrero et al., 2019), efforts to resolve the phylogenetic relationships of *Discocactus* using a limited number of traditional molecular markers were unsuccessful due to a lack of phylogenetic signal (Santos, 2013). In recent years, significant advancements have been made in the use of genome-scale sequence data for phylogenetic analyses (McKain et al., 2018; Smith et al., 2020). Indeed, such data has improved our understanding of species relationships and circumscriptions in the cactus family (Bombonato et al., 2020; Merklinger et al., 2021; Romeiro-Brito et al., 2023; Taylor et al., 2023; Zappi et al., 2024). The same approach would enhance our understanding of species relationships and classifications within the genus *Discocactus*, contributing to solving taxonomic conflicts.

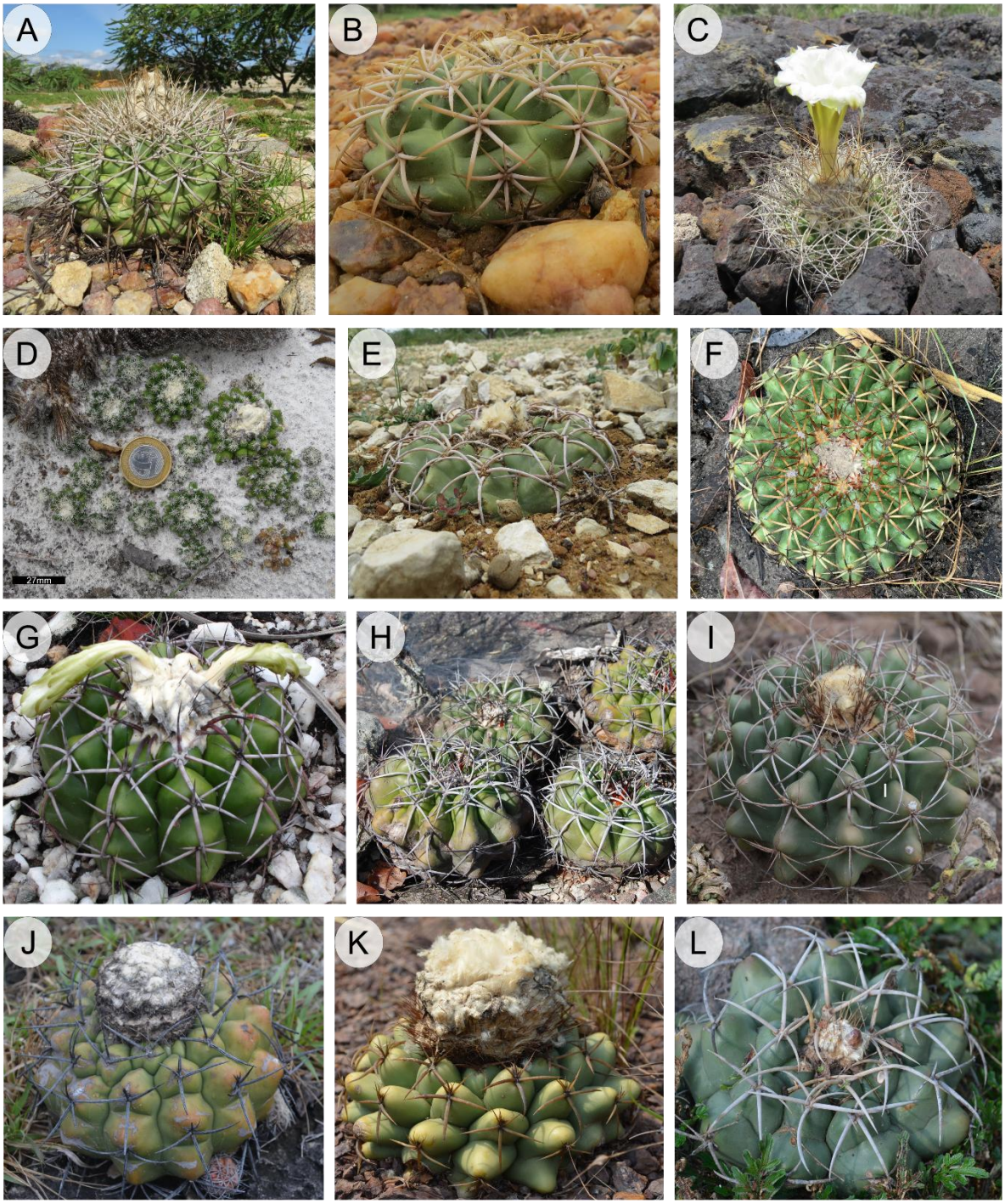


Figure 1. Representatives of genus *Discocactus*. A) *D. zehntneri.* B) *D. petr-halfari.* C) *D. boomianus.* D) *D. buenekeri*, a Brazilian coin of one Real (diameter: 27 mm) is used as a scale. E) *D. bahiensis.* F) *D. fariae-peresii*. G) *D. placentiformis* H) *D. diersianus*. I) *D. hartmannii.* J) *D. boliviensis.* K) *D. heptacanthus.* L) *D. catingicola.* Photo credits: A to E - Telhe, M. C; F - Zappi, D. C.; G to L - Olsthoorn, G.

Discocactus is one of the few cactus genera predominantly found in the Savanna biome of Brazil, particularly in the Cerrado domain, although some species also inhabit the SDTF within the Caatinga domain (Taylor & Zappi, 2004). Most cactus genera do not occur in Neotropical Savannas due to frequent fires that often cause overheating of their succulent tissues, leading to mortality or severe damage (Barthlott et al., 2015; Ritz et al., 2007; Taylor & Zappi, 2004). The occurrence of *Discocactus* across the fire-prone Cerrado domain is likely due to their fire-adapted traits. The discoid form and ground-hugging habit of *Discocactus* help shield the plant from the most intense effects of burning, as the region right above the ground tends to be cooler during a fire event due to drawn-in air resulting from the convection of the heated gases (Whelan, 1995; Taylor & Zappi, 2004).

Studies investigating plant lineages that diversified across arid and semi-arid neotropical habitats, such as SDTF and Savanna, often focus on the lineage's adaptation to fire. This is because savannas regularly experience fires, which can act as a significant environmental filter, whereas SDTFs do not (e.g. Simon et al., 2009; Silva & Batalha, 2010; Simon & Pennington, 2012 and references therein). Although it represents an important limiting factor, fire may not be the sole environmental filter affecting these lineages.

Despite both habitats being classified as semi-arid or arid and experiencing seasonal stress, they differ significantly in climate, seasonality, and temperature gradients. The Caatinga domain, an SDTF biome, remains arid year-round with 300 to 800 mm of erratic rainfall in a short, unpredictable rainy season and temperatures ranging from 20°C to 32°C, experiencing higher hyperaridity levels during winter. In contrast, the Cerrado domain, a Savanna biome, receives 1,200 to 1,800 mm of rain annually during a distinct wet season from October to March, with temperatures ranging from 18°C to 28°C. These differences in precipitation, aridity, and temperature gradient can hinder plant species establishment and distribution (Ringelberg et al., 2023). While the fire-adapted traits in *Discocactus* are well recognized, the influence of other environmental filters, such as climate and temperature, on the diversification and evolution of *Discocactus* across these biomes remains unclear.

Identifying functional traits and their adaptive value is a key focus in plant biology and ecology (Caruso et al., 2020 and references therein). However, determining which traits influence plant adaptation, and testing hypotheses regarding their ecological and evolutionary aspects can be challenging (Anderegg, 2023). Recently, machine-learning approaches have emerged as valuable exploratory tools for identifying potential adaptive traits (e.g. Pigot et al., 2020; Lanna et al., 2022). These methods enable researchers to explore a wide range of traits, both continuous and categorical

(Simon et al., 2023), allowing the prediction of traits that may have enabled the colonization of different habitats, thus enhancing our understanding of species diversification.

Here, we present the first comprehensive phylogenomic hypothesis for the genus *Discocactus*, clarifying species relationships and classification. Additionally, we employed predictive machine learning and phylogenetic comparative methods to identify key traits that might have facilitated the colonization of diverse habitats within the semi-arid and arid Neotropical regions, providing new insights into the diversification processes experienced by this genus.

3.2 Methods

3.2.1 Genomic data collection and Bioinformatics

Our ingroup sampling comprises 29 specimens, including all the 15 *Discocactus* species currently recognized (*sensu* Hunt et al., 2006; Flora e Funga do Brasil, 2025). Initially, the sampling was designed to account for all geographical distributions and variations described for each species, however, due to the COVID-19 pandemic sanitary crises, field expeditions and sampling were limited. Whenever feasible, we obtained at least two specimens for each species, trying to cover the widest distribution possible. Additionally, for species with a larger number of infraspecific taxa described, we collected at least one of each, seeking to capture the morphological variation among populations. For instance, we ensured the representation of four infraspecific taxa from *D. heptacanthus* (Rodrigues) Britton & Rose (see Table S1). Our outgroup comprises one genus of the subfamily Opuntioideae and 14 genera from the subfamily Cactoideae (see suppl. Table S1).

Genomic DNA was extracted from root tissue using high-salt CTAB 3% protocol (Romeiro-Brito et al. 2023) for all samples. Library preparation for the enriched loci targeted by the Cactaceae591 probe set (Romeiro-Brito et al., 2022), as well as the sequencing on an Illumina NovaSeq 6000 platform (2 × 150 bp), were performed by RAPiD Genomics (Gainesville, Florida, USA).

For all raw Cactaceae591 reads, we removed adapters, low-quality reads (phred < 20) and short sequences (reads < 60bp) using *fastp* v.0.23.3 (Chen et al. 2018). Next, we mapped and assembled the reads using HybPiper v. 2.1.6 with the *--run_intronerate* which allows the recovery of supercontigs (on-target + flanking nuclear regions; Johnson et al., 2016). Though the Cacataceae591 probe set was designed for single-copy regions, some regions may be duplicated for some taxa in our sampling. Thus, we excluded all putative paralogs identified following the pipeline used by Romeiro-Brito et al. (2022). Multiple sequence alignments were performed using MAFFT v. 7.520 (Katoh & Standley, 2013), and the poorly aligned sequences were trimmed using

TrimAL v. 1.4 (Capella-Gutiérrez et al., 2009) with the function *-strict.* This function is an automatic algorithm that performs a *gappyout* trimming with a subsequent trimming based on similarity threshold. Considering the effects of missing data in the following analysis, we only allowed regions with up to 20% of missing samples in our final dataset.

3.2.2 Phylogenetics inferences and divergence time estimations

Phylogenetic inferences were carried out using two different approaches, the maximum likelihood (ML) and the multispecies coalescent (MSC). Maximum likelihood phylogenetic inferences were performed using IQ-TREE v. 2.2 (Minh & al., 2020) for two different datasets: (i) a dataset with all loci concatenated in a supermatrix; (ii) for each locus independently (single locus). For each ML inference, the best substitution model was selected using the ModelFinder Plus option, and branch support was measured using the ultrafast bootstrap (Hoang et al., 2017) with 10,000 replicates. The MSC inference was performed using the weighted ASTRAL algorithm (wASTRAL) with the function *astral-hybrid*, which accounts for both low bootstrap values and branch lengths (wASTRAL; Zhang & Mirarab, 2022). For this analysis, the ML gene trees were used as input, and node support bootstrap values under 70 were considered low. The wASTRAL weights each gene tree quartet based on the branch lengths and node supports, reducing the impact of inconsistent quartets in the recovery of the species tree (Zhang & Mirarab, 2022).

Divergence time estimation was performed on MEGA11 (Tamura et al., 2021) using the RelTime method (Tamura et al., 2012) which is based on the relative rate framework. Because there are no fossil records for the Cactaceae family, a secondary calibration was performed assuming the ages previously estimated by Hernandez-Hernandez et al., (2014) to the subfamily Cactoideae (crown age: 10.94 – 21.85 Ma) with a normal distribution (mean: 16.13; SD: 2.90). This analysis was performed fixing the MSC tree topology and using all loci recovered in the final dataset as inputs.

3.2.3 Trait data collection and biome definition

We remotely collected geographic occurrences from SpeciesLink (https://specieslink.net/) and GBIF (https://www.gbif.org/) databases and eliminated spatial errors using the 'CoordinateCleaner' package (Zizka et al., 2019) in R v.4.3.1 (R Core Team, 2023). Additional filtering was applied by plotting the occurrences of each *Discocactus* species on a map and manually excluding records outside their known distribution. The retrieved occurrences were then classified according to their biome occurrence based on the Ecoregions 2017© Resolve database

(Dinerstein et al., 2017) using the R package 'speciesgeocodeR' (Töpel et al., 2017). To categorize the biome occurrence of each *Discocactus* species we used the biomes Seasonally Dry Tropical Forests (herein shortened as SDTF) and Tropical & Subtropical Grasslands, Savannas & Shrublands (herein shortened as Savanna) delimited by Dinerstein et al. (2017). A species was considered present in a biome if at least 50% of its occurrences were within that biome. The classification was then refined using habitat characteristics and known biome distribution from the literature (e.g. Taylor & Zappi, 2004; Flora e Funga do Brasil, 2025).

The biotic dataset was obtained for each *Discocactus* species based on the collection of morphological data records, such as stem, flower, and seed sizes, rib patterns, spine color, along with ecological variables (details Table S2) through a literature survey and taxonomic descriptions of currently recognized species (e.g., Taylor and Zappi, 2004; Hunt et al., 2006; Lowry, 2016, Taylor and Zappi, 2018).

3.2.4 Identification of key predictors of species occurrence

Morphological data were analyzed using the random forest (RF) method (Breiman, 2001) and the 'caret' package (Kuhn, 2008) in R v. 4.3.1 to predict species occurrence in SDTF and Savanna biomes. The RF method, a supervised machine-learning algorithm, builds multiple decision trees to evaluate the predictive power of specific traits on the desired response (Kuhn, 2008). This approach allows for the simultaneous exploration of various traits, regardless of their distribution (e.g., binary, multi-categorical, or continuous), even when the most important trait or combination of traits is uncertain (Breiman, 2001).

Since the RF methods do not accommodate missing data, we employed two strategies for the analysis: (i) an original dataset where any species with missing values were excluded, and (ii) an inputted dataset in which all species were retained, and the missing values were predicted using the 'missForest' package (Stekhoven, 2011) in R. To avoid circularity, the response variable (biome) was excluded from the dataset before implementing the 'missForest' function.

Both datasets (original and inputted) were used to build and evaluate the predictive models with the 'caret' package, using 10,000 trees, the leave-one-out cross-validation (LOOCV) to assess the model quality, and the mean decrease accuracy (MDA) to measure variable importance. In addition to MDA, we assessed variable importance using the package 'Boruta' (Kursa & Rudnicki, 2010). This package enables the evaluation of important variables, adding a set of variables uncorrelated to the response variable. First, this algorithm obtains the uncorrelated variables by

shuffling the values of the original dataset, then, it compares the importance of the model variables with the uncorrelated random variables, known in the literature as "shadow variables" (Kursa & Rudnicki, 2010). Only the variables with importance greater than the most important shadow variable were considered important predictors. For cases where variable importance was unresolved, we used the function *TentativeRoughFix* in 'Boruta' to confirm or reject them. Finally, variables identified as important predictors by both RF predictive model and 'Boruta' analysis were reconstructed and mapped onto the *Discocactus* phylogeny using the *contmap* function in the 'phytools' package (Revell, 2012).

To investigate the correlation between traits recovered as important predictors in our RF model and the environmental characteristics of the biomes where each species occurs, we conducted a phylogenetic linear regression analysis. We extracted 30-second resolution climatic variables from the WorldClim2.1 (Fick & Hijmans, 2017) and Envirem (Title, & Bemmels, 2018) databases for each species, using the occurrence dataset previously described (see Trait data collection and biome definition section). We selected annual mean temperature (bio1), isothermality (bio3), annual precipitation (bio12), precipitation seasonality (bio15), and aridity index considering the critical role of precipitation and temperature in the composition of Neotropical flora (Santos et al., 2012; Neves et al., 2015). Moreover, despite both biomes (SDTF and Savanna) being seasonal, they differ in temperature, precipitation, and aridity levels (Luebert, 2021). We estimated the correlation between the mean climatic values for each taxon with important biological traits using the package 'phylolm' (Ho et al., 2018) in R.

3.3 Results

3.3.1 Genomic sampling, phylogenetic relationships and divergence time

After data processing, removing putative paralogs, and excluding regions with more than 20% missing samples, our final dataset (supercontig) includes 29 *Discocactus* specimens across 415 genomic regions. This dataset comprises 681,216 bp, of which 153,238 are variable sites, and 55,281 are parsimoniously informative sites. Both ML and MSC approaches supported *Discocactus* as a well-resolved and monophyletic genus (LPP > 0.9 and UFBoot: 100; Fig. 2). The main incongruence between ML and MSC trees is regarding the position of *Discocactus boomianus* Buining & Brederoo. While in the ML tree this species is recovered as a sister of the remaining *Discocactus* species, in the MSC tree this position is replaced by the clade composed of *D. zehntenri* Britton & Rose and *D. petr-halfari* Zachar, rendering *D. boomianus* as an inner lineage.

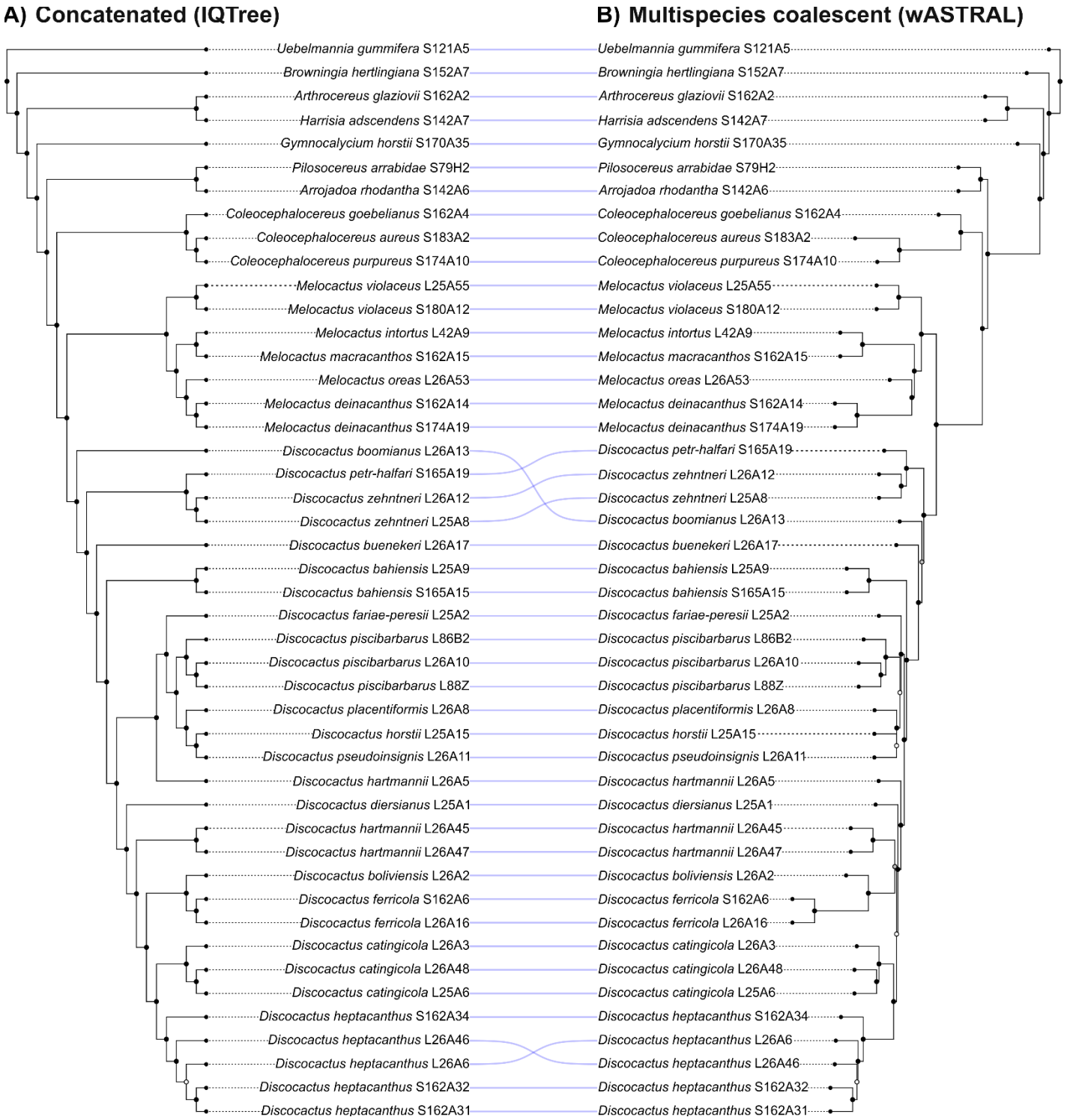


Figure 2. Comparison of *Discocactus* phylogenies inferred using **(A)** the concatenated maximum likelihood and **(B)** the multispecies coalescent approach. Nodes are categorized by support level: highly-supported branches (LPP ≥ 0.95; UFBoot ≥ 95) are marked with black circles, moderately-supported branches (LPP 0.94–0.75) with gray circles, and low-supported branches (LPP < 0.75;

UFBoot < 95) with white circles. The sample code of each specimen (Table S1) is indicated after the species names.

Both approaches consistently recovered *D. bahiensis* Britton & Rose as a sister to the Savanna clade, which comprises the species that are endemic to Savanna (LPP: 1, UFBoot: 100). Additionally, a well-supported clade comprising *D. fariae-peresii* P.J.Braun, *D. placentiformis* (Lehm.) K.Schum., *D. piscibarbarus* N.P.Taylor & Olsthoorn, *D. pseudoinsignis* N.P.Taylor & Zappi and *D. horstii* Buining & Brederoo was consistently recovered (LPP: 0.99; UFBoot:100). Within this clade, *D. fariae-peresii* is recovered as a sister lineage to all other species from this clade. *Discocactus horstii* and *D. pseudoinsignis* are consistently recovered as sister lineages. Moreover, all phylogenetics approaches recovered *D. hartmannii* (K.Schum.) Britton & Rose (L26A5) from Paraguay as distinct from the Brazilian *D. hartmannii* specimens (LPP:0.99; UFBoot:100). Regardless of the approach, *D. diersianus* is recovered as a sister to the clade comprising *D. hartmannii, D. boliviensis, D. ferricola, D. catingicola, and D. heptacanthus*. *Discocactus ferricola* and *D. boliviensis* were consistently recovered as sister lineages (LPP:1; UFBoot:100), as well as *D. heptacanthus* and *D. catingicola*.

The RelTime estimation places the crown age of *Discocactus* in the Pliocene (95% CI: 4.96–2.93 million years ago [Ma]; suppl. Table S4; Figure 3). Most lineages within the genus, particularly those endemics to the Savanna, originated along the Plio-Pleistocene transition and early Pleistocene (e.g., the Savanna clade, mean: 2.6 Ma, Figure 2).

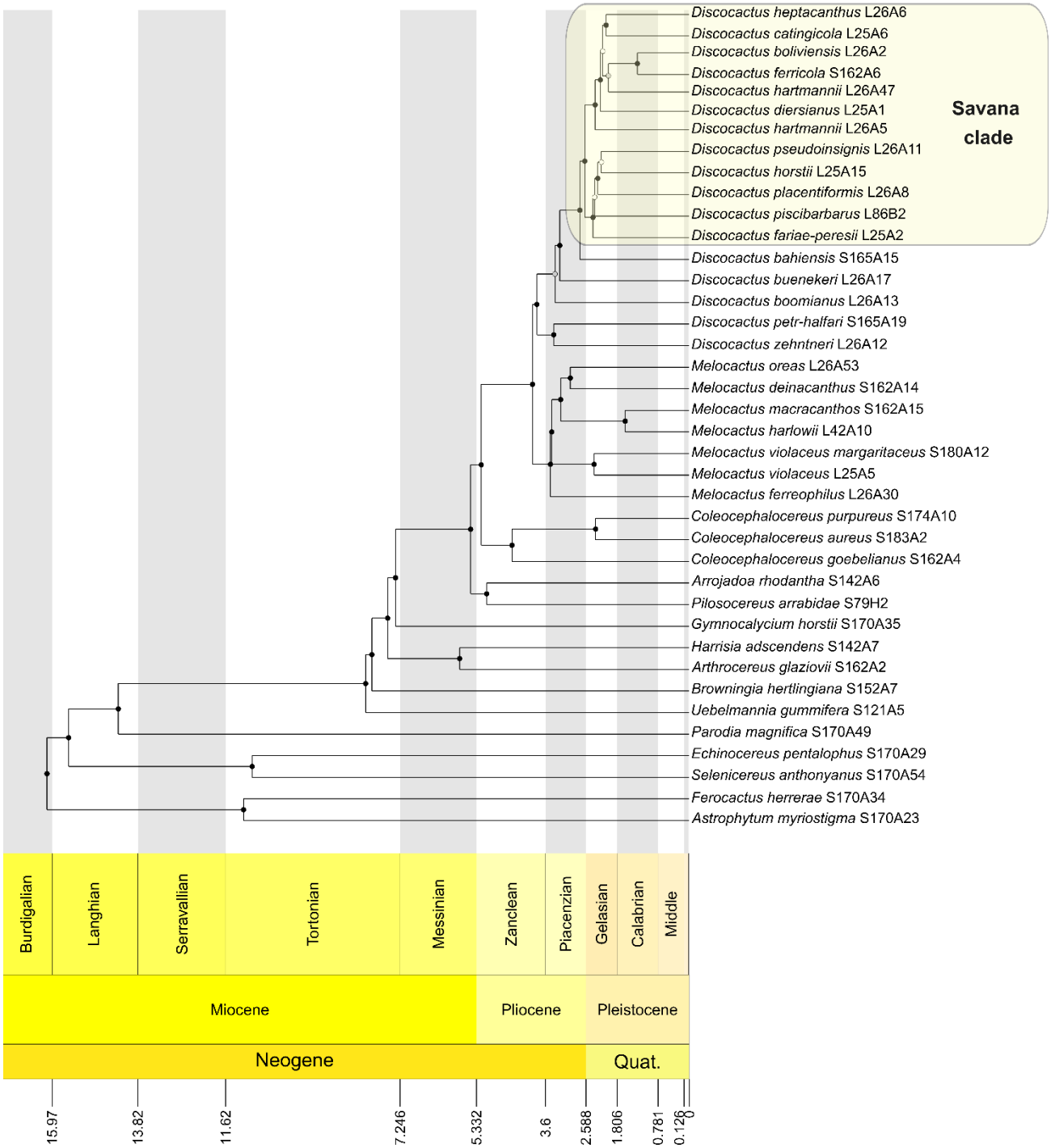


Figure 3. Time-calibrated tree of *Discocactus* inferred with the RelTime method in MEGA11. Branches with high support (LPP ≥ 0.95) are indicated by black circles, those with moderate support (LPP 0.94–0.75) by gray circles, and those with low support (LPP < 0.75) by white circles. The sample code of each specimen (Table S1) is indicated after the species names.

3.3.2 Predictive analysis and ancestral trait reconstruction

The final biotic dataset assessed consisted of 22 traits (see Table S2). For the model using the original dataset, the optimal *mtry* value – representing the number of traits considered at each node in the decision trees – was 2, while for the model using the inputted dataset this value was 12. The accuracy of these models was 0.50 and 0.52, respectively. Regardless of the dataset, our models and 'Boruta' analysis identified stem diameter and flower length as the traits with the greatest predictive power to a species occurrence in SDTF or Savanna biomes (Figure 4).

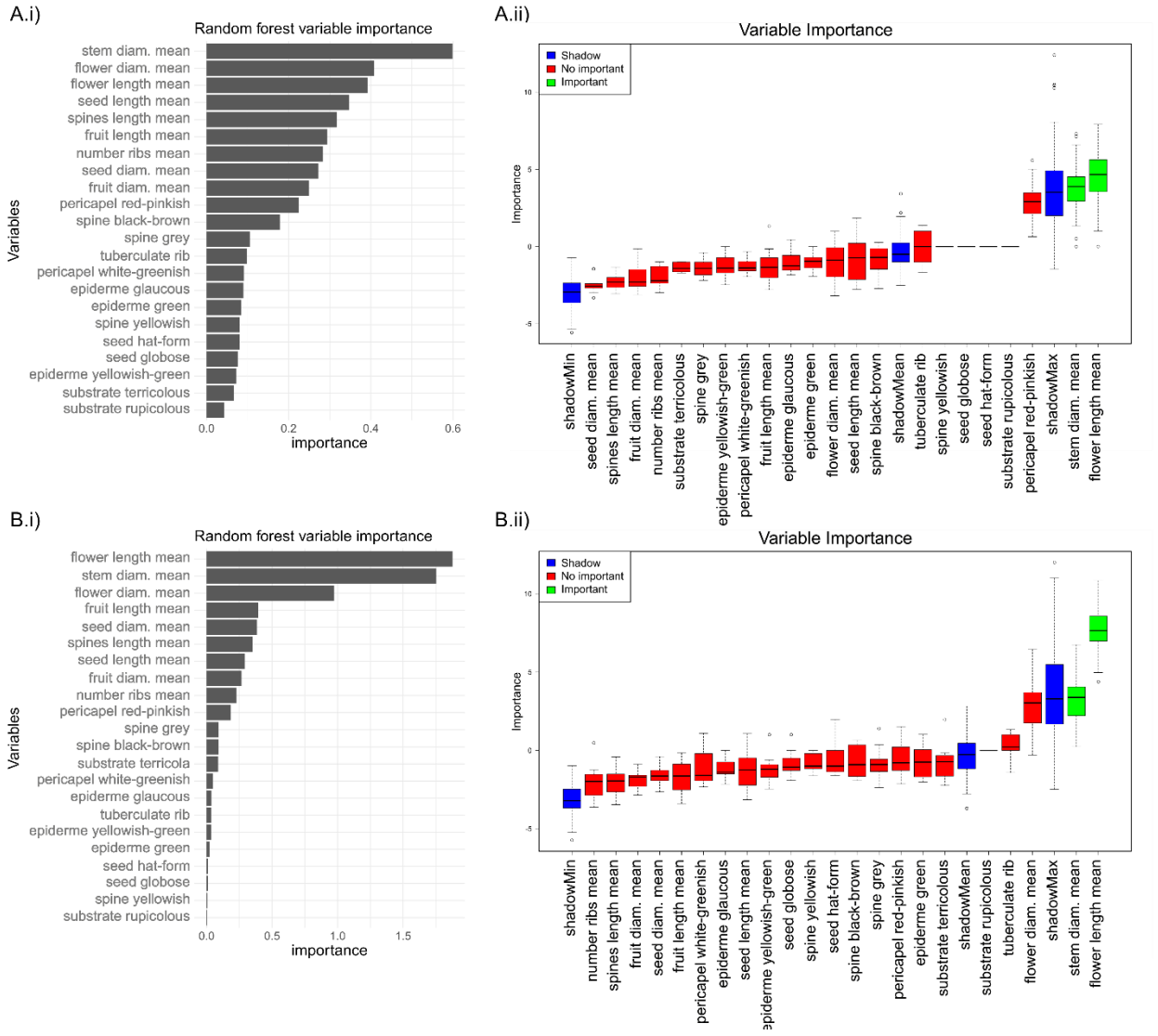


Figure 4. Key predictor traits of *Discocactus* species occurrence in SDTF or Savanna biomes, based on mean accuracy decrease (A.i and B.i), and *boruta* analysis (A.ii and B.ii), estimated with the random forest classification model. Results are shown for both the original dataset (A) and the inputted dataset (B).

The continuous ancestral trait reconstruction for key predictive variables is shown in Figures 5 and 6. Median values were recovered as the ancestral state for flower length, with lineages endemic to SDTF having the shortest flower length and the Savanna endemic lineages having the largest flower length (Figure 5.A). Nonetheless, five of the twelve lineages in the Savanna clade exhibit medium to shorter flower lengths. For the stem diameter, the ancestral reconstruction shows lower values as the ancestral state (Figure 6.A), which is retained by early diverging lineages. A shift to medium stem diameters is observed during the divergence of *D. bahiensis* and the Savanna clade. This clade, comprising most of the Savanna endemic lineages, exhibits the largest stem diameters. The only exception is *D. horstii,* which is endemic to Savanna but exhibits a smaller stem diameter.

When contrasting the flower length and stem diameter with the climatic conditions where each *Discocactus* species occurs, significant correlations were found between them. We recovered a statistically significant correlation ($p < 0.05$) between flower length and isothermality, annual precipitation, precipitation seasonality, and aridity ($p < 0.05$; Figure 5). However, the correlation between flower length and the annual mean temperature was not statistically significant ($p = 0.21$). For stem diameter, significant correlations were observed between isothermality and annual precipitation ($p < 0.05$; Figure 6), but no significant correlations were found for other climatic variables. Overall, both stem diameter and flower length are smaller when environmental conditions are hotter and drier (Figures 5 and 6).

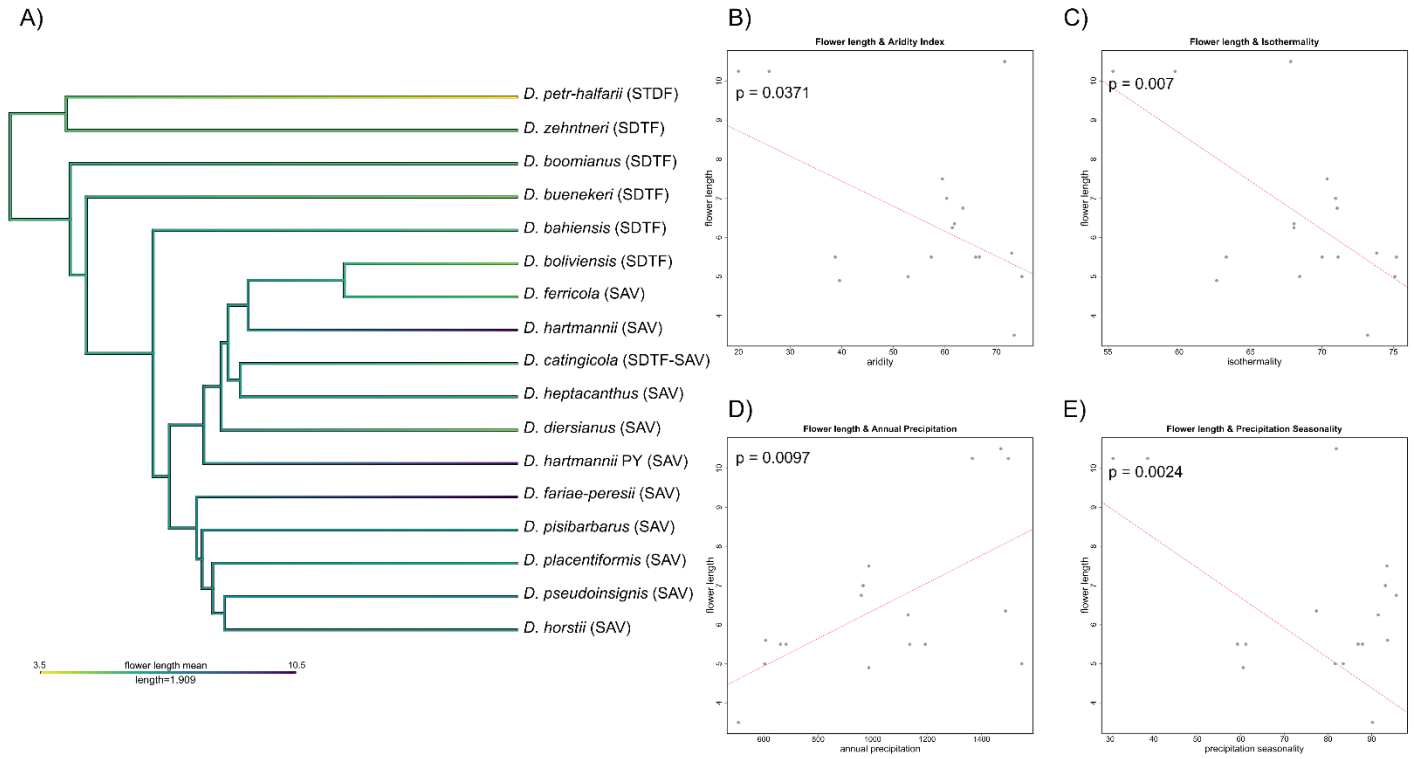


Figure 5. A) Continuous ancestral character state reconstruction for flower length. Phylogenetics regression of flower length only for the statistically significant variables: B) aridity index, C) isothermality (bio3), D) annual precipitation (bio12), E) precipitation seasonality (bio15). Biome classifications for each species are provided after their names (SAV, Savanna; SDTF, dry forest). The *D. hartmannii* specimen from Paraguay is followed by the PY (Paraguay abbreviation) to its differentiation.

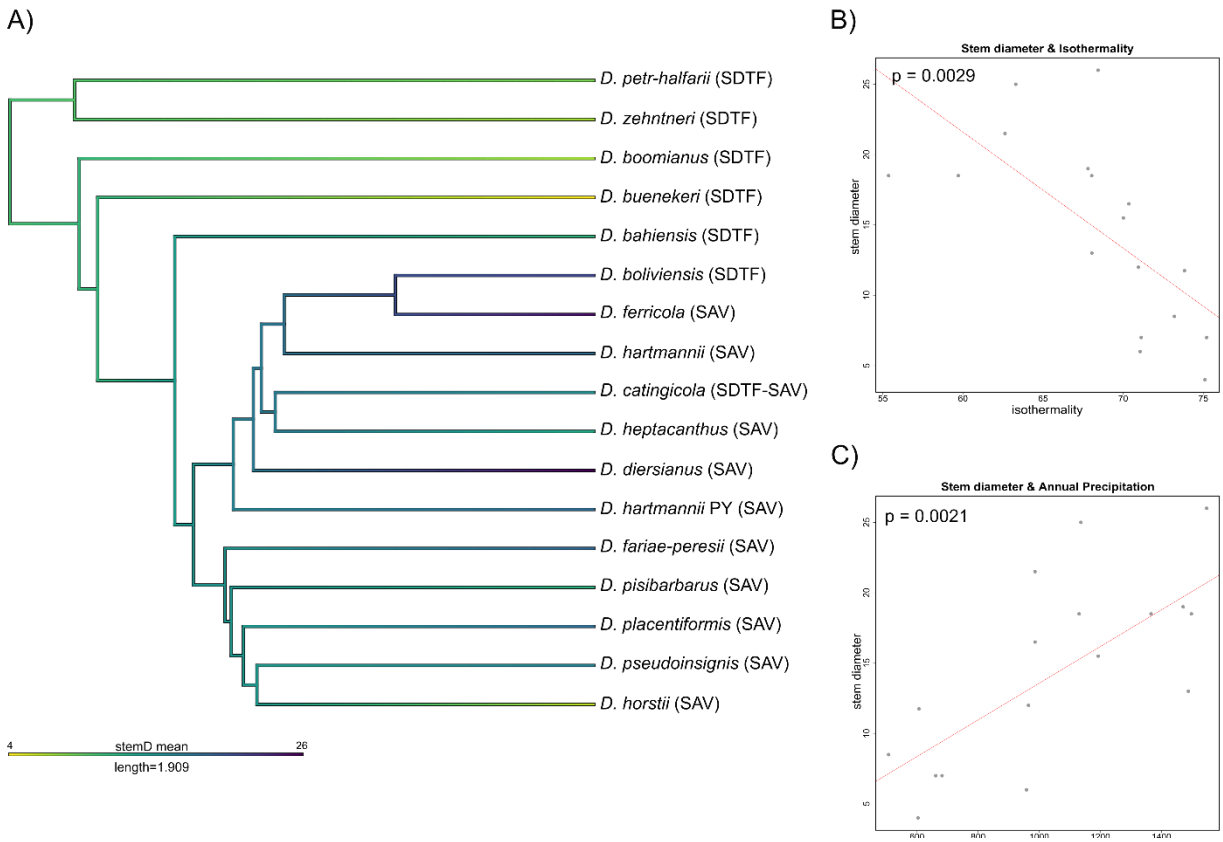


Figure 6. A) Continuous ancestral character state reconstruction for stem diameter. Phylogenetics regression of stem diameter for only the statistically significant variables: B) isothermality (bio3), C) annual precipitation (bio12). Biome classifications for each species are provided after their names (SAV, Savanna; SDTF, dry forest). The *D. hartmannii* specimen from Paraguay is followed by the PY (Paraguay abbreviation) to its differentiation.

3.4 Discussion

3.4.1 Phylogenetic inferences and divergence time in Discocactus

This work represents the first robustly supported species-level phylogeny for the genus *Discocactus*. Previous attempts in this regard were hindered by the lack of resolution using traditional molecular markers (Santos, 2013). Target sequencing capture has increasingly and efficiently been used to elucidate phylogenetic relationships in plants, becoming a valuable tool for taxonomic revisions (e.g., Loiseau et al., 2019; Larridon et al., 2021; Nicol et al., 2024). The Cactaceae591 probe set associated with MSC has proven to be useful in resolving the phylogenetic relationships within *Discocactus*.

Despite the overall high resolution, an instance of phylogenetic incongruence was observed across concatenated and MSC methods, as commonly observed in genomic datasets (e.g., Wu et al., 2018; Cai et al., 2021; Romeiro-Brito et al., 2022; Telhe et al., 2025). These differences stem

from distinct assumptions inherent in each method. While concatenation assumes evolutionary congruence among genes, the coalescent method accounts for gene tree heterogeneity to infer the species tree (Maddison, 1997; Edwards et al., 2016). Given that Cactaceae is a recent and rapidly diverging family (Hernandez-Hernandez et al., 2014), with hybridization events commonly reported (Machado, 2008), and lineages with extensive incomplete lineage sorting leading to gene tree heterogeneity (Copetti et al., 2017), the use of the MSC method associated with genomic data enhances the likelihood of recovering a more accurate and robust phylogeny (Jiang et al., 2020).

The estimated mean ages for *Discocactus* suggest a Pliocene origin, with intense speciation occurring during the Plio-Pleistocene transition and early Pleistocene (Gelasian), particularly within the Savanna clade. These timeframes coincide with the final uplift of the Central Brazilian Plateau (Del'Arco & Bezerra, 1989) in the Cerrado domain, which promoted the formation of new landscape features, increasing soil heterogeneity, and the compartmentalization between plateaus and depressions (Colli, 2005). Additionally, during the Pleistocene climatic oscillations alternating between glaciation and interglacial cycles have occurred, as well as a reduction in atmospheric CO_2 (Tripati, Roberts & Eagle 2009; Guillermic et al., 2022). Topographic variation, soil heterogeneity, and climatic variability are among the most important factors generating habitat diversity and promoting species diversification in the Neotropical region (Rull & Carnaval, 2020 and references therein). Similar diversification timeframes have been estimated for several Cactaceae genera (e.g., *Cereus* Mill., Franco et al., 2017; *Eulychnia* Phil., Merklinger et al., 2021; *Mammillaria* Haw., Cornejo-Romero et al., 2014; *Pilosocereus* Byles & G.D.Rowley, Bonatelli et al., 2014, Romeiro-Brito et al., 2023a; *Melocactus* Link & Otto, Telhe et al., 2025). The age recovered in this study is in concordance with other Cactaceae studies, suggesting that both orogenic and Quaternary climatic changes, as well as the atmospheric CO_2 reductions, were important drivers for cactus diversification.

3.4.2 Phylogenetic relationships and taxonomic classification

The intricate phylogenetic relationships among *Discocactus* species have now been disentangled, revealing that this genus, predominantly found in the Savanna biome (Cerrado domain), includes earlier diverging species that are exclusively found in the SDTF biome, such as the Caatinga domain. Our MSC inference shows that *D. zehntneri* and *D. petr-halfari* form a clade sister to the rest of the genus. Notably, both species are restricted to the Caatinga domain (Taylor & Zappi, 2004), the largest SDTF of the Neotropical region (Pennington et al., 2000; Silva, 2017). Following this, the phylogeny reveals that *D. boomianus*, *D. buenekeri* W. R. Abraham, and *D.*

bahiensis, all restricted to the Caatinga domain (Flora e Funga do Brasil, 2025), form subsequent lineages. Sister to *D. bahiensis* is a clade comprising all the remaining *Discocactus* species, mostly endemic to the Savannas, highlighting a clear pattern of SDTF-Savanna distinction within the genus.

Discocactus zehntneri subsp. *boomianus* (Buining & Brederoo) N. P. Taylor & Zappi (*D. boomianus*), which occurs in sandstone or occasionally ironstone (*canga*), is consistently recovered as a distinct taxon from *D. zehntneri*, which occurs in gneiss outcrops and gravelly soils. This result indicates a need for a taxonomic rearrangement. Originally described by Buining & Brederoo in 1971 as *D. boomianus*, it was later synonymized as *D. zehntneri* subsp. *boomianus* by N. P. Taylor & Zappi (1991). Morphological differences, such as a cephalium with numerous thinner bristles, and thin spines arranged in a dense spination that usually obscures the stem, along with our phylogenetic findings, support the reinstatement of *D. boomianus* as an accepted species.

The close relationship between *D. zehntneri* and *D. petr-halfarii* is not unexpected, since *D. petr-halfarii* has been considered a subspecies of *D. zehntneri* (*D. zehntnerii* subsp. *petr-halfarii* (M. Zachar) M. R. Santos & M. C. Machado). Like other *Discocactus* species, *D. petr-halfarii* has an intricate taxonomic history. It was first described as *D. petr-halfarii* by Zachar in 2008, but in the same year, Braun & Esteves suggested it should be considered as a synonym of *D. bahiensis* (Braun & Esteves, 2008). Later, Santos et al., (2015) placed it as a subspecies of *D. zehntneri* based on molecular information previously generated by Santos (2013). More recently, Taylor et al., (2022) reinstated *D. petr-halfarii* as a species due to its morphological and ecological characteristics. Even though *D. petr-halfarii* indeed resembles *D. zehntneri* in its spination pattern and spiraled tubercles, it also shares similarities with *D. bahiensis*, both showing a thickened taproot and occurring in stony and loamy soils. Our phylogenetic inferences untangled the unclear relationship of *D. petr-halfarii*, recovering it as sister to *D. zehntneri* and distantly related to *D. bahiensis*.

The phylogenetic relationship of the geophytic *Discocactus buenekeri* was previously unclear due to its highly distinct morphology. While some authors suggested a close relationship with *D. zehntneri* and *D. boomianus* (Abraham, 1987), our analysis reveals *D. buenekeri* as a distinct and isolated lineage. This species is easily distinguished from other *Discocactus* species by its unique morphological characteristics, including stem, spine, and seed sizes, ribs dissolution, and the presence of scales on the entire flower receptacle (Abraham, 1987). Additionally, this species

shows a unique geophytic habit, being partially to completely covered by the sand in which it occurs, making it difficult to find in the dry season (see Figure 1.D).

The subsequent species recovered in the phylogeny is *D. bahiensis*, a species once considered a potential conspecific with *D. heptacanthus. Discocactus bahiensis* was considered conspecific due to intermediate characteristics observed in a population from the municipality of Paramirim, Bahia (Machado, 2005). Recently, the Paramirim population with intermediate characteristics was recognized as a population belonging to *D. bahiensis* (*D. bahiensis* subsp. *gracilis* P.J.Braun & Esteves; Taylor et al., 2022). Its occurrence near rivers in alluvial stony-loamy soils is an important ecological feature of *D. bahiensis* which is usually helpful in its identification. Our phylogenetic analysis recovered *D. bahiensis* as a distinct lineage, sister to the clade comprising most of the *Discocactus* species restricted to the Savanna biome.

Discocactus diersianus is closely allied to the clade comprising *D. hartmannii, D. ferricola, D. boliviensis, D. catingicola, and D. heptacanthus.* This species occurs in granite and quartz outcrops in Goiás and Tocantins. Although its phylogenetic position is well-supported, this species is morphologically variable, having at least three synonymized taxa (*sensu* Flora e Funga do Brasil, 2025). Among its synonymized forms, it is possible to find populations with spines within the cephalium instead of the common bristles, shorter flowers (e.g. *D. cephaliaciculosu*s Buining & Brederoo synonym; Buining & Brederoo, 1975), and smaller plants with fewer ribs (e.g. *D. lindaianus* Diers & Esteves synonym; Diers & Esteves, 1981). The *D. diersianus* specimen sampled in this study refers to the form described as *D. lindaianus*. Given the morphological variation among populations and the absence of a type locality specimen, further studies with a broader sampling of this species are needed to better understand species limits and its relationship within the genus.

The close relationship between *D. heptacanthus* and *D. catingicola* has been previously noted. *Discocactus catingicola* was initially treated as a subspecies of *D. heptacanthus* (*D. heptacanthus* subsp. *catingicola*) by Taylor & Zappi (1997) but later reinstated as an accepted species by Machado (2005). Although Machado (2005) recognized *D. catingicola* as an accepted species, he observed that some populations of *D. heptacanthus* from Central Brazil (Goiás state) were almost morphologically indistinguishable from *D. catingicola*. Besides the morphological similarities found in some populations, these species show a similar ecology. Both species are considered more generalists regarding their soil preferences, being found in gravel, sandy, rocky soils, and even in *canga* soils. *Discocactus heptacanthus* is one of the most variable species, with many variations described (Taylor & Zappi, 2004; Hunt et al., 2006). This morphological

variability led Braun & Esteves (1993) to suggest that *D. heptacanthus* could represent a species complex. Our data further supports *D. heptacanthus* as a monophyletic clade, including its variations.

In our phylogenetic inferences, the *D. hartmannii* specimen from Paraguay was resolved as a separate lineage from all other *D. hartmannii* specimens from Brazil. *Discocactus hartmannii* can be found in gravel, rocky, or sandy soils, while in Paraguay they are usually found in sandy soils (Santos, 2013). Although the recovery of the separated lineages is strongly supported, it should be interpreted with caution. Considering the morphological variation observed in *D. hartmannii* across its range (Hunt et al., 2006; Santos, 2013) and the limited sampling from Paraguay in our phylogeny, further studies with a broader sampling are needed to better understand species limits and status of *D. hartmannii*.

The close relationship between *D. boliviensis* and *D. ferricola*, resolved as sister lineages (Figure 2), is consistent with the findings of Santos (2013). Despite the general lack of phylogenetic resolution among *Discocactus* species, Santos (2013) also recovered *D. boliviensis* and *D. ferricola* as sister lineages with high support. Both species are found near Santa Cruz in Bolivia, but *D. ferricola* shows a wider distribution extending into Brazil (Corumbá - Mato Grosso do Sul). These sister species are easily distinguished by their morphological and ecological differences. Morphologically, they differ in their tubercules shape, in which *D. ferricola* shows a rhomboid form, while in *D. boliviensis* they are more pentagonal to rounded (Santos, 2013). Additionally. *D. ferricola* is found on ironstone outcrops (*canga*), while *D. boliviensis* is found on limestone outcrops (Buining & Bergink 1975; Hunt et al., 2006).

Discocactus fariae-peresii, described by Braun in 2016, had an unclear relationship within the genus. Morphologically, it closely resembles *D. placentiformis* but differs primarily in having a greater number of ribs and less sinuous rib structure. Geographically, the two species are also distinct. While *D. placentiformis* is found in south-eastern Brazil (Minas Gerais state), *D. fariae-peresii* occurs in the central-west region (Goiás state). Our study recovered *D. fariae-peresii* as a distinct lineage from *D. placentiformis*, supporting its recognition as an accepted species.

Discocactus placentiformis, a morphologically very variable and a more generalist species regarding its soil preferences (Taylor & Zappi, 2004; Santos, 2013), was recovered with high support as a sister to the clade clustering *D. pseudoinsignis* and *D. horstii*. Given the morphological

variation observed among *D. placentiformis* populations, further studies with a broader sampling are needed to better understand its morphological diversity and species limits.

The close relationship between *D. pseudoinsignis* and *D. horstii* has been previously suggested due to their morphological similarities (Taylor & Zappi, 2004). Both species share the presence of straight ribs, while the ribs in other species are tuberculated (Taylor & Zappi, 2004; Hunt et al., 2006). Additionally, *D. horstii* is considered a dwarf, neotenic ally of *D. pseudoinsignis* (Taylor & Zappi, 2004). Both species are considered rare with a limited distribution in south-eastern Brazil (Minas Gerais state). For example, *D. horstii* range covers around 100 km^2, known for a few locations (≅ three locations). Regarding their ecology, *D. horstii* is found in quartz gravel and sandy soils, while *D. pseudoinsignis* is found in pure quartz soils.

The recently described species, *D. piscibarbarus* was recovered as a distinct lineage in our phylogenetic inferences, supporting its recognition as an accepted species. This species occurs in the marginal distribution of *D. placentiformis* but is easily distinguished by its much smaller stature, stem color, round-edged ribs, delicate spination, and exclusive occurrence on ironstone outcrops (*canga*) (Taylor et al., 2023). These features are somewhat similar to those of *D. ferricola*, another species restricted to *canga* (Taylor et al., 2023). Given the geographical and phylogenetic distance between *D. pscibarbarus* and *D. ferricola*, their shared substrate occurrence may play a role in shaping their similar morphology.

The results of this study support the taxonomic circumscription of 16 accepted *Discocactus* species, which recognizes the morphological variation in certain taxa as variations, leading to the synonymization of many taxa previously described as species (Hunt et al., 2006; Koroktova et al., 2024; Flora e Funga do Brasil, 2025). Due to the small population sizes, the limited number of populations per species, and habitat specificity, more than 70% of *Discocactus* species are classified as Endangered or, to some degree, threatened (IUCN, 2024). Additionally, many species are restricted to rocky outcrops subject to intense mining due to mineral resources (e.g., iron and limestone), which makes their populations highly vulnerable. These findings highlight the need for the development of conservation strategies for this genus.

3.4.3 Key traits predicting the occurrence of Discocactus species in Neotropical SDTFs or Savannas biomes

The role of plant traits and their potential adaptive value has interested plant biologists and ecologists for decades (Caruso et al., 2020 and references therein). However, the uncertainty about which trait may influence plant adaptation limits us from testing hypotheses regarding traits and their evolutionary and ecological process and patterns (Anderegg, 2023). Here, we explored different categories of biotic data, from morphological to ecological characteristics, and identified which one potentially influences *Discocactus* species colonization and persistence among Neotropical open and arid to semi-arid regions, the SDTFs and Savanna biomes.

We found that *Discocactus* species primarily distributed in the SDTF biome have smaller flowers (small length size) compared to those in the Savanna biome. While differences in flower size are often correlated with pollinators, this is only one of the several factors influencing floral traits (Galen, 1999; Lambrecht & Dawson, 2007; Roddy et al., 2021). Environmental factors affecting resource availability are also recognized as key drivers of floral trait variation, given the inherent resource cost of flowers to the plant (Galen, 1999; Caruso et al., 2019; Roddy et al., 2021). Temperature and water availability, in particular, are known to influence flower development and size (Galen, 1999; Suni et al., 2023, Roddy et al., 2023). Variation in flower size seems to be linked to plant water balance (Lambrecht & Dawson, 2007; Lambrecht, 2013), as flowers lose a significant amount of water through transpiration (Galen, 1999; Teixido et al., 2019). Therefore, having smaller and fewer flowers may minimize this loss (Herrera, 2005; Lambrecht & Dawson, 2007). Indeed, several studies have identified this pattern, a decrease in flower size in hotter and drier conditions (e.g., Galen, 2000; Herrera, 2005; Lambrecht, 2013; Campbell & Powers, 2015; Lambrecht et al., 2017; Kuppler & Kotowska, 2021; Kuppler et al., 2021; Descamps et al., 2021). Our findings provide empirical evidence for the role of temperature and water availability in shaping floral traits in *Discocactus* species occupying different biomes. Differences in flower phenology may also be associated with pollinators, but this hypothesis was not explored in this study as the current knowledge about the pollinators of the genus is limited to unspecified moths, without detailed data regarding the pollination of different *Discocactus* species.

Discocactus species primarily found in the SDTF biome have smaller stem diameters compared to those in the Savanna biome. Cactus stems function as photosynthetic organs, and their surface area is somewhat analogous to the leaf area in other plants (Williams et al., 2014). Like leaves, the evolutionary diversification of stem dimensions is likely constrained by ecological and biophysical trade-offs (Williams et al., 2014). Cacti utilize the CAM photosynthetic pathway, which reduces water loss by keeping stomata closed during the day and open only at night (Nobel,

2003; Lüttge, 2004). While advantageous in dry habitats, this photosynthetic pathway limits the plant's ability to dissipate heat during the day, when thermal stress is higher (Hultine et al., 2023). Consequently, stem temperature can rise significantly, sometimes by 15-20°C above the ambient air temperature (Gibbs & Patten, 1970; Smith, 1978). This issue is particularly severe in globose and dwarf cacti, as their photosynthetic surfaces are closer to the ground, where temperatures are higher (Smith et al.,1984; Nobel et al., 1986; Drezner, 2011).

Elevated stem temperatures can impair photosynthetic metabolism, reducing nocturnal acid accumulation and causing chloroplast membrane degradation (Smith et al.,1984; Hultine et al., 2023). Extreme temperatures, averaging 57°C, can cause up to a 50% reduction in photosynthetic activity (Nobel, 2003 and references therein). One strategy to mitigate stem overheating is the reduction of stem diameter, as smaller diameters enhance convective heat dissipation due to the correlation between the effectiveness of heat dissipation and the square diameter of a cacti stem (Smith et al., 1984; Nobel, 1978). Studies within cacti have shown that stem diameters decrease under hotter, and more arid conditions when investigating temperature and latitudinal gradients (Nobel, 1978; 1980; Smith et al., 1984; Delgado-Fernández, 2017). Among *Discocactus* species, variations in stem diameter across different habitats likely reflect this trade-off between stem diameter and environmental conditions.

While temperature and precipitation are key factors influencing plant reproduction and growth (Moles et al., 2014; Hatfield & Prueger, 2015), a more complex scenario involving interactions among climate, soil conditions, and community interactions likely offers a better explanation for cactus evolution (Godinez-Alvarez et al., 2003; Guerrero et al., 2012; Fonseca-García et al., 2016; Guerrero et al., 2019 and references therein). However, this complexity does not preclude the possibility of isolating specific interactions to enhance our understanding of the evolutionary processes in plant species across different habitats.

3.4.3 Taxonomic Adjustments

Discocactus is a genus of Cactaceae with one of the most controversial taxonomic histories, in which the phylogenomic inferences revealed a need for taxon rearrangement. One species of *Discocactus* was reinstated as follows:

Discocactus boomianus Buining & Bredero in Succulenta (Netherlands) 50: 26 (1971). Holotype: Brazil, Bahia, Município Morro do Chapéu, *Horst* 222 (U). = *Discocactus zehntneri* subsp. *boomianus* (Buining & Bredero) Taylor & Zappi in Bradleya 9: 86 (1991).

Diagnosis/Notes: The reinstated species, *D. boomianus*, stands out as an independent evolutionary lineage in our phylogenetic inferences. It is morphologically distinct from *D. zehntneri*, showing a cephalium with numerous thinner bristles, and thin spines arranged in a dense spination, usually obscuring the stem. Ecologically, *D. boomianus* is generally found in arenitic rocks, often with an accumulation of gravel or in pure quartz sand, with a few populations also occurring in ironstone (*canga*) in north Bahia (Chapada Diamantina). In contrast, *D. zehntneri* is usually found in gneiss rocks and gravelly soil in north Bahia (Chapada Diamantina), Sergipe, and in sandstone in the northwest of Ceará.

3.5 Acknowledgments

We express our sincere gratitude to Dr. Marlon Machado, Dr. Diego Gonzaga, and the *Coleção de Cactos e Suculentas* at the Instituto de Pesquisas Jardim Botânico do Rio de Janeiro for generously providing cactus samples used in this study. This study was funded by FAPESP, which provided grants supporting doctoral students and supervisors (MCT, 2019/11233-2; MRB, 2018/06937-8; EMM, 2018/03428-5 and 2019/03211-9; FFF, 2020/15161-3). DCZ and EMM hold productivity grants from CNPq (304178/2021-7 and 310963/2022-6, respectively). The authors declare that the publication of this article does not involve any conflicts of interest.

3.6 Funding

This study was financially supported by FAPESP, which provided fellowships for doctoral students (Processes 2019/11233-2 to MCT and 2018/06937-8 to MRB) and research grants (2019/03211-9 to EMM, 2020/15161-3 to FFF and 2018/03428-5 to EMM and FFF). DCZ and EMM hold productivity grants from CNPq (304178/2021-7 and 310963/2022-6, respectively).

3.7 References

Abraham, W. R. .1987. Ein außergewöhnlicher neuer *Discocactus*: *Discocactus buenekeri* Abraham. *Kakteen und andere Sukkulenten*, 38 (11), p. 282-285.

Anderegg, L. D. (2023). Why can't we predict traits from the environment?. *New Phytologist*, *237*(6), 1998-2004.

Barthlott, W., Burstedde, K., Geffert, J. L., Ibisch, P. L., Korotkova, N., Miebach, A., Rafiqpoor, M. D., Stein, A., & Mutke, J. (2015). *Biogeography and Biodiversity of Cacti* (Vol. 7). Deutschen KakteenGesellschaft e.V.

Bezerra-Silva, A., Albuquerque-Lima, S., Gomes, V. G. N., Fagundes, A. C. D. A., Gomes, M. T. D., Silva, M. T. D., ... & Funch, L. S. (2024). When are cacti found with flowers and fruits? Estimation of the reproductive phenology of the genus *Xiquexique* based on herbarium data. *Diversity*, *16*(2), 79.

Bombonato, J. R., do Amaral, D. T., Silva, G. A. R., Khan, G., Moraes, E. M., da Silva Andrade, S. C., ... & Franco, F. F. (2020). The potential of genome-wide RAD sequences for resolving rapid radiations: a case study in Cactaceae. *Molecular phylogenetics and evolution*, 151, 106896.

Bonatelli, I. A., Perez, M. F., Peterson, A. T., Taylor, N. P., Zappi, D. C., Machado, M. C., ... & Moraes, E. M. (2014). Interglacial microrefugia and diversification of a cactus species complex: phylogeography and palaeodistributional reconstructions for *Pilosocereus aurisetus* and allies. *Molecular Ecology*, 23(12), 3044-3063.

Braun, P. J. & Esteves P. E. 2008. *Discocactus bahiensis* Britton & Rose subsp. *petr-halfari* (Zachar) P.J.Braun & Esteves. *Kakteen und Andere Sukkulenten*, vol. 57, n. 9.

Braun, P. J. & Esteves, P. E. (1993). *Discocactus heptacanthus* var. *semicampaniflorus* and *Discocactus heptacanthus* subsp. *melanochlorus*. Kakteen und Andere Sukkulenten 44(5): 103.

Braun, P. J. (2016). *Discocactus fariae-peresii* P.J. Braun. Kakteen und Andere Sukkulenten 67(1): 33, p. 1–6.

Breiman, L. (2001). Random forests. *Machine learning*, *45*, 5-32.

Buining, A. F. H & Brederoo, A. J. 1975. *Discocactus cephaliaciculosus* Buining & Brederoo. *Kakteen und andere Sukkulenten*, 26 (5), p. 97-100.

Buining, A. F. H. (1980). *The genus Discocactus Pfeiffer: a revision of known, and description of new species*. The Buining-fonds, a foundation of Succulenta, Venlo, Netherlands.

Bustamante, E., & Búrquez, A. (2008). Effects of plant size and weather on the flowering phenology of the organ pipe cactus (*Stenocereus thurberi*). *Annals of Botany*, *102*(6), 1019-1030.

Cai, L., Xi, Z., Lemmon, E. M., Lemmon, A. R., Mast, A., Buddenhagen, C. E., ... & Davis, C. C. (2021). The perfect storm: gene tree estimation error, incomplete lineage sorting, and ancient gene flow explain the most recalcitrant ancient angiosperm clade, Malpighiales. *Systematic Biology*, 70(3), 491-507.

Campbell, D. R., & Powers, J. M. (2015). Natural selection on floral morphology can be influenced by climate. *Proceedings of the Royal Society B: Biological Sciences*, *282*(1808), 20150178.

Caruso, C. M., Eisen, K. E., Martin, R. A., & Sletvold, N. (2019). A meta-analysis of the agents of selection on floral traits. *Evolution*, *73*(1), 4-14.

Caruso, C. M., Mason, C. M., & Medeiros, J. S. (2020). The evolution of functional traits in plants: is the giant still sleeping?. *International Journal of Plant Sciences*, *181*(1), 1-8.

Colli, G. R. (2005). As origens ea diversificação da herpetofauna do Cerrado. *Cerrado: ecologia, biodiversidade e conservação*, 247-264.

Copetti, D., Búrquez, A., Bustamante, E., Charboneau, J. L., Childs, K. L., Eguiarte, L. E., ... & Sanderson, M. J. (2017). Extensive gene tree discordance and hemiplasy shaped the genomes of North American columnar cacti. *Proceedings of the National Academy of Sciences*, 114(45), 12003-12008.

Cornejo-Romero, A., Medina-Sánchez, J., Hernández-Hernández, T., Rendón-Aguilar, B., Valverde, P. L., Zavala-Hurtado, A., ... & Vargas-Mendoza, C. F. (2014). Quaternary origin and genetic divergence of the endemic cactus *Mammillaria pectinifera* in a changing landscape in the Tehuacán Valley, Mexico. *Genetics and Molecular Research*, 13(1), 73-88.

Del'Arco, J.F. and Bezerra, P.E.L. (1989) Geologia. In *Geografia do Brasil - Região Centro-Oeste.* (A.C. Duarte, ed.) pp. 35±51. Rio de Janeiro: FIBGE-Directoria de Geociências.

Delgado-Fernández, M., Garcillán, P. P., & Ezcurra, E. (2017). The giant columnar cactus *Pachycereus pringlei* adaptively modifies its stem shape from the dry tropics into the arid mid-latitude deserts. *Journal of Arid Environments*, *146*, 10-17.

Descamps, C., Quinet, M., & Jacquemart, A. L. (2021). The effects of drought on plant–pollinator interactions: What to expect?. *Environmental and Experimental Botany*, *182*, 104297.

Diers, L. & Esteves, E. P. 1981. *Discocactus lindaianus* Diers & Esteves. *Cactus & Succulent Journal*, vol. 53, p. 56-60.

Dinerstein, E., Olson, D., Joshi, A., Vynne, C., Burgess, N. D., Wikramanayake, E., ... & Saleem, M. (2017). An ecoregion-based approach to protecting half the terrestrial realm. *BioScience*, *67*(6), 534-545.

Drezner, T. D. (2011). Cactus surface temperatures are impacted by seasonality, spines and height on plant. *Environmental and experimental botany*, *74*, 17-21.

Edwards S.V., Xi Z., Janke A., Faircloth B.C., McCormack J.E., Glenn T.C., Zhong B., Wu S., Lemmon E.M., Lemmon A.R., Leaché A.D., Liu L., Davis C.C. 2016. Implementing and testing the multispecies coalescent model: a valuable paradigm for phylogenomics. *Molecular Phylogenetics and Evolution*, 94:447–462

Flora e Funga do Brasil. (2025). *Flora e Funga do Brasil*. Jardim Botânico do Rio de Janeiro. Available online: https://floradobrasil.jbrj.gov.br. (Accessed, February 7th, 2025).

Fonseca-García, C., Coleman-Derr, D., Garrido, E., Visel, A., Tringe, S. G., & Partida-Martínez, L. P. (2016). The cacti microbiome: interplay between habitat-filtering and host-specificity. *Frontiers in Microbiology*, *7*, 150.

Franco, F. F., Amaral, D. T., Bonatelli, I. A., Romeiro-Brito, M., Telhe, M. C., & Moraes, E. M. (2022). Evolutionary genetics of cacti: research biases, advances and prospects. *Genes*, 13(3), 452.

Franco, F. F., Silva, G. A. R., Moraes, E. M., Taylor, N., Zappi, D. C., Jojima, C. L., & Machado, M. C. (2017). Plio-Pleistocene diversification of *Cereus* (Cactaceae, Cereeae) and closely allied genera. *Botanical Journal of the Linnean Society*, 183(2), 199-210.

Galen, C. (1999). Why do flowers vary? The functional ecology of variation in flower size and form within natural plant populations. *Bioscience*, *49*(8), 631-640.

Galen, C. (2000). High and dry: drought stress, sex-allocation trade-offs, and selection on flower size in the alpine wildflower Polemonium viscosum (Polemoniaceae). *The American Naturalist*, *156*(1), 72-83.

Galen, C. (2005). It never rains but then it pours: the diverse effects of water on flower integrity and function. In *Reproductive allocation in plants* (pp. 77-95). Academic Press.

Gibbs, J. G., & Patten, D. T. (1970). Plant temperatures and heat flux in a Sonoran Desert ecosystem. *Oecologia*, *5*, 165-184.

Godinez-Alvarez, H., Valverde, T., & Ortega-Baes, P. (2003). Demographic trends in the Cactaceae. *The Botanical Review*, *69*(2), 173-201.

Guerrero, P. C., Carvallo, G. O., Nassar, J. M., Rojas-Sandoval, J., Sanz, V., & Medel, R. (2012). Ecology and evolution of negative and positive interactions in Cactaceae: lessons and pending tasks. *Plant Ecology & Diversity*, *5*(2), 205-215.

Guerrero, P. C., Majure, L. C., Cornejo-Romero, A., & Hernández-Hernández, T. (2019). Phylogenetic relationships and evolutionary trends in the cactus family. *Journal of Heredity*, *110*(1), 4-21.

Guerrero, P. C., Majure, L. C., Cornejo-Romero, A., & Hernández-Hernández, T. (2019). Phylogenetic relationships and evolutionary trends in the cactus family. *Journal of Heredity*, *110*(1), 4-21.

Guillermic, M., Misra, S., Eagle, R., & Tripati, A. (2021). Atmospheric CO 2 estimates for the Miocene to Pleistocene based on foraminiferal δ 11 B at Ocean Drilling Program Sites 806 and 807 in the Western Equatorial Pacific. *Climate of the Past Discussions*, 2021, 1-40.

Hatfield, J. L., & Prueger, J. H. (2015). Temperature extremes: Effect on plant growth and development. *Weather and climate Extremes*, *10*, 4-10.

Hernández-Hernández, T., Brown, J. W., Schlumpberger, B. O., Eguiarte, L. E., & Magallón, S. (2014). Beyond aridification: multiple explanations for the elevated diversification of cacti in the New World Succulent Biome. *New Phytologist*, *202*(4), 1382-1397.

Herrera, J. (2005). Flower size variation in *Rosmarinus officinalis*: individuals, populations and habitats. *Annals of Botany*, *95*(3), 431-437.

Ho, L. S. T., Ane, C., Lachlan, R., Tarpinian, K., Feldman, R., Yu, Q., ... & Ho, M. L. S. T. (2016). Package 'phylolm'. *See http://cran. r-project. org/web/packages/phylolm/index. html (accessed February 2018)*.

Hultine, K. R., Hernández-Hernández, T., Williams, D. G., Albeke, S. E., Tran, N., Puente, R., & Larios, E. (2023). Global change impacts on cacti (Cactaceae): current threats, challenges and conservation solutions. *Annals of botany*, *132*(4), 671-683.

Hunt, D. R., Taylor, N. P., & Charles, G. (2006). *The new cactus lexicon*. DH Books.

Hunt, D., & Taylor, N. (Eds.). (1991). Notes on miscellaneous genera of Cactaceae. *Bradleya*, *9*, 81-92.

Ichimura, K., & Suto, K. (1998). Environmental Factors Controlling Flower Opening and Closing in a Portulaca Hybrid. *Annals of Botany*, *82*(1), 67-70.

Jiang, X., Edwards, S. V., & Liu, L. (2020). The multispecies coalescent model outperforms concatenation across diverse phylogenomic data sets. *Systematic Biology*, *69*(4), 795-812.

Khan, S., Jaral, S., Kumari, P., & Verma, S. (2024). Effect of elevated temperature on the rate of flowering anthesis and seed set in *Olea ferruginea*. *Ecological Frontiers*, *44*(2), 282-288.

Kuhn, Max. "Caret package." *Journal of statistical software* 28.5 (2008): 1-26.

Kuppler, J., & Kotowska, M. M. (2021). A meta-analysis of responses in floral traits and flower–visitor interactions to water deficit. *Global Change Biology*, *27*(13), 3095-3108.

Kuppler, J., & Kotowska, M. M. (2021). A meta-analysis of responses in floral traits and flower–visitor interactions to water deficit. *Global Change Biology*, *27*(13), 3095-3108.

Kuppler, J., Wieland, J., Junker, R. R., & Ayasse, M. (2021). Drought-induced reduction in flower size and abundance correlates with reduced flower visits by bumble bees. *AoB Plants*, *13*(1), plab001.

Kursa, M. B., & Rudnicki, W. R. (2010). Feature selection with the Boruta package. *Journal of Statistical Software*, *36*, 1-13.

Lambrecht, S. C. (2013). Floral water costs and size variation in the highly selfing Leptosiphon bicolor (Polemoniaceae). *International Journal of Plant Sciences*, *174*(1), 74-84.

Lambrecht, S. C., & Dawson, T. E. (2007). Correlated variation of floral and leaf traits along a moisture availability gradient. *Oecologia*, *151*, 574-583.

Lambrecht, S. C., Morrow, A., & Hussey, R. (2017). Variation in and adaptive plasticity of flower size and drought-coping traits. *Plant Ecology*, *218*, 647-660.

Lanna, F. M., Colli, G. R., Burbrink, F. T., & Carstens, B. C. (2022). Identifying traits that enable lizard adaptation to different habitats. *Journal of Biogeography*, *49*(1), 104-116.

Luebert, F. (2021). The two South American dry diagonals. *Frontiers of Biogeography*, *13*(4).

Lüttge, U. (2004). Ecophysiology of crassulacean acid metabolism (CAM). *Annals of botany*, *93*(6), 629-652

Machado, M. C. (2008). What is the role of hybridization in the evolution of the Cactaceae?. *Bradleya*, *2008*(26), 1-18.

Machado, M. C., Zappi, D. C., Taylor, N. P., & Borba, E. L. (2005). Taxonomy and conservation of the *Discocactus* Pfeiff. (Cactaceae) species occurring in the state of Bahia, Brazil. *Bradleya*, *2005*(23), 41-56.

Maddison, W. P. (1997). Gene trees in species trees. *Systematic Biology*, *46*(3), 523-536.

McKain, M. R., M. G. Johnson, S. Uribe-Convers, D. Eaton, and Y. Yang. (2018). Practical considerations for plant phylogenomics. *Applications in Plant Sciences* 6(3): e1038

Mendonça Filho, C. V., Souza, J. P., Lopes, L. L., & Antonini, Y. (2024). Long-term monitoring of the columnar cactus *Cipocereus minensis* reveals unforeseeable reproductive phenology. *Journal of Arid Environments*, *224*, 105202.

Merklinger, F. F., Böhnert, T., Arakaki, M., Weigend, M., Quandt, D., & Luebert, F. (2021). Quaternary diversification of a columnar cactus in the driest place on earth. *American Journal of Botany*, 108(2), 184-199.

Miranda-Jácome, A., & Sosa, V. J. (2023). Inter-and intra-annual variation in the pulsed flowering phenology of the columnar cactus *Pilosocereus leucocephalus* and its relation to temperature, rainfall and plant size. *Flora*, *305*, 152339.

Moles, A. T., Perkins, S. E., Laffan, S. W., Flores-Moreno, H., Awasthy, M., Tindall, M. L., ... & Bonser, S. P. (2014). Which is a better predictor of plant traits: temperature or precipitation?. *Journal of Vegetation Science*, *25*(5), 1167-1180

Neves, D. M., Dexter, K. G., Pennington, R. T., Bueno, M. L., & Oliveira Filho, A. T. (2015). Environmental and historical controls of floristic composition across the South American Dry Diagonal. *Journal of Biogeography*, *42*(8), 1566-1576.

Nobel, P. S. (1978). Surface Temperatures of Cacti-Influences of Environmental and Morphological Factors. *Ecology*, *59*(5), 986-995.

Nobel, P. S. (1980). Morphology, surface temperatures, and northern limits of columnar cacti in the Sonoran Desert. *Ecology*, *61*(1), 1-7.

Nobel, P. S. (2003). *Environmental biology of agaves and cacti*. Cambridge University Press

Nobel, P. S., Geller, G. N., Kee, S. C., & Zimmerman, A. D. (1986). Temperatures and thermal tolerances for cacti exposed to high temperatures near the soil surface. *Plant, Cell & Environment*, *9*(4), 279-287.

Pennington, R. T., Prado, D. E., & Pendry, C. A. (2000). Neotropical seasonally dry forests and Quaternary vegetation changes. *Journal of Biogeography*, *27*(2), 261-273.

Pennington, R. T., & Ratter, J. A. (Eds.). (2006). *Neotropical savannas and seasonally dry forests: plant diversity, biogeography, and conservation*. CRC press.

Pfeiffer, L. (1837) *Discocactus insigni*s Pfeiffer. Allgemeine Gartenzeitung 5: 241–242.

Pigot, A. L., Sheard, C., Miller, E. T., Bregman, T. P., Freeman, B. G., Roll, U., ... & Tobias, J. A. (2020). Macroevolutionary convergence connects morphological form to ecological function in birds. *Nature Ecology & Evolution*, *4*(2), 230-239.

Ringelberg et al. (2023). Precipitation is the main axis of tropical plant phylogenetic turnover across space and time. *Science Advances*, *9*(7), eade4954.

Roddy A.B., Guilliams C.M., Fine P.V.A., Mambelli S., Dawson T.E., Simonin K.A. (2023) Flowers are leakier than leaves but cheaper to build. *New Phytologist*, 239, 2076–2082.

Roddy, A. B., Martínez-Perez, C., Teixido, A. L., Cornelissen, T. G., Olson, M. E., Oliveira, R. S., & Silveira, F. A. (2021). Towards the flower economics spectrum. *New Phytologist*, *229*(2), 665-672.

Romeiro-Brito, M., Khan, G., Perez, M. F., Zappi, D. C., Taylor, N. P., Olsthoorn, G., ... & Moraes, E. M. (2023)a. Revisiting phylogeny, systematics, and biogeography of a Pleistocene radiation. *American Journal of Botany*, 110(3), e16134.

Romeiro-Brito, M., Taylor, N. P., Zappi, D. C., Telhe, M. C., Franco, F. F., & Moraes, E. M. (2023). Unravelling phylogenetic relationships of the tribe Cereeae using target enrichment sequencing. *Annals of Botany*, 132(5), 989-1006.

Romeiro-Brito, M., Telhe, M. C., Amaral, D. T., Franco, F. F., & Moraes, E. M. (2022). A target capture probe set useful for deep-and shallow-level phylogenetic studies in Cactaceae. *Genes*, *13*(4), 707.

Rull, V. & Carnaval, A. C. (2020). *Neotropical diversification: patterns and processes*. Springer International Publishing.

Rull, V. (2011). Neotropical biodiversity: timing and potential drivers. *Trends in ecology & evolution*, *26*(10), 508-513.

Santos, M. R. *Filogenia molecular, taxonomia, biogeografia e conservação de Discocactus Pfeiff. (Cactaceae)*. 2013. 120 f. [Doctoral thesis - Universidade Federal de Viçosa]. LOCUS UFV Repository http://www.locus.ufv.br/handle/123456789/7588.

Santos, M. R., Machado, M. C., Garcia, F. C. P., Taylor, N. P. (2015). Taxonomic adjustments in *Discocactus* (Cactaceae). *Phytotaxa*, *207*(2), 209-212.

Santos, R. M., Oliveira-Filho, A. T., Eisenlohr, P. V., Queiroz, L. P., Cardoso, D. B., & Rodal, M. J. (2012). Identity and relationships of the Arboreal Caatinga among other floristic units of seasonally dry tropical forests (SDTFs) of north-eastern and Central Brazil. *Ecology and Evolution*, *2*(2), 409-428.

Silva, G. A. R., Antonelli, A., Lendel, A., Moraes, E. D. M., & Manfrin, M. H. (2018). The impact of early Quaternary climate change on the diversification and population dynamics of a South American cactus species. *Journal of Biogeography*, *45*(1), 76-88

Silva, I. A., & Batalha, M. A. (2010). Phylogenetic structure of Brazilian savannas under different fire regimes. *Journal of Vegetation Science*, *21*(6), 1003-1013.

Silva, J. M. C. D., Leal, I. R., & Tabarelli, M. (Eds.). (2017). Caatinga—The largest tropical dry forest region in South America. Springer.

Simon, M. F., & Pennington, T. (2012). Evidence for adaptation to fire regimes in the tropical savannas of the Brazilian Cerrado. *International Journal of Plant Sciences*, *173*(6), 711-723.

Simon, M. F., Grether, R., de Queiroz, L. P., Skema, C., Pennington, R. T., & Hughes, C. E. (2009). Recent assembly of the Cerrado, a neotropical plant diversity hotspot, by in situ evolution of adaptations to fire. *Proceedings of the National Academy of Sciences*, *106*(48), 20359-20364.

Simon, S. M., Glaum, P., & Valdovinos, F. S. (2023). Interpreting random forest analysis of ecological models to move from prediction to explanation. *Scientific Reports*, 13(1), 3881.

Smith SD, Pennell MW, Dunn CW, Edwards SV. (2020). Phylogenetics is the new genetics (for most of biodiversity). *Trends in Ecology and Evolution*, 35:(5):415–25

Smith, S. D., Didden-Zopfy, B., & Nobel, P. S. (1984). High-temperature responses of North American cacti. *Ecology*, *65*(2), 643-651.

Smith, W. K. (1978). Temperatures of desert plants: another perspective on the adaptability of leaf size. *Science*, *201*(4356), 614-616.

Stekhoven, D. J. (2011). Using the missForest package. *R package*, 1-11.

Suni, S., Kuwana, A., & Ibañez, N. (2023). Responses of floral traits to water limitation across soil types and along climatic gradients. *Flora*, *304*, 152295.

Taylor, N. P., & Zappi, D. C. (1997). Nomenclatural adjustments and novelties in Brazilian Cactaceae. *Cactaceae Consensus Initiatives*, *3*, 7-8.

Taylor, N. P., Albuquerque-Lima, S., Telhe, M., & Zappi, D. C. (2022). Further additions and corrections to Cacti of Eastern Brazil. *Bradleya*, *2022*(40), 61-92.

Taylor, N.P & Zappi, D.C. (1991). *Discocactus zehntneri* subsp. *boomianus*. *Bradleya* 9: 86.

Taylor, N.P. & Zappi, D.C. (2004). *Cacti of Eastern Brazil*. Richmond: Royal Botanic Gardens, Kew.

Teixido, A. L., Leite-Santos, V. B., Paiva, É. A., & Silveira, F. A. (2019). Water-use strategies in flowers from a neotropical savanna under contrasting environmental conditions during flowering. *Plant Physiology and Biochemistry*, *144*, 283-291.

Telhe, M.C., Taylor, N.P., Romeiro-Brito, M., Zappi, D.C., Olsthoorn, G., Franco, F.F., Moraes, E. M. (2025). Phylogenomic and ecological systematics of *Melocactus* (Cactaceae). *Manuscript accepted for publication in Systematic Botany*.

Tripati, A. K., Roberts, C. D., & Eagle, R. A. (2009). Coupling of CO2 and ice sheet stability over major climate transitions of the last 20 million years. *Science*, 326(5958), 1394-1397.

Whelan, R. J. (1995). *The ecology of fire*. Cambridge university press.

Williams, D. G., Hultine, K. R., & Dettman, D. L. (2014). Functional trade-offs in succulent stems predict responses to climate change in columnar cacti. *Journal of Experimental Botany*, *65*(13), 3405-3413.

Wu M., Kostyun J.L., Hahn M.W., Moyle L.C. 2018. Dissecting the basis of novel trait evolution in a radiation with widespread phylogenetic discordance. *Molecular Ecology*, 27:3301–3316.

Zachar, M. (2008). Nowość z Brazylii-*Discocactus petr-halfari*. *Kaktusy i Inne*, *2*, 43-47.

3.8 Supplementary Material

Table S1 Sampling information of *Melocactus* and outgroup species used for phylogenetic analysis. Most species included in this study are considered vulnerable or endangered, for this reason, no coordinates are included.

scientificName	scientificNameAuthorship	recordNumber	country	county	locality	identifiedBy
Discocactus Pfeiff. (ingroup)						
Discocactus bahiensis	Britton & Rose	S165A15	Brazil	Bahia	Povoado Carnaíba do Sertão	Taylor, N.
Discocactus bahiensis	Britton & Rose	L25A9	Brazil	Ceará	Lavras de Mangabeira	Olsthoorn, G
Discocactus boliviensis	Backeb	L26A2	Bolívia	Santa Cruz	Porto Suarez	Olsthoorn, G
Discocactus catingicola	Buining & Brederoo	L26A3	Brazil	Goiás	Itumbiara	Olsthoorn, G
Discocactus catingicola	Buining & Brederoo	L25A6	Brazil	Piauí	Bom Jesus	Olsthoorn, G
Discocactus catingicola	Buining & Brederoo	L26A48	Brazil	Mato Grosso do Sul	Porto Murtinho	Olsthoorn, G
Discocactus diersianus	Esteves	L25A1	Brazil	Goiás	Alto Paraíso de Goiás	Olsthoorn, G
Discocactus fariae-peresii	P.J.Braun	L25A2	Brazil	Goiás	Planaltina	Olsthoorn, G
Discocactus ferricola	Buining & Brederoo	L26A16	Brazil	Mato Grosso do Sul	Corumba	Olsthoorn, G
Discocactus ferricola	Buining & Brederoo	S162A6	Brazil	Mato Grosso do Sul	Corumba	Olsthoorn, G
Discocactus hartmannii	(K.Schum.) Britton & Rose	L26A5	Paraguay	San Pedro	San Estariola	Olsthoorn, G
Discocactus hartmannii	(K.Schum.) Britton & Rose	L26A47	Brazil	Mato Grosso do Sul	Caracol	Olsthoorn, G
Discocactus hartmannii	(K.Schum.) Britton & Rose	L26A45	Brazil	Mato Grosso do Sul	Terenos	Taylor, N.
Discocactus heptacanthus	(Rodrigues) Britton & Rose	L26A6	Brazil	Mato Grosso	Cuiabá	Olsthoorn, G

Discocactus heptacanthus var. semicampaniflorus	*(Buining & Brederoo) P.J.Braun & Esteves*	L26A46	Brazil	Mato Grosso do Sul	Coxim	Olsthoorn, G
Discocactus heptacanthus var. squamibaccatus	*(Buining & Brederoo) P.J.Braun & Esteves*	S162A32	Brazil	Tocantins	Jaú do Tocantins	Olsthoorn, G
Discocactus heptacanthus var. prominentigibbus	*(Buining & Brederoo) P.J.Braun & Esteves*	S162A31	Brazil	Tocantins	Formosa do Araguaia	Olsthoorn, G
Discocactus heptacanthus var. cangaensis	*(Buining & Brederoo) P.J.Braun & Esteves*	S162A34	Brazil	Goiás	Diorama	Olsthoorn, G
Discocactus horstii	*Buining & Brederoo*	L44A1				Gonzaga, D.
Discocactus placentiformis	(Lehm.) K.Schum.	L26A8	Brazil	Minas Gerais	Porteirinha	Olsthoorn, G
Discocactus psibarbarus	N.P.Taylor & Olsthoorn	L86B2	Brazil	Minas Gerais	Rio Pardo de Minas	Taylor, N.
Discocactus psibarbarus	N.P.Taylor & Olsthoorn	L26A10	Brazil	Minas Gerais	Salinas	Olsthoorn, G
Discocactus pseudoinsignis	N.P.Taylor & Zappi	L26A11	Brazil	Minas Gerais	Cristália	Olsthoorn, G
Discocactus zehntneri	Britton & Rose	L25A8	Brazil	Ceará	Guaraciaba do Norte	Olsthoorn, G
Discocactus zehntneri	Britton & Rose	L26A12	Brazil	Bahia	Sento Sé	Olsthoorn, G
Discocactus zehntneri ssp. boomianus	(Buining & Brederoo) N.P.Taylor & Zappi	L26A13	Brazil	Bahia	Sento Sé	Olsthoorn, G
Discocactus buenekeri	*W.R.Abraham*	*L26A17*	Brazil	Bahia	Povoado Alegre	Olsthoorn, G
Discocactus zehntneri ssp. petr-halfari	(Zachar)M.R.Santos & M.C.Machado	S165A19	Brazil	Bahia	Juazeiro	Taylor, N.
Outgroup						
Arrojadoa rhodantha	Britton & Rose	142A6	Brazil	Minas Gerais	Matias Cardoso	Olsthoorn, G
Arthrocereus glaziovii	(K.Schum.) N.P.Taylor & Zappi	S162A2	Brazil	Bahia	Caeté	Zappi, D.
Astrophytum myriostigma	Lem.	S170A23	Mexico	-	-	Zappi, D.

Browningia hertlingiana	(Backeb.) Buxb.	S152A7	Peru	-	-	KEW
Coleocephalocereus aureus	F.Ritter	S183A2	Brazil	Minas Gerais	Jacinto	Olsthoorn, G
Coleocephalocereus goebelianus	(Vaupel) Buining	S162A4	Brazil	Minas Gerais	Mato Verde	Mendes, D. B.; Machado, M.
Coleocephalocereus purpureus	(Buining & Brederoo) F.Ritter	S174A10	Brazil	Minas Gerais	Itinga	Olsthoorn, G
Melocactus intortus	(Mill.) Urb.	L42A9	Puerto Rico	-	-	Taylor, N.
Melocactus deinacanthus	Buining & Brederoo	S162A14	Brazil	Bahia	Bom Jesus da Lapa	Olsthoorn, G
Melocactus deinacanthus	Buining & Brederoo	S174A19	Brazil	Bahia	Juá	Olsthoorn, G
Melocactus harlowii	Britton & Rose) Vaupel	L42A10	Cuba	Santiago	Santa Maria de Loreto	Taylor, N.
Melocactus macracanthos	(Salm-Dyck) Link & Otto	S162A15	Curaçao	-	Shete Boka National Park	Olsthoorn, G
Melocactus violaceus	Pfeiff.	L26A55	Brazil	Rio de Janeiro	Saquarema	Olsthoorn, G
Melocactus oreas	Miq.	L26A53	Brazil	Bahia	Milagres	Olsthoorn, G
Melocactus violaceus ssp margaritaceus	N. P. Taylor	S180A12	Brazil	Pernambuco	Catimbau	Taylor, N.
Echinocereus pentalophus	Lem.	S170A29	Mexico	-	-	Zappi, D.
Ferocactus herrerae	J.G.Ortega	S170A34	Mexico	-	-	Zappi, D.
Gymnocalycium horstii	Buining	S170A35	Brazil	-	-	Zappi, D.
Harrisia adscendens	(Gürke) Britton & Rose	S142A7	Brazil	Bahia	Malhada	
Opuntia monacantha	(Willd.) Haw.	S170A46	Brazil	Rio de Janeiro	Itaipu	Zappi, D.
Parodia magnifica	(F.Ritter) F.H.Brandt	S170A49	Brazil	-	-	Zappi, D.
Pilosocereus arrabidae	(Lem.) Byles & G.D.Rowley	S79H2	Brazil	Espirito Santo	São Mateus	
Selenicereus chrysocardium	(Alexander) Kimnach	S170A54	Mexico	-	-	Zappi, D.
Uebelmannia gummifera	(Backeb. & Voll) Buining	S121A5	Brazil	Minas Gerais	Itamarandiba	Aona, L.Y.S.

Table S2. Biotic variables used to build the predictive model. Discrete variables are coded as A (present) and B (absent). NA = Not available.

Species	**Biome**	**epiderm**			**stem D**	**N_ribs**	**tube r-ri**	**spines**				**flower**		**Fruit**		**pericarpel_**		**seed**				**substrate**	
-	-	green	yello-green	glaucous	-	-	-	leng	grey	yellow	black-brown	leng	diam	leng	diam	red-pink	white-green	leng	diam	glob	hat	rupic	terric
D. bahiensis	dryForest	B	A	A	11.75	14.5	A	2.9	B	B	A	5.6	4.2	3.75	0.9	A	A	1.5	1.7	B	A	A	A
D. boliviensis	dryForest	A	B	B	21.5	12.5	A	2.55	A	B	A	4.9	4.8	2.9	0.9	A	B	1.95	1.85	NA	NA	A	B
D. buenekeri	dryForest	A	B	B	4	16.5	A	0.45	A	B	B	5	5	3	NA	B	A	1.85	2.1	B	A	A	B
D. catingicola	dryForest	A	B	B	15.5	12.5	A	3	A	B	B	5.5	5	4.25	1.05	A	A	1.75	1.75	B	A	A	A
D. catingicola	savanna	A	B	B	15.5	12.5	A	3	A	B	B	5.5	5	4.25	1.05	A	A	1.75	1.75	B	A	A	A
D. diersianus	savanna	A	B	B	26	15.5	A	5.65	B	B	A	5	4.5	3.8	0.85	B	A	1.2	1.7	B	A	A	B
D. fariae-peresii	savanna	A	B	B	19	18.5	A	1.75	A	A	B	10.5	7	3.5	0.7	A	B	2	1.5	A	B	NA	NA
D. ferricola	savanna	A	B	B	25	14	A	4	A	B	A	5.5	3.5	3.5	0.7	B	A	1.5	1.5	B	A	A	A
D. hartmannii	savanna	B	A	B	18.5	17.5	A	2.75	B	B	A	10.25	7	3.75	0.9	A	B	1.85	1.9	B	A	A	A
D. heptacanthus	savanna	B	B	A	13	10	A	2.75	A	B	B	6.35	4.25	2.85	0.9	A	A	1.5	1.4	B	A	A	A
D. horstii	savanna	A	B	B	6	18.5	B	0.325	B	B	A	6.75	6	3	0.4	B	A	1.05	0.95	B	A	A	A
D. psibarbarus	savanna	A	B	B	12	14.5	A	2	B	B	A	7	5.5	3	1	B	A	1.8	1.8	NA	NA	A	B
D. placentiformis	savanna	A	B	B	18.5	17.5	A	2.85	NA	NA	NA	6.25	5.8	4	1	A	A	1.7	1.7	B	A	A	A

D. pseudoinsignis	savanna	A	B	B	16.5	12.5	B	4.2	A	B	B	7.5	6	4.5	0.7	A	B	1.2	1	B	A	B	A
D. zehntneri	dryForest	A	B	B	7	14	A	4.5	A	B	B	5.5	4.5	3.2	0.75	A	A	0.95	1.15	B	A	A	B
D. boomianus	dryForest	A	B	B	7	19.5	A	2.85	A	B	B	5.5	3.8	3.2	0.9	A	B	1.7	1.8	A	B	A	A
D. petr-halfari	dryForest	A	B	B	8.5	12	A	1.75	A	A	B	3.5	2.45	3.5	0.5	A	B	1	1	B	A	A	A

Description of biotic traits obtained for *Discocactus* species to build the predictive model:

epiderme color (discrete) - green; yellowish to greenish.

stemdD (continuous) – mean stem diameter (cm).

N_ribs (continuous) – mean number of ribs

tuber-ri (discrete) - tuberculated ribs; straight ribs.

spines leng (continuous) – mean length of spines

color spine (discrete) - grey; yellowish; black to brown.

flower leng (continuous) – mean flower length (cm).

flower diam (continuous) – mean flower diameter (cm).

fruit leng (continuous) – mean fruit length (cm).

fruit diam (continuous) – mean fruit diameter (cm).

pericarpel color (discrete) - red to pinksh; white to greenish.

seed leng (continuous) – mean seed length (mm).

seed diam (continuous) – mean seed diameter (mm).

seed shape (discrete) - globose; hat-shaped.

substrate (discrete) - rupicolous; terricolous.

CHAPTER 3: Biome shifts or geographical range shifts? Exploring their role in *Melocactus* and *Discocactus* diversification

Milena C. Telhe[1], Monique Romeiro-Brito[2], Fernando F. Franco[1], Thais N. C. Vasconcelos[3], Evandro M. Moraes[1]

1. Departamento de Biologia. Universidade Federal de São Carlos, Sorocaba, São Paulo 18052-780, Brazil.
2. Natural History Department, Florida Museum of Natural History, Gainesville, FL 32611, USA.
3. Ecology and Evolutionary Biology Department, University of Michigan, 500 S State St, Ann Arbor, MI 48109, United States of America.

Draft manuscript

Abstract

The Cactaceae family, renowned for its distinctive succulent growth forms, thrives predominantly in arid and semi-arid regions across the Neotropics. Historically, the Savannas of this region were considered inhospitable to most cactus species due to the prevalence of seasonal fires. However, recent studies have revealed the Neotropical Savanna, particularly the South American Cerrado domain, as the ancestral range for cactus genera widely distributed like *Cereus* and *Pilosocereus*. This suggests that the biotic interchange between Savannas and adjacent semi-arid biomes may have played a pivotal role in their diversification. Although the origin and biome shifts of the sister genera *Melocactus* and *Discocactus* have yet to be investigated, they may exemplify similar dynamics. Both genera exhibit morphological adaptations to fire-prone biomes, such as the Neotropical Savannas, but also occur in fire-free semi-arid biomes. These traits may have facilitated biotic interchanges between these biomes, potentially influenced by the selective pressures of fire regimes. To unravel the intricate dynamic of biome shifts, integrating phylogenetic niche modeling with biogeographic inferences offers a robust approach, as it conjunctly accounts for niche evolution and past climate dynamics. This study employs an integrative approach to explore the role of biome shifts between Neotropical Savannas and fire-free biomes, such as Seasonally Dry Tropical Forests (SDTF), in shaping the diversification of *Melocactus* and *Discocactus*. Pleistocene range shifts between SDTFs and Savannas were identified, though many of these events did not result in niche evolutionary shifts and, therefore, are not characterized as a biome shift. Additionally, no correlation was found between adaptive traits and changes in diversification rates within lineages occupying environments that impose significant survival challenges. The results obtained suggest that finer ecological heterogeneities, such as soil heterogeneity, within seasonal Neotropical biomes may have driven *Melocactus* and *Discocactus* diversification. This highlights the need for more caution in drawing conclusions about biome shifts and their role in lineage diversification.

Keywords: Cactaceae, Cerrado, Caatinga, Biome shifts, phylogenetic niche modeling

4.1 Introduction

Cactaceae is the most celebrated radiation of succulent plants and is remarkable for its variable growth forms (Arakaki et al., 2011, Hernandez-Hernandez et al., 2014). This astonishing family is intrinsically associated with arid and semi-arid Neotropical habitats (Arakaki et al., 2011). Despite their remarkable presence across these habitats, the seasonally dry Neotropical Savannas have traditionally been perceived as avoided by most cactus species due to the challenges imposed by regular fire (Taylor & Zappi, 2004; Ritz et al., 2007; Barthlott et al., 2015). The seasonal fires characteristic of these Savannas pose a significant threat to cacti, as it can lead to fatal overheating of their succulent tissues. (Taylor & Zappi, 2004; Ritz et al., 2007; Barthlott et al., 2015). Although this biome may be unfavorable to cacti survive, recent studies have recovered the largest Neotropical Savanna, the South American Cerrado domain, as an important ancestral range for two species-rich and widely distributed cactus genera (*Cereus* Mill. and *Pilosocereus* Byles & G.D.Rowley), suggesting the biotic interchange between the Savannas and other arid and semi-arid Neotropical biomes as a key factor to their diversification (Amaral et al., 2021; Franco et al., 2017; Lavor et al., 2019; Romeiro-Brito et al., 2023a).

Besides *Cereus* and *Pilosocereu*s, the genera *Melocactus* Link & Otto and *Discocactus* Pfeiff. are among the few cactus living in Neotropical Savannas (Taylor & Zappi, 2004). Although these sister genera (Silva et al., 2018; Romeiro-Brito et al., 2023) exhibit intriguing distribution patterns, they have not yet been investigated regarding their origins and biome shifts (biotic interchanges between biomes). *Discocactus* is predominantly distributed in Savannas, specifically within the Cerrado domain, with few species endemic to the Seasonally Dry Tropical Forests (SDTFs) (Taylor & Zappi, 2004; Flora Brasil, 2020). In contrast, *Melocactus* primarily inhabits SDTFs, which are fire-free biomes (Taylor & Zappi, 2004). Both genera show species with a globose-discoid growth form, a habit suggested as an adaptation promoting resistance to fire (Taylor & Zappi, 2004). Their sister phylogenetic relationship, coupled with their distribution patterns and shared adaptive traits, suggests an interchange between Neotropical Savannas and SDTFs.

Numerous biome shifts in the Neotropical region have been identified and are believed to have played a key role in species adaptation and diversification (Antonelli et al., 2018; Zizka, 2019), including in some Cactaceae lineages (Amaral et al., 2021). Studies investigating the impact of biome shifts have provided valuable evolutionary insights, suggesting that greater

ecological opportunities between biomes in this region may have driven species diversification (Simon & Pennington, 2012; Koenen et al., 2013; Cardillo et al., 2017; Antonelli et al, 2018; Zizka, 2019). Consequently, understanding the role of biome shifts on the diversification rates of Neotropical lineages is essential to understanding the origin and persistence of biodiversity in this area.

Most of the studies investigating biome shifts and their role in the diversification of Neotropical lineages rely on traditional biogeographical analyses, and usually without explicitly testing changes in the diversification rates (e.g. Souza-Neto et al., 2016; Nieto-Blázquez et al., 2017; Antonelli et al., 2018; Amaral et al., 2021; Pérez-Escobar et al., 2022). While these studies provide valuable insights, they often equate shifts in ancestral ranges across different biomes with actual biome shifts (e.g. Antonelli et al., 2018; Pérez-Escobar et al., 2022). However, true biome shifts should involve evolutionary changes that enable lineages to transcend biome boundaries, such as niche evolution (Donoghue & Edwards, 2014). The use of traditional biogeographic analysis to discuss biome shift can pose an even more intricate challenge if we consider that biomes in the Neotropics are composed of mosaics of vegetation types where distinct microhabitats support specialized lineages (Fine & Lohmann, 2018).

Another caveat is that many ancestral range estimation approaches focus on the contemporary ecology of lineages, potentially overlooking niche evolution during its evolutionary history or the effects of past climate fluctuations on species distribution (Donoghue & Edwards, 2014). Furthermore, these approaches often fail to account for how paleo-ecological dynamics may have reshaped the geographical distribution of biomes and how it can impact the ancestral biome estimates (Donoghue & Edwards, 2014; Landis, Edwards & Donoghue, 2021). To better understand biome shifts and their role in species diversification, studies should integrate explicit tests of diversification rate changes, consider the dynamism of ecological niches over time, and incorporate past climate dynamics into ancestral range estimations.

An alternative approach to overcome these challenges is combining biogeographic inferences with phylogenetic niche modeling, offering complementary sources of evidence (Wiens, 2011). Phylogenetic niche modeling integrates ecological niche modeling and dated phylogenies to examine the niche evolution of lineages (Evans et al., 2009). Recently, Guillory & Brown (2021) introduced the R package ‘Machuruku’, designed to infer ancestral distributions and model phylogenetic niches with precise quantification of lineage responses to

climate, allowing the projection of ancestral niches into the paleoclimate data. This method, when used alongside traditional biogeographical inferences, allows for a comprehensive investigation of biome shifts by incorporating niche evolution and past climate fluctuations in the Neotropics (Guillory & Brown, 2021).

To gain deeper insights into the factors contributing to the diversity of seasonally dry and semi-arid Neotropical regions, we investigated the role of biome shift events between fire-prone and fire-free environments in the diversification of the cactus sister genera *Melocactus* and *Discocactus*. By integrating biogeography inferences, phylogenetic niche modeling, adaptive trait analysis, and diversification rates estimates, we aim to determine when and where these biome shifts occurred and assess their potential influence on the diversification of those two sister lineages.

4.2 Methods

4.2.1 Data Sampling and Processing

Our focal group corresponds to two genera of the Cereinae subtribe (Cereeae tribe; sensu Romeiro-Brito et al., 2023), *Melocactus* and *Discocactus*, and its sampling comprises 45 specimens corresponding to 72% of *Melocactus* and 93% of *Discocactus* currently described diversity (*sensu* Flora e Funga do Brasil, 2025; Hunt et al., 2006). To improve the accuracy of the diversification rate estimates, a broader taxonomic sampling beyond our two focal genera was conducted, addressing the inverse relationship between clade size and variance in evolutionary rate estimates (Rabosky, 2019). Therefore, genomic data were sampled from 174 species, which included 31 species from 18 genera as outgroups, representing the Opuntioideae and Cactoideae subfamilies (Cacteae tribe, Notocactaceae tribe, Cereeae tribe), and 143 species from 12 genera within the Cereinae subtribe. Our main hypotheses focus on the role of biome shifts in the diversification of *Melocactus* and *Discocactus*, both of which exhibit fire-adapted traits. Therefore, all analyses performed to address this question were conducted using a pruned phylogeny including only these two genera. The genomic data used were previously generated in recent studies (Romeiro-Brito et al 2022, Romeiro-Brito et al., 2023; Telhe et al., 2025; Telhe et al., *in prep*). These data were generated by target sequencing of genomic libraries hybridized with the Cactaceae591 probe-set (Romeiro-Brito et al 2022), which targets the exons and introns of nuclear genes and two plastid genes.

All raw sequencing reads were pre-processed using *fastp* v. 0.23.3 (Chen et al., 2018) to remove adapters, low-quality reads (phred < 20), and reads shorter than 60 bp. After pre-processing the data, the reads were mapped and assembled using *HybPiper* v. 2.1.6 (Johnson et al., 2016) and the reference sequences from Romeiro-Brito et al. (2022). Although the Cactaceae591 probe set targets single-copy regions, some genes may be duplicated in certain taxa. Thus, we followed the workflow outlined in Romeiro-Brito et al. (2022) to identify and exclude putative paralogs. The final dataset was aligned using *MAFFT* v. 7.520 (Katoh & Standley, 2013), and poorly aligned and indel-rich regions were trimmed with *TrimAL* v.1.4 (Capella-Gutiérrez et al., 2009) using an automatic selection algorithm (*-strict*). To further refine the dataset, we filtered hypervariable regions and outliers sequences using the *spruceup* (Borowiec, 2019) with a lognormal distribution and a cutoff of 0.95. Finally, to minimize the effects of missing data, we allowed a maximum of 20% of missing samples for each nuclear region.

Phylogenetic inference and Divergence Time Estimation

Phylogenetic inferences were conducted using a multispecies coalescent model to address gene tree discordance, a common issue in the Cactaceae family (Copetti et al., 2017; Romeiro-Brito et al., 2023). We employed the coalescent summary approach implemented in *weighted Astral* (*wASTRAL*; Zhang & Mirarab, 2022) with the *astral-hybrid* function, assuming low bootstrap values below 70. Independent gene trees, used as input on *wAstral*, were inferred using *IQTree* v. 2.0.7 (Minh et al., 2020) with the *ModelFinder Plus* fuction and the ultrafast bootstrap with 10,000 replicates.

Divergence times were estimated using the relative rate framework implemented in *RelTime* (Tamura et al., 2012) available in *MEGA11* (Tamura, Stecher & Kumar, S, 2021). This framework reduces the computational latency and, therefore, *RelTime* was chosen to handle the full genomic dataset efficiently without computational limitations. Since there are no fossil records for the Cactaceae family, we performed secondary calibration assuming the ages previously estimated by Hernandez-Hernandez et al (2014) for the Cactoideae subfamily (crown age: 10.94–21.85 Ma) with a normal distribution and including all loci from our final dataset. The tree estimated by *wASTRAL* was used as input in this analysis.

4.2.2 Ancestral range estimations and phylogenetic niche reconstruction

Geographical occurrence data for all species from our focal group were retrieved from GBIF (http://www.gbif.org) and *speciesLink* (https://specieslink.net/search/) databases. To remove duplicates and any spatial errors (e.g., non-terrestrial records, herbarium entries, and zero latitude or longitude points), occurrences were filtered using the 'CoordinateCleaner' package (Zizka et al., 2019) in R v. 4.3.1 (R Core Team, 2023). The retained records were plotted on a biome map and manually inspected to ensure they matched known species distribution (*sensu* Flora e Funga do Brasil, 2025; Barthlott et al., 2015). Any record outside of the expected distribution range was excluded.

Ancestral range estimations were conducted using a maximum-likelihood approach implemented in 'BioGeoBEARS' (Matzke, 2013). Six models were evaluated, including DEC; DIVA-like, and BayArea-like, accounting or not for the "jump speciation" parameter. Model selection was based on Akaike Information Criterion (AIC) scores. We defined three operational areas based on the Neotropical biomes described by Dinerstein et al (2017): (1) Deserts & Xeric Shrublands, (2) Tropical & Subtropical Dry Broadleaf Forests (3) Tropical & Subtropical Grasslands, Savannas & Shrublands. Species distributions were assigned to these areas based on the presence of at least 50% of their geographic records within one of the operational areas, using the R package 'speciesgeocodeR' (Töpel et al., 2017). Habitat characteristics and known biome distribution of each species from the literature (e.g. Taylor & Zappi, 2004; Flora e Funga do Brasil, 2025; POWO, 2024) were used to validate each classification. The tree inferred with the *RelTime* method on *MEGA11* was used as input for this analysis after pruning to retain only the *Melocactus* and *Discocactus* clades.

To determine if the ancestral range shifts inferred from the previous analysis corresponded to evolutionary changes that enabled lineages to transcend biome boundaries (i.e., niche shifts), we reconstructed the niche evolution using the 'Machuruku' R package (Guillory & Brown, 2021). This package enables the inference of ancestral distributions and phylogenetic niche modeling, allowing precise quantification of lineage responses to climate. We collected 19 climatic variables at a 2.5 arcsec (~5km) resolution from the *WorldClim* v2.1 database (Fick & Hijmans, 2017). Variables contributing at least 40% to each species model were selected using the function *machu.top.env* with the "contrib" method. The retained variables were then used to estimate the species response curves with the function *machu.1.tip.resp* and to estimate species niche parameters using the function *machu.2.ace* at four paleoclimate periods: Mid-Holocene (8.326-4.2 ka; Fordham et al., 2017), Last Glacial Maximum (21 ka; Karger et al.,

2021), Last Interglacial Maximum (130 ka; Otto-Bliesner, et al., 2006), mid-Pliocene Warm Period (3.205 Ma; Hill, 2015). Species with fewer than 10 geographical occurrences, required by *machu.1.tip.resp*, had additional occurrences randomly generated. A 5km buffer was created around each existing occurrence using the *buffer* tool in *QGis* v.3.32.2 (QGIS Development Team, 2023), and random points were generated within this buffer using the tool *Random points in polygons* and ensuring a minimum distance of 2km between points.

To evaluate how different species' responses are along the environmental gradient (niche dissimilarity), we employed the Bhattacharyya coefficient (Bhattacharyya, 1943, Winner et al., 2018) to compare the response curves estimated by 'Machuruku' between two lineages along a single climatic variable. A species pair with a Bhattacharyya coefficient lower than 0.3 for at least one climatic variable was considered to have different niches.

This combined approach allowed us to incorporate both the historical climatic conditions and spatial distribution of species into our analysis of ancestral range shifts and niche evolution in *Melocactus* and *Discocactus*.

4.2.3 Correlations between speciation rate, trait, and biomes

To investigate the association between adaptive traits, biome shifts, and diversification rates, we applied the Missing State Speciation and Extinctions (MiSSE) method (Vasconcelos, O'Meara, & Beaulieu, 2022) to estimate diversification rates and subsequently correlate these rates with trait and biome shifts.

First, fire resilience traits were scored for each species in our focal group by documenting the presence or absence of morphological adaptations to fire, the depressed-globose habit, based on a comprehensive review of the literature and taxonomic descriptions (see Table S1 for detailed trait data). Second, biome shifts were quantified for each lineage by examining range shifts inferred from the 'BioGeoBEARS' analysis in conjunction with niche shifts estimated by the 'Machuruku' analysis and Bhattacharyya's coefficient. Both range shifts and biome shifts were scored for each lineage by evaluating all nodes from root to tip.

Diversification rates were estimated using the R package 'hisse' with the *MiSSE* method, which is an extension of the *HISSE* framework and allows for trait-independent diversification rate estimation (Vasconcelos, O'Meara, & Beaulieu, 2022). The models were structured to allow both turnover and extinction fractions to vary between rate classes. We

accounted for incomplete taxon sampling with a sampling fraction of 0.72. We employed the *MiSSEGreedy* function to streamline model selection, evaluating chunks of four models at a time. The model selection process was stopped when the difference in the AICc between successive model chunks was less than two units (stop.deltaAICc = 2), indicating that further models would not substantially improve the fit. Redundant models were then pruned, and the tip diversification rates were estimated based on model averaging across the non-redundant models.

Once the tip diversification rates were estimated, we explored the association between trait states, biome occurrence, and speciation rates using phylogenetic ANOVA, implemented with the function *phylANOVA* from the R package 'phytools' (Revell, 2012). Additionally, to assess the correlation between biome shifts and speciation rate across lineages, we performed a phylogenetic logistic regression using the *phyloglm* function from the R package 'phylolm' (Ho et al., 2018). This analysis allowed us to test whether the changes in diversification rates are associated with biome shifts, providing insights into the role of biome shifts in shaping the evolutionary history of the focal group.

4.3 Results

4.3.1 Data Sampling, Phylogenetic inference, and Divergence Time Estimation

After processing the raw data, excluding regions with more than 20% missing samples, and removing putative paralogs, the final dataset consisted of 398 regions, comprising 663,958 base pairs (median per region = 1459 bp). Phylogenetic analyses based on this dataset yielded well-supported relationships across the tree (LPP > 0.9; Fig. 1). The tribe Cereeae was recovered as a well-supported clade, with strong support for its major subtribes (LPP = 0.99). Within the subtribe Cereinae, most clades also displayed high support (LPP > 0.9), including the focal Melocactus-Discocactus clade (LLP = 1).

Divergence time estimations revealed that the crown age of the Cereeae tribe is approximately 7.11 Ma (95% CI: 11.4 - 4.4 Ma; Fig. 1, Table S2). Major clades within the tribe Cereeae, including those corresponding to the subtribes Trichocereinae, Gymnocalycinae, and Cereinae, diverged during the Miocene, while most Cereinae genera underwent diversification during Pliocene-Pleistocene (e.g., *Discocactus*; Fig. 2) and the Pleistocene (e.g., *Melocactus*; Fig. 2).

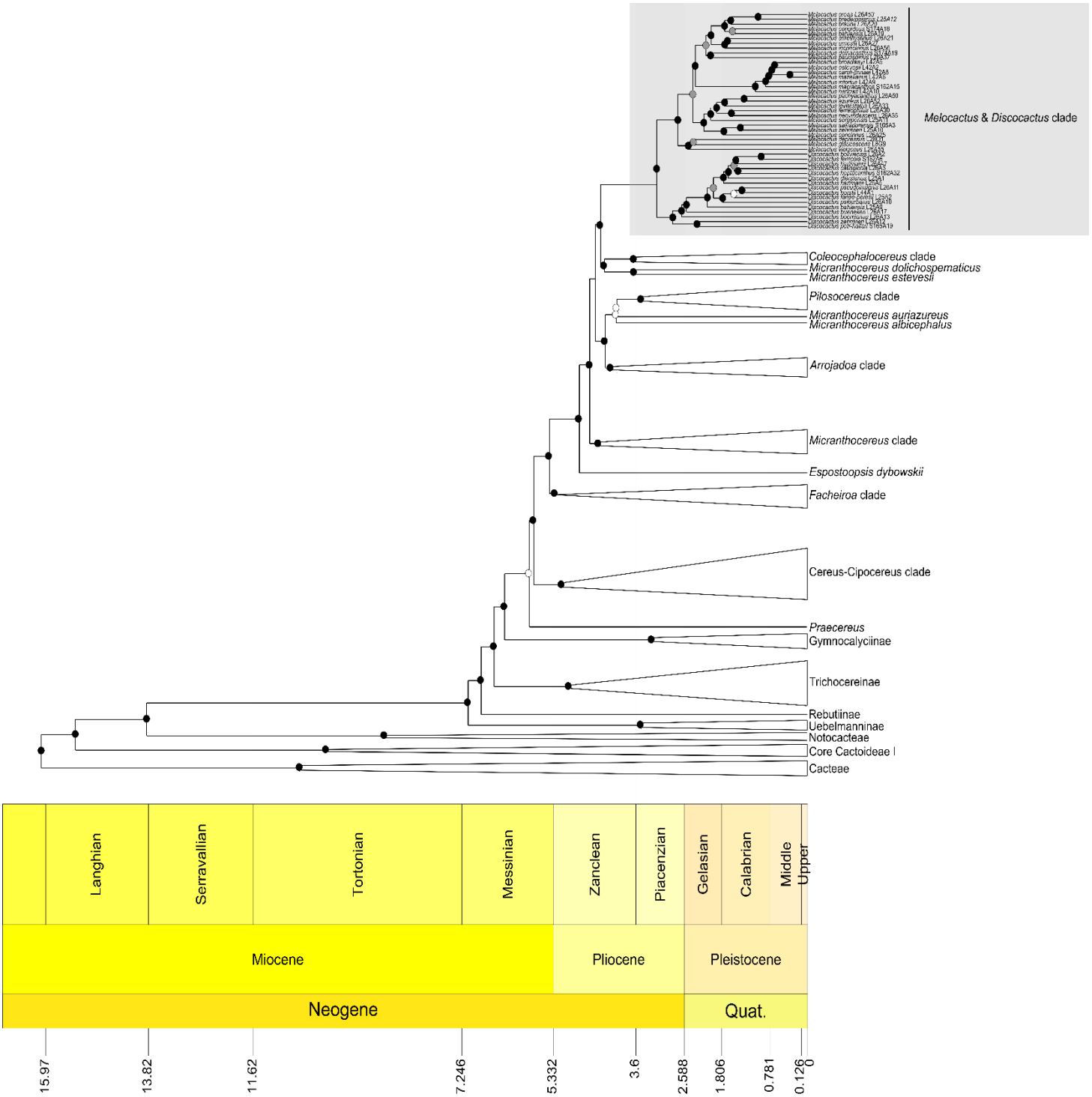


Figure 1. The divergence time of the tribe Cereeae estimated with the *RelTime* method on *MEGA11*. The Melocactus-Discocactus clade is highlighted in a gray box.

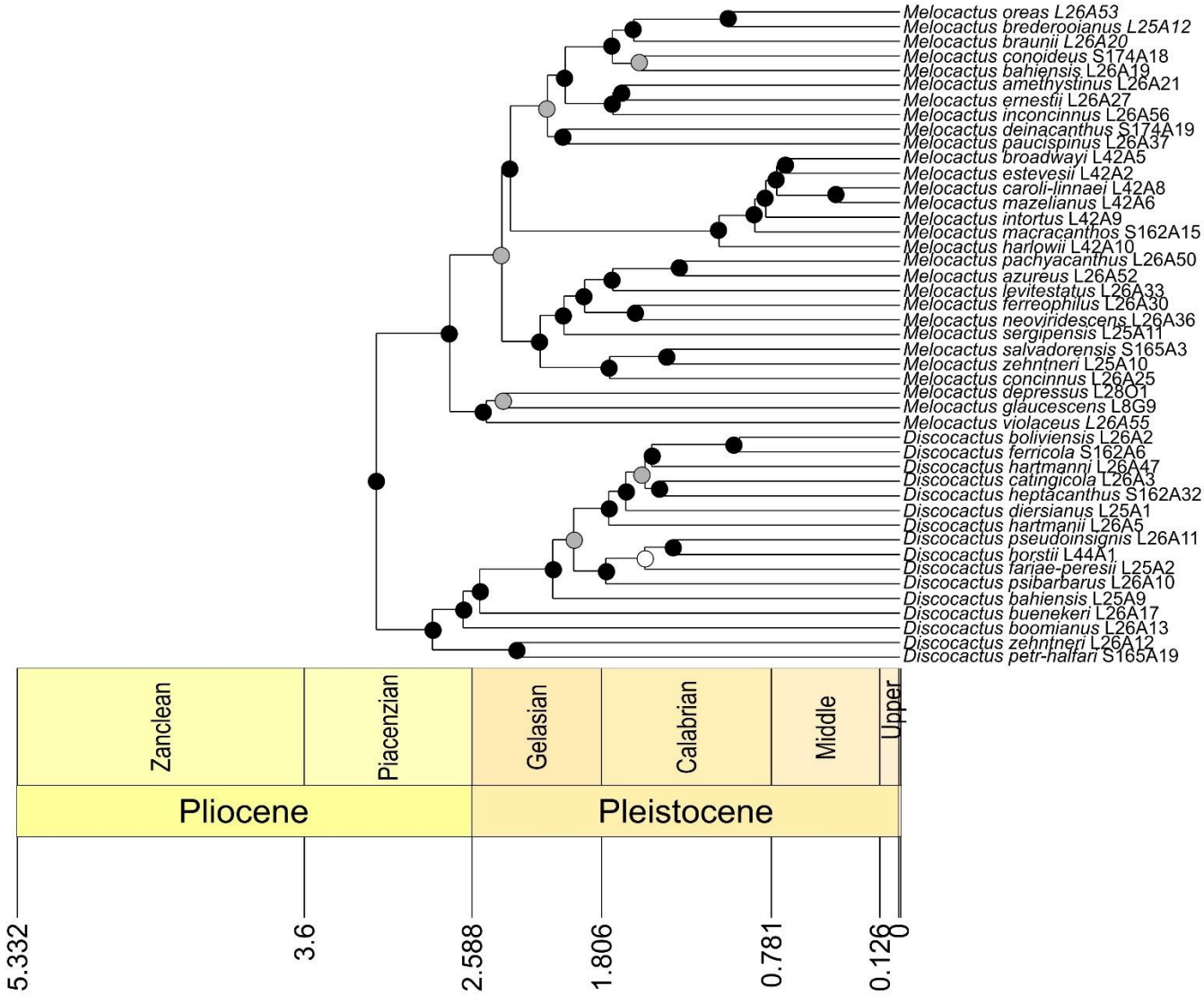


Figure 2. Zoom on the *Melocactus* & *Discocactus* clade from Figure 1. The divergence time of *Melocactus* and *Discocactus* estimated with the *RelTime* method on *MEGA11*. The sample code of each specimen is indicated after the species names.

4.3.2 Ancestral range estimations

The DEC+J model was selected as the best-fit biogeographical model based on AIC weights (AICwt = 1.0), with DEC+J and DIVA+J models having AIC values with less than two units of difference. The ancestral area of the Melocactus-Discocactus clade is most likely restricted to SDTFs (Fig. 3). The divergence between *Melocactus* and *Discocactus*, took place in SDTF. Approximately 1.6 Ma, *Melocactus* lineages experience a jump dispersal from SDTFs to Savannas (e.g., *M. conoideus*), followed by range expansions where present-day lineage inhabits both biomes (e.g. *M. bahinesis*). The clade comprising *M. harlowii-M. broadwayi* lineages experienced most of the range shift events within the genus, with multiple jump

dispersal from SDTF to both Savannas and Deserts & Xeric Shrublands (e.g., *M. estevessii*, *M. mazelianus*, *M. macracanthos*). All other *Melocactus* species currently found outside SDTF occupy both SDTF and Savannas (e.g., *M. concinnus*, *M. zehntneri*, *M. paucispinus*), having expanded their range to Savanna from an ancestral SDTF distribution.

Discocactus also had an ancestral range in SDTF, with its first derived lineages retaining this range. Around 1.8 Ma, a jump dispersal allowed *Discocactus* to colonize the Savannas, where most species are currently distributed. *Discocactus catingicola* represents a case of range expansion back to SDTF, which currently occupies both biomes. A jump-back dispersal to SDTF also occurred during the divergence of *D. boliviensis*, which is now endemic to this biome.

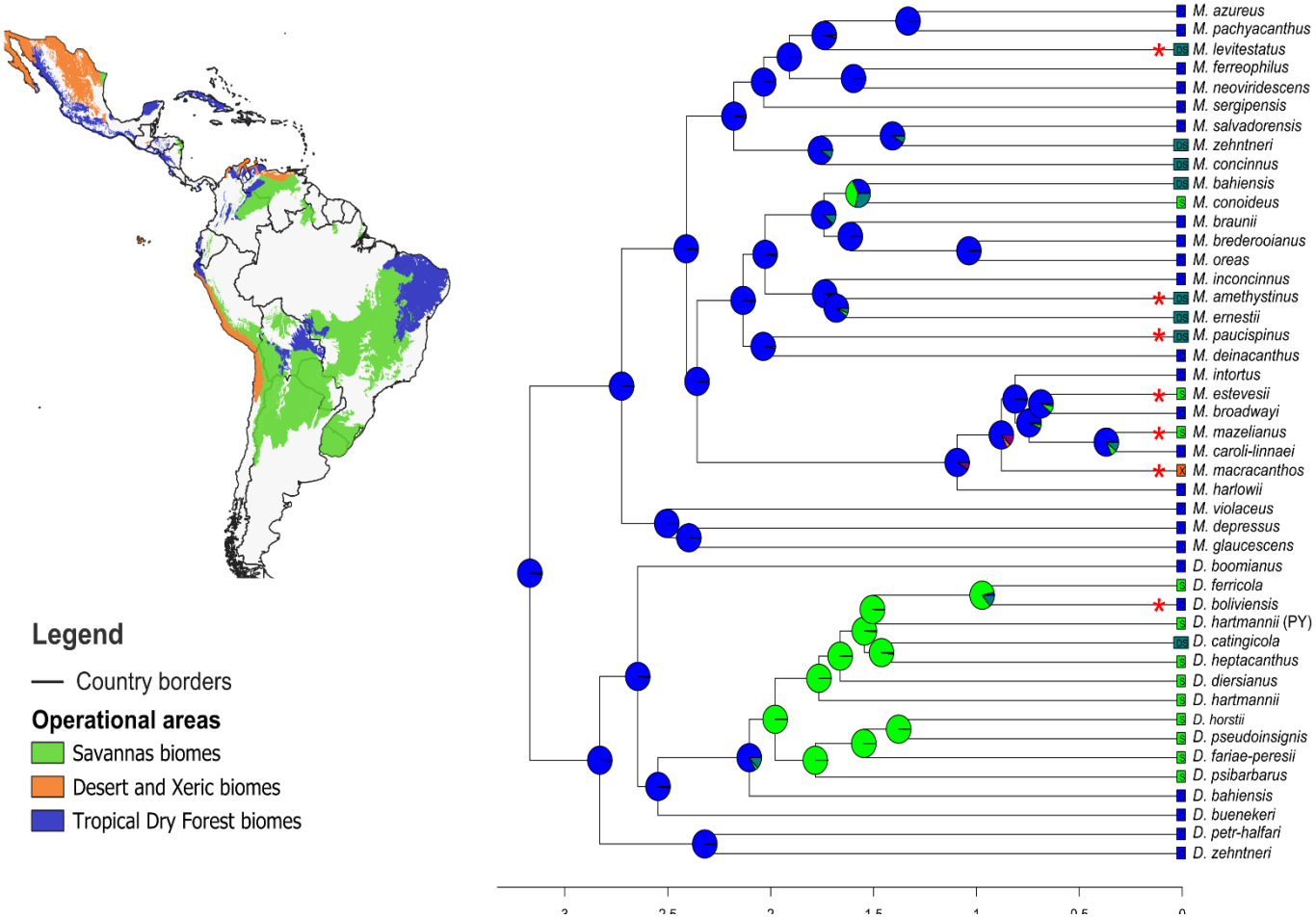


Figure 3. Biogeographic reconstruction of *Melocacus* and *Discocactus* using *BioGeoBEARS* R package. The red asterisk (*) indicates range shifts that correspond to niche shifts based on machuruku analysis. The *D. hartmannii* specimen from Paraguay is followed by the PY (Paraguay abbreviation) for its differentiation

4.3.3 Correlations between speciation rate, trait, and biomes

No significant shifts in speciation rates were identified between major clades (Fig. S1; Table S3). However, lineages with lower to moderate speciation rates ($\lambda < 0.21$; Table S3) were predominantly found in the *Praecereus*, *Cereus-Cipocereus*, *Brasilicereus*, and *XiqueXique* clades. In contrast, higher speciation rates were observed in the *Pilosocereus* clade and the focal group comprising *Discocactus* and *Melocactus*, with the clade comprising *M. harlowii-M. broadwayi* lineages exhibiting particularly high rates ($0.41 < \lambda < 5.36$; Table S3).

Ancestral range estimation for the *Melocactus* and *Discocactus* identified 25 range shifts, including range expansions and jump dispersals (Table S4). Of these, only seven range shifts corresponded to climatic niche shifts, as estimated by *machuruku* (shifts in response curves for variables used in the models; Fig. S2-S15 and TableS5), indicating biome shifts. Notably, only one *Disocactus* lineage experienced a biome shift, while all other biome shift events were experienced by *Melocactus*, with the clade comprising *M.harlowii-M.intortus* lineages undergoing most of the shift events. Despite this, there was no correlation (p-value > 0.05) between the occurrence of biome shifts and changes in diversification rates (Fig. 4).

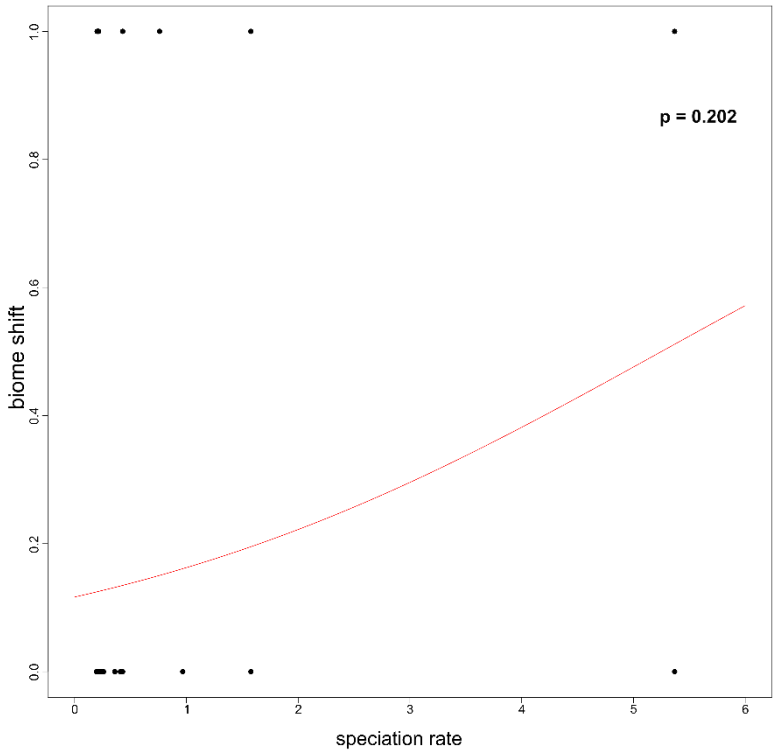


Figure 4. Phylogenetic logistic regression of speciation rates and the occurrence of biome shifts in the *Melocactus* and *Discocactus* lineages.

When comparing fire-resistance traits and their association with speciation rates in fire-prone biomes (Savannas), the boxplot suggests that lineages exhibiting these adaptive traits display higher speciation rates compared to those lacking such fire-adapted traits (Fig. 5). However, the phylogenetic ANOVA revealed no statistically significant difference in speciation rates between these groups (p = 0.6; Table 1).

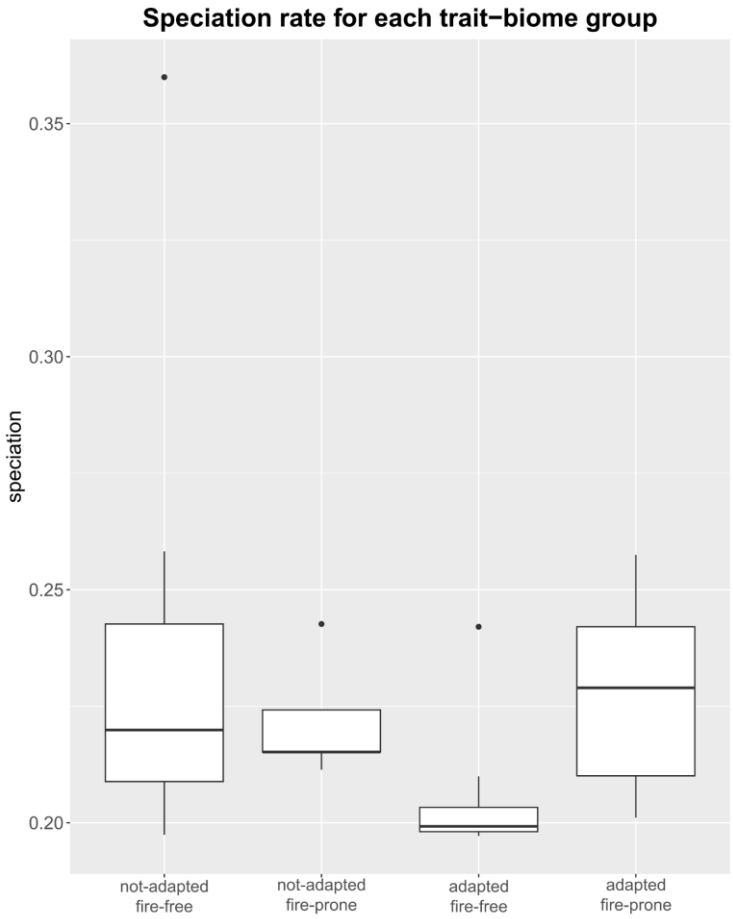


Figure 5. Boxplots showing the distribution of speciation rates across lineages with varying traits, grouped by their occurrence in different biomes.

Table 1. Results of phylogenetic ANOVA assessing the correlation between speciation rates and trait differences among lineages occurring in different biomes.

Anova parameters	**Sum Sq**	**Mean Sq**	**F value**	**Pr(>F)**
x	5.219114	1.739705	1.505169	0.6201
Residual	47.388627	1.155820		

4.4 Discussion

4.4.1 Phylogenetic inferences and ancestral range estimations

Genomic data have increasingly been utilized in plant evolutionary studies (Sun et al., 2022), providing robust phylogenetic resolutions even in lineages with recent and rapid diversification, in which incomplete lineage sorting is a huge source of gene tree discordance (Acha & Majure, 2022; Lagomarsino et al., 2022; Romeiro-Brito et al., 2022; Taylor et al., 2023). In this study, the target sequencing approach effectively recovered well-supported phylogenetic relationships within the tribe Cereeae, corroborating earlier findings from genetic analyses (Hernández-Hernández, et al., 2014; Silva et al., 2018; Romeiro-Brito et al., 2023). Unraveling the phylogenetic relationships between lineages is a prerequisite to any evolutionary studies. Advances in genomic techniques, particularly the development of customized probe sets like Cactaceae591, have enabled the reconstruction of well-supported phylogenetic hypotheses (e.g., Romeiro-Brito et al., 2023; Taylor et al., 2023; Zappi et al., 2024). These phylogenetic hypotheses are essential for addressing questions related to lineage origins, trait evolution, and diversification rates, which are key focal points of this study.

Our divergence time estimates align with previous studies on the Cactaceae family (Hernández-Hernández et al., 2014), although we recovered generally older divergence dates for the Cereinae subtribe. These results are consistent with those estimated by Silva et al. (2018) and Romeiro-Brito et al. (*in prep.*). The Cereeae tribe originated in the Miocene, an epoch marked by a global decline in atmospheric CO_2 and a shift toward cooler, dryer climates (Tripati et al., 2009; Herbert et al., 2016). These environmental changes were associated with the expansion of arid habitats in the New World and the origin and diversification of succulent lineages (Arakaki et al., 2011; Hernandez-Hernandez et al., 2014). Several genera within the tribe Cereeae diversified primarily during the Pliocene-Pleistocene, with *Melocactus* and *Discocactus* diversifying in the Pleistocene. Other studies in Cactaceae using genetic or genomic sequencing data have recovered the diversification of genera such as *Pilosocereus* Byles & G.D.Rowley (Bonatelli et al., 2014; Menezes et al., 2016; Franco et al., 2017; Romeiro-Brito et al., 2023a), *Cereus* Mill (Franco et al., 2017), and *Mammilaria* Haw. (Cornejo-Romero et al., 2014) during this epoch, indicating a significant period of cactus diversification.

The diversification of most *Melocactus* and *Discocactus* lineages during the Pleistocene and the estimation of the Seasonally Dry Tropical Forests (SDTF) as their ancestral area

coincides with the climatic and environmental changes of the Pleistocene, an epoch characterized by the expansion of arid habitats. The Quaternary climatic oscillations further contributed to vegetation dynamics, causing SDTFs to expand and contract in response to the climatic fluctuating conditions (Ab'saber, 1977; Pennington et al., 2000; Carnaval et al., 2014; Sobral-Souza et al., 2015). In this period, when experiencing an expansion, the SDTF may have reached up to the surrounds of a core area of the Cerrado domain (a Savanna biome) in central Brazil (Ab'saber, 1977; Prado & Gibbs, 1993). This dynamic is supported by historical vegetation distributions and phylogenetic patterns in other SDTF lineages (Prado & Gibbs, 1993; de Queiroz et al., 2017; Corbett et al., 2020; Hurbath et al., 2021), although some studies suggest different interpretations (e.g., Werneck et al., 2011). The historical climatic instability and shifting vegetation likely explain the biogeographic patterns observed in this study, such as the range expansions of *Discocactus* from SDTF to Savanna. During periods of favorable climatic conditions, lineages may have expanded into new areas like the Savanna biome, taking advantage of the temporary expansion of suitable habitats. These shifts in vegetation likely facilitated the dispersal and establishment of lineages in regions beyond their ancestral ranges.

Our findings add to the growing body of evidence that the origin and assembly hypothesis of the Cerrado domain, a neotropical Savanna biome (Jaramillo, 2023), is more recent than the formation of the adjacent biomes. Several studies based on phylogenetic age estimates have suggested that many Cerrado lineages diversified over the last 4 Ma and are derived from adjacent regions, as Amazonian forest and the Caatinga domain (e.g. Simon et al 2009; Souza-Neto et al., 2016; Azevedo et al., 2020). The range expansion of *Discocactus* into the Savanna, followed by *in situ* diversification, contributed to the plant species endemism and diversity found in the Cerrado domain. It is worth mentioning that while we infer geographic range shifts from SDTF to Savanna, we do not conclude ecological niche shifts - this will be discussed further in the subsequent section.

4.4.2 Biome Shifts, Traits, and Diversification

We combined biogeographic inferences with phylogenetic niche modeling to overcome the challenges of inferring biome shifts using traditional biogeographical methods alone. This integrative approach allowed us to investigate biome shifts while accounting for niche evolution and the effects of past climate oscillations. By combining these analyses, we recovered fewer biome shifts compared to the results from purely biogeographic methods. Moreover, we found no significant correlation between biome shifts and the diversification

rates of *Melocactus* and *Discocactus* lineages. Even when taking a closer look into the role of biome shifts, investigating environmental challenges to the survival of Cactaceae species (e.g. fire) and the presence of adaptive traits, we could not find a positive correlation between the presence of these traits and differences in the diversification rate, contrasting with findings in other plants groups (e.g. Koenen et al., 2013). Given the environmental heterogeneity of Neotropical biomes, characterized by a mosaic of vegetation types, soils, geomorphology, geology, and microclimates (Moreira et al., 1999; Coe & Sousa, 2014), it is possible that *Melocactus* and *Discocactus* lineages experienced geographic range shifts into new areas with microhabitats that resembled the climate conditions similar to their ancestral.

Rather than being driven by biome shifts, the in-situ diversification of *Melocactus* and *Discocactus* within the SDTF and Savanna, respectively, appears to have been shaped by dynamic vegetation changes—both expansions and contractions—combined with the emergence of diverse topographic and edaphic conditions (Fernandes et al., 2022, Pérez-Escobar et al., 2022). These environmental dynamics were influenced by the expansion of arid conditions during the late Pliocene (Amidon et al., 2017) and subsequent climate cycles throughout the Quaternary (Schaefer et al., 2023). The origin and diversification of *Melocactus* and *Discocactus*, dating mostly to the Pleistocene and late Pleistocene, coincided with significant geological changes in northeastern Brazil. During this period occurred the formation of the continental sand dunes (Barreto and Suguio, 1993; Mabesoone, 1994; Barreto, 1996), driven by changes in the course of Rio São Francisco due to the Araripe plateau uplifting and the wind action (Mabesoone, 1994; Maia et al., 2010; Barros-Corrêa et al., 2019), alongside occurred the exhumation of karstic systems in northeastern Brazil (Silva et al., 2017). Additionally, during the Quaternary, processes such as pediplanation further differentiated sedimentary and crystalline soils, exposing additional crystalline surfaces (Ab'saber 1974; Bigarella et al., 2016). Particularly, the rise of these heterogeneous edaphic conditions across arid habitats in the Neotropical region, including the Caatinga domain (SDTF) and Cerrado domain (Savanna) (Salgado et al., 2015; Steinke, 2021), likely provided further ecological opportunities for lineages diversification.

Many *Melocactus* and *Discocactus* species exhibit edaphic specialization, being endemic to specific substrates (e.g., *M. violaceus* and *D. pseudoinsignis* on sand; *M. levitestatus* and *D. boliviensis* on limestone). Edaphic specialization has already been recognized as a driver of genetic structure, diversification, and endemism in several plant lineages (Rajakaruna, 2018;

Hulshof & Spasojevic, 2020). The edaphic heterogeneity across the dry habitats studied here is characterized by edaphic conditions derived from distinct parental materials, creating variability in soil fertility, texture, and water retention. These differences offer ecological opportunities that may rapidly promote edaphic specialization, ultimately leading to diversification (Rajakaruna, 2004; Rajakaruna, 2018; Azevedo et al., 2023).

The presence of rocky outcrops in these biomes has also been suggested to play a key role in the diversification of cactus lineages (e.g., Bonatelli et al., 2014; Menezes et al., 2016; Perez et al., 2016; Juárez-Miranda et al., 2021). During wetter interglacial periods of the Pleistocene, these rocky outcrops might have worked as refugia for dry-adapted lineages (e.g., Bonatelli et al. 2014; Maswangany et al., 2017; de Aguiar-Campos et al., 2020). The combination of Pleistocene climatic oscillations, which drove the expansion and contraction of lineage distributions, along with increasingly heterogeneous edaphic conditions, may have triggered the diversification of *Melocactus* and *Discocactus*.

In recent years, there has been a growing interest in understanding the role of biome shifts in shaping the Neotropical diversity (Ledo & Colli, 2017; Antonelli et al., 2018; Fine & Lohmann, 2018; Rull, 2020, and references therein). Although these studies have provided valuable insights, most focus on biome shifts was inferred through traditional biogeographical analyses, which primarily consider the contemporary ecology of lineages when inferring past range shifts (Antonelli et al., 2018; Pérez-Escobar et al., 2022). These approaches often fail to account for the influence of paleo-ecological dynamics on range evolution (Landis, Edwards & Donoghue, 2021). Furthermore, Neotropical biomes are not homogeneous units but rather mosaics of vegetation types, suggesting the existence of distinct habitat patches within a biome that sustains lineages adapted to them (Fine & Lohmann, 2018).

If we had relied solely on traditional biogeographical methods, we would have inferred a higher number of biome shifts in *Melocactus* and *Discocactus*, likely overestimating the role of biome shifts in their diversification. This underscores the importance of caution when drawing conclusions about biome shifts and their contribution to lineage diversification. It is essential to distinguish between geographic shifts and true biome shifts, the latter must involve evolutionary changes, such as niche evolution. As Donoghue & Edwards (2014) suggested, addressing this question through multiple and independent lines of evidence, particularly by considering past environmental conditions, can provide more robust insights into the role of biome shifts.

4.5 Acknowledgments

We are deeply grateful to Dr. Marlon Machado, Dr. Diego Gonzaga, and the *Coleção de Cactos e Suculentas* at the Instituto de Pesquisas Jardim Botânico do Rio de Janeiro for donating cactus samples used in this study. The authors declare that the publication of this article does not involve any conflicts of interest.

4.6 Funding

This study was financially supported by FAPESP, which provided fellowships for doctoral students (Processes 2019/11233-2 and 2022/11683-0 to MCT and 2018/06937-8 to MRB) and research grants (2019/03211-9 to EMM, 2020/15161-3 to FFF and 2018/03428-5 to EMM and FFF). DCZ and EMM hold productivity grants from CNPq (304178/2021-7 and 310963/2022-6, respectively).

4.5 References

Ab'Saber, A. N. (1974). O domínio morfoclimático semi-árido das caatingas brasileiras. *Geomorfologia*, (43), 1-39.

Ab'Sáber, A. N. (1977). *Espaços ocupados pela expansão dos climas secos na América do Sul, por ocasião dos períodos glaciais quaternários.* Universidade de São Paulo, São Paulo, Brasil.

Acha, S., & Majure, L. C. (2022). A new approach using targeted sequence capture for phylogenomic studies across Cactaceae. *Genes*, *13*(2), 350.

Amaral, D.T., Minhós-Yano, I., Oliveira, J.V.M., Romeiro-Brito, M., Bonatelli, I.A.S., Taylor, N.P., Zappi, D.C., Moraes, E.M., Eaton, D.A. and Franco, F.F. (2021). Tracking the xeric biomes of South America: The spatiotemporal diversification of Mandacaru cactus. *Journal of Biogeography*, *48*(12), 3085-3103.

Amidon, W.H., Fisher, G.B., Burbank, D.W., Ciccioli, P.L., Alonso, R.N., Gorin, A.L., Silverhart, P.H., Kylander-Clark, A.R. and Christoffersen, M.S. (2017). Mio-Pliocene aridity in the south-central Andes associated with Southern Hemisphere cold periods. *Proceedings of the National Academy of Sciences*, 114(25), 6474-6479

Antonelli, A., Zizka, A., Carvalho, F. A., Scharn, R., Bacon, C. D., Silvestro, D., & Condamine, F. L. (2018). Amazonia is the primary source of Neotropical biodiversity. *Proceedings of the National Academy of Sciences*, *115*(23), 6034-6039.

Arakaki, M., Christin, P. A., Nyffeler, R., Lendel, A., Eggli, U., Ogburn, R. M., ... & Edwards, E. J. (2011). Contemporaneous and recent radiations of the world's major succulent plant lineages. *Proceedings of the National Academy of Sciences*, *108*(20), 8379-8384.

Azevedo, J.A., Collevatti, R.G., Jaramillo, C.A., Strömberg, C.A., Guedes, T.B., Matos-Maraví, P., Bacon, C.D., Carillo, J.D., Faurby, S. and Antonelli, A. (2020). On the young savannas in the land of ancient forests. *Neotropical diversification: Patterns and processes*, 271-298.

Azevedo, L., Zappi, D.C., de Oliveira, D.M.G., Meyer, L., Nic Lughadha, E., Clegg, R., Meireles, L.D., de Melo, P.H.A., Pennington, R.T. and Neves, D.M (2023). On the rocks: Biogeography and floristic identity of rocky ecosystems in eastern South America. *Journal of Systematics and Evolution*

Bhattacharyya, A. (1943). On a measure of divergence between two statistical populations defined by their probability distribution. *Bulletin of the Calcutta Mathematical Society*, *35*, 99-110.

Barreto, A. M. F. (1997). *Interpretação paleoambiental do sistema de dunas fixadas do médio Rio São Francisco, Bahia* (Tese de Doutorado). Universidade de São Paulo, São Paulo, Brasil.

Barreto, A. M. F., & Suguio, K. (1993). Considerações sobre a idade e a paleogeografia das paleodunas do médio Rio São Francisco, Bahia. In *Congresso da Associação Brasileira de Estudos do Quaternário-ABEQUA, IV, São Paulo. Resumos Expandidos* (p. 11).

Barros Corrêa, A. C., de Azevêdo Cavalcanti Tavares, B., de Lira, D. R., da Silva Mutzenberg, D., & de Souza Cavalcanti, L. C. (2019). The semi-arid domain of the Northeast of Brazil. *The Physical Geography of Brazil: Environment, Vegetation and Landscape*, 119-150.

Barthlott, W. (2015). *Biogeography and biodiversity of cacti*. Univ.-Verlag Isensee.

Bigarella, J. J., de Meis, M. R. M., & da Silva, J. X. (2016). Pediplanos, pedimentos e seus depósitos correlativos no Brasil. *Espaço Aberto*, *6*(2), 165-196.

Bombonato, J.R., do Amaral, D.T., Silva, G.A.R., Khan, G., Moraes, E.M., da Silva Andrade, S.C., Eaton, D.A., Alonso, D.P., Ribolla, P.E.M., Taylor, N. and Zappi, D. (2020). The potential of genome-wide RAD sequences for resolving rapid radiations: a case study in Cactaceae. *Molecular phylogenetics and evolution*, *151*, 106896.

Bonatelli, I.A., Perez, M.F., Peterson, A.T., Taylor, N.P., Zappi, D.C., Machado, M.C., Koch, I., Pires, A.H. and Moraes, E.M. (2014). Interglacial microrefugia and diversification of a cactus species complex: phylogeography and palaeodistributional reconstructions for *Pilosocereus aurisetus* and allies. *Molecular Ecology*, *23*(12), 3044-3063.

Borowiec, M. L. (2019). Spruceup: fast and flexible identification, visualization, and removal of outliers from large multiple sequence alignments. *Journal of Open Source Software*, *4*(42), 1635.

Capella-Gutiérrez, S., Silla-Martínez, J. M., & Gabaldón, T. (2009). trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*, *25*(15), 1972-1973.

Cardillo, M., Weston, P.H., Reynolds, Z.K., Olde, P.M., Mast, A.R., Lemmon, E.M., Lemmon, A.R. and Bromham, L. (2017). The phylogeny and biogeography of Hakea (Proteaceae) reveals the role of biome shifts in a continental plant radiation. *Evolution*, 71(8), 1928-1943.

Carnaval, A.C., Waltari, E., Rodrigues, M.T., Rosauer, D., VanDerWal, J., Damasceno, R., Prates, I., Strangas, M., Spanos, Z., Rivera, D. and Pie, M.R. (2014). Prediction of phylogeographic endemism in an environmentally complex biome. *Proceedings of the Royal Society B: Biological Sciences*, *281*(1792), 20141461

Chen, S., Zhou, Y., Chen, Y., & Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*, *34*(17), i884-i890.

Coe, H. H. G., & Sousa, L. O. F. (2014). The Brazilian "Caatinga": ecology and vegetal biodiversity of a semiarid region. *Dry forests: Ecology, species diversity and sustainable management*, *1*, 81-103.

Copetti, D., Búrquez, A., Bustamante, E., Charboneau, J.L., Childs, K.L., Eguiarte, L.E., Lee, S., Liu, T.L., McMahon, M.M., Whiteman, N.K., Wing, R.A., Wojciechowski, M. F., & Sanderson, M. J. (2017). Extensive gene tree discordance and hemiplasy shaped the genomes of North American columnar cacti. *Proceedings of the National Academy of Sciences*, *114*(45), 12003-12008.

Corbett, E. C., Bravo, G. A., Schunck, F., Naka, L. N., Silveira, L. F., & Edwards, S. V. (2020). Evidence for the Pleistocene Arc Hypothesis from genome-wide SNPs in a Neotropical dry forest specialist, the Rufous-fronted Thornbird (Furnariidae: *Phacellodomus rufifrons*). *Molecular ecology*, *29*(22), 4457-4472.

Cornejo-Romero, A., Medina-Sánchez, J., Hernández-Hernández, T., Rendón-Aguilar, B., Valverde, P.L., Zavala-Hurtado, A., Rivas-Arancibia, S.P., Pérez-Hernández, M.A., López-Ortega, G., Jiménez-Sierra, C. and Vargas-Mendoza, C.F. (2014). Quaternary origin and genetic divergence of the endemic cactus *Mammillaria pectinifera* in a changing landscape in the Tehuacán Valley, Mexico. *Genetics and Molecular Research*, *13*(1), 73-88.

de Aguiar-Campos, N., Maia, V. A., da Silva, W. B., de Souza, C. R., & dos Santos, R. M. (2020). Can fine-scale habitats of limestone outcrops be considered litho-refugia for dry forest tree lineages?. *Biodiversity and Conservation*, *29*(3), 1009-1026.

de Queiroz, L. P., Cardoso, D., Fernandes, M. F., & Moro, M. F. (2017). Diversity and evolution of flowering plants of the Caatinga domain. *Caatinga: the largest tropical dry forest region in South America*, 23-63.

Dinerstein, E., Olson, D., Joshi, A., Vynne, C., Burgess, N. D., Wikramanayake, E., ... & Saleem, M. (2017). An ecoregion-based approach to protecting half the terrestrial realm. *BioScience*, *67*(6), 534-545.

Dolan, A. M., Haywood, A. M., Hunter, S. J., Tindall, J. C., Dowsett, H. J., Hill, D. J., & Pickering, S. J. (2015). Modelling the enigmatic late Pliocene glacial event—Marine Isotope Stage M2. *Global and Planetary Change*, *128*, 47-60.

Donoghue, M. J., & Edwards, E. J. (2014). Biome shifts and niche evolution in plants. *Annual Review of Ecology, Evolution, and Systematics*, *45*(1), 547-572.

Evans, M. E., Smith, S. A., Flynn, R. S., & Donoghue, M. J. (2009). Climate, niche evolution, and diversification of the “bird-cage” evening primroses (Oenothera, sections Anogra and Kleinia). *The American Naturalist*, *173*(2), 225-240.

Fernandes, M. F., Cardoso, D., Pennington, R. T., & de Queiroz, L. P. (2022). The origins and historical assembly of the Brazilian Caatinga seasonally dry tropical forests. *Frontiers in Ecology and Evolution*, *10*, 723286.

Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International journal of climatology*, *37*(12), 4302-4315.

Fine, P. V., & Lohmann, L. G. (2018). Importance of dispersal in the assembly of the Neotropical biota. *Proceedings of the National Academy of Sciences*, *115*(23), 5829-5831.

Flora do Brasil. (2020). Flora do Brasil. Jardim Botânico do Rio de Janeiro. Available online: http://floradobrasil.jbrj.gov.br/ (Accessed, June 20th, 2024)

Fordham, D. A., Saltré, F., Haythorne, S., Wigley, T. M., Otto-Bliesner, B. L., Chan, K. C., & Brook, B. W. (2017). PaleoView: a tool for generating continuous climate projections spanning the last 21 000 years at regional and global scales. *Ecography*, *40*(11), 1348-1358.

Franco, F. F., Silva, G. A. R., Moraes, E. M., Taylor, N., Zappi, D. C., Jojima, C. L., & Machado, M. C. (2017). Plio-Pleistocene diversification of *Cereus* (Cactaceae, Cereeae) and closely allied genera. *Botanical Journal of the Linnean Society*, *183*(2), 199-210.

Guillory, W. X., & Brown, J. L. (2021). A new method for integrating ecological niche modeling with phylogenetics to estimate ancestral distributions. *Systematic Biology*, *70*(5), 1033-1045.

Herbert, T. D., Lawrence, K. T., Tzanova, A., Peterson, L. C., Caballero-Gill, R., & Kelly, C. S. (2016). Late Miocene global cooling and the rise of modern ecosystems. *Nature Geoscience*, *9*(11), 843-847.

Hernández-Hernández, T., Brown, J. W., Schlumpberger, B. O., Eguiarte, L. E., & Magallón, S. (2014). Beyond aridification: multiple explanations for the elevated diversification of cacti in the New World Succulent Biome. *New phytologist*, *202*(4), 1382-1397.

Hulshof, C. M., & Spasojevic, M. J. (2020). The edaphic control of plant diversity. *Global Ecology and Biogeography*, 29(10), 1634-1650.

Hunt, D. R., Taylor, N. P., and Charles, G. (2006). *New cactus lexicon*. Royal Botanic Gardens, Kew.

Hurbath, F., Stubbs, R. L., Cordeiro, I., & Cellinese, N. (2021). Biogeography of succulent spurges from Brazilian Seasonally Dry Tropical Forest (SDTF). *Taxon*, *70*(1), 153-169.

Johnson, M.G., Gardner, E.M., Liu, Y., Medina, R., Goffinet, B., Shaw, A.J., Zerega, N.J. and Wickett, N.J. (2016). HybPiper: Extracting coding sequence and introns for phylogenetics from high-throughput sequencing reads using target enrichment. *Applications in plant sciences*, *4*(7), 1600016.

Juárez-Miranda, A. I., Cornejo-Romero, A., & Vargas-Mendoza, C. F. (2021). Population expansion and genetic structure in *Cephalocereus nizandensis* (Cactaceae), a microendemic cactus of rocky outcrops of the Tehuantepec basin, Mexico. *Plant Ecology and Evolution*, *154*(2), 217-230.

Karger, D. N., Nobis, M. P., Normand, S., Graham, C. H., & Zimmermann, N. E. (2021): CHELSA-TraCE21k v1. 0. Downscaled transient temperature and precipitation data since the last glacial maximum. *Climate of the Past Discussions*, 1-27.

Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular biology and evolution*, *30*(4), 772-780.

Koenen, E.J.M., De Vos, J.M., Atchison, G.W., Simon, M.F., Schrire, B.D., De Souza, E.R., De Queiroz, L.P. and Hughes, C.E. (2013). Exploring the tempo of species diversification in legumes. *South African Journal of Botany*, *89*, 19-30.

Korotkova, N., Aquino, D., Arias, S., Eggli, U., Franck, A., Gómez-Hinostrosa, C., ... & Berendsohn, W. G. (2021). Cactaceae at Caryophyllales. org–a dynamic online species-level taxonomic backbone for the family. *Willdenowia*, *51*(2), 251-270.

Lagomarsino, L. P., Frankel, L., Uribe-Convers, S., Antonelli, A., & Muchhala, N. (2022). Increased resolution in the face of conflict: phylogenomics of the Neotropical

bellflowers (Campanulaceae: Lobelioideae), a rapid plant radiation. *Annals of Botany*, *129*(6), 723-736.

Landis, M., Edwards, E. J., & Donoghue, M. J. (2021). Modeling phylogenetic biome shifts on a planet with a past. *Systematic biology*, *70*(1), 86-107.

Lavor, P., Calvente, A., Versieux, L. M., & Sanmartin, I. (2019). Bayesian spatio-temporal reconstruction reveals rapid diversification and Pleistocene range expansion in the widespread columnar cactus *Pilosocereus*. *Journal of Biogeography*, *46*(1), 238-250.

Ledo, R. M. D., & Colli, G. R. (2017). The historical connections between the Amazon and the Atlantic Forest revisited. *Journal of biogeography*, *44*(11), 2551-2563.

Mabesoone, J. M. (1994). *Sedimentary Basins of Northeast Brazil*. Recife: Editora Universitária UFPE.

Maia, R. P., Bezerra, F. H. R., & Claudino-Sales, V. (2010). Geomorfologia do Nordeste: concepções clássicas e atuais acerca das superfícies de aplainamento nordestinas. *Revista de Geografia*, *27*, 6-19.

Maswanganye, K. A., Cunningham, M. J., Bennett, N. C., Chimimba, C. T., & Bloomer, P. (2017). Life on the rocks: Multilocus phylogeography of rock hyrax (Procavia capensis) from southern Africa. *Molecular Phylogenetics and Evolution*, *114*, 49-62.

Matzke, N. J. (2013). *Probabilistic historical biogeography: new models for founder-event speciation, imperfect detection, and fossils allow improved accuracy and model-testing*. University of California, Berkeley.

Menezes, M. O., Zappi, D. C., Moraes, E. M., Franco, F. F., Taylor, N. P., Costa, I. R., & Loiola, M. I. (2016). Pleistocene radiation of coastal species of *Pilosocereus* (Cactaceae) in eastern Brazil. *Journal of Arid Environments*, *135*, 22-32.

Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., Von Haeseler, A., & Lanfear, R. (2020). IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Molecular biology and evolution*, *37*(5), 1530-1534.

Moreira, A., Moutinho, P., Nepstad, D., & Oliveira, A. (1999). Evolution of the Brazilian phytogeography classification systems: implications for biodiversity conservation. *Ciência e Cultura*, *51*(5/6), 331-348.

Nieto-Blázquez, M. E., Antonelli, A., & Roncal, J. (2017). Historical Biogeography of endemic seed plant genera in the Caribbean: Did GAAR landia play a role?. *Ecology and Evolution*, *7*(23), 10158-10174.

Otto-Bliesner, B. L., Marshall, S. J., Overpeck, J. T., Miller, G. H., Hu, A., & CAPE Last Interglacial Project members. (2006). Simulating Arctic climate warmth and icefield retreat in the last interglaciation. *science*, *311*(5768), 1751-1753.

Pennington, R. T., Prado, D. E., and Pendry, C. A. (2000). Neotropical seasonally dry forests and Quaternary vegetation changes. *J. Biogeogr.* 27, 261–273. doi: 10.1046/j.1365-2699.2000.00397.x

Perez, M. F., Bonatelli, I. A. S., Moraes, E. M., & Carstens, B. C. (2016). Model-based analysis supports interglacial refugia over long-dispersal events in the diversification of two South American cactus species. *Heredity*, *116*(6), 550-557.

Pérez-Escobar, O. A., Zizka, A., Bermúdez, M. A., Meseguer, A. S., Condamine, F. L., Hoorn, C., Hooghiemstra, H., Pu, Y., Bogarín, D., Boschman, L. M., Pennington, T. R., Antonelli, A., & Chomicki, G. (2022). The Andes through time: evolution and distribution of Andean floras. *Trends in Plant Science*, *27*(4), 364-378.

POWO (2024). *Plants of the World Online*. Facilitated by the Royal Botanic Gardens, Kew. Published on the Internet; https://powo.science.kew.org/

Prado, D. E., & Gibbs, P. E. (1993). Patterns of species distributions in the dry seasonal forests of South America. *Annals of the Missouri Botanical Garden*, 902-927.

QGIS Development Team, (2023). *QGIS Geographic Information System*. Open Source Geospatial Foundation Project. http://qgis.org

R Core Team (2023). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>.

Rabosky, D. L. (2019). Phylogenies and diversification rates: variance cannot be ignored. *Systematic Biology*, *68*(3), 538-550.

Rajakaruna, N. (2004). The edaphic factor in the origin of plant species. *International Geology Review*, *46*(5), 471-478.

Rajakaruna, N. (2018). Lessons on evolution from the study of edaphic specialization. *The Botanical Review*, *84*, 39-78.

Revell, L. J. (2012). phytools: an R package for phylogenetic comparative biology (and other things). *Methods in ecology and evolution*, (2), 217-223.

Ritz, C. M., Martins, L., Mecklenburg, R., Goremykin, V., & Hellwig, F. H. (2007). The molecular phylogeny of *Rebutia* (Cactaceae) and its allies demonstrates the influence of paleogeography on the evolution of South American mountain cacti. *American journal of Botany*, *94*(8), 1321-1332.

Romeiro-Brito, M., Taylor, N. P., Zappi, D. C., Telhe, M. C., Franco, F. F., & Moraes, E. M. (2023). Unravelling phylogenetic relationships of the tribe Cereeae using target enrichment sequencing. *Annals of Botany*, mcad153.

Romeiro-Brito, M., Telhe, M. C., Amaral, D. T., Franco, F. F., & Moraes, E. M. (2022). A target capture probe set useful for deep-and shallow-level phylogenetic studies in Cactaceae. *Genes*, *13*(4), 707.

Romeiro-Brito, M., Khan, G., Perez, M.F., Zappi, D.C., Taylor, N.P., Olsthoorn, G., Franco, F.F. and Moraes, E.M., (2023a). Revisiting phylogeny, systematics, and biogeography of a Pleistocene radiation. *American Journal of Botany*, *110*(3), e16134.

Rull, V. (2020). Neotropical diversification: historical overview and conceptual insights. *Neotropical diversification: Patterns and processes*, 13-49.

Salgado, A. A. R., Bueno, G. T., Diniz, A. D., & Marent, B. R. (2015). Long-term geomorphological evolution of the Brazilian territory. *Landscapes and landforms of Brazil*, 19-31.

Silva, G. A. R., Antonelli, A., Lendel, A., Moraes, E. D. M., & Manfrin, M. H. (2018). The impact of early Quaternary climate change on the diversification and population dynamics of a South American cactus species. *Journal of Biogeography*, *45*(1), 76-88.

Silva, O. L., Bezerra, F. H., Maia, R. P., & Cazarin, C. L. (2017). Karst landforms revealed at various scales using LiDAR and UAV in semi-arid Brazil: Consideration on karstification processes and methodological constraints. *Geomorphology*, *295*, 611-630.

Simon, M. F., & Pennington, T. (2012). Evidence for adaptation to fire regimes in the tropical savannas of the Brazilian Cerrado. *International Journal of Plant Sciences*, *173*(6), 711-723.

Simon, M. F., Grether, R., de Queiroz, L. P., Skema, C., Pennington, R. T., & Hughes, C. E. (2009). Recent assembly of the Cerrado, a neotropical plant diversity hotspot, by in situ evolution of adaptations to fire. *Proceedings of the National Academy of Sciences*, *106*(48), 20359-20364.

Sobral-Souza, T., Lima-Ribeiro, M. S., & Solferini, V. N. (2015). Biogeography of Neotropical Rainforests: past connections between Amazon and Atlantic Forest detected by ecological niche modeling. *Evolutionary Ecology*, *29*, 643-655

Souza-Neto, A. C., Cianciaruso, M. V., & Collevatti, R. G. (2016). Habitat shifts shaping the diversity of a biodiversity hotspot through time: insights from the phylogenetic structure of Caesalpinioideae in the Brazilian Cerrado. *Journal of Biogeography*, *43*(2), 340-350.

Steinke, V. A. (2021). Proposal for a Geobiodiversity Index Applied to the Morphoclimatic Domain of Cerrado—Brazil. *Geoheritage*, 13(3), 59.

Sun, Y., Shang, L., Zhu, Q. H., Fan, L., & Guo, L. (2022). Twenty years of plant genome sequencing: achievements and challenges. *Trends in Plant Science*, *27*(4), 391-401.

Tamura, K., Battistuzzi, F. U., Billing-Ross, P., Murillo, O., Filipski, A., & Kumar, S. (2012). Estimating divergence times in large molecular phylogenies. *Proceedings of the National Academy of Sciences*, *109*(47), 19333-19338.

Tamura, K., Stecher, G., & Kumar, S. (2021). MEGA11: molecular evolutionary genetics analysis version 11. *Molecular biology and evolution*, *38*(7), 3022-3027.

Taylor, N. P., & Zappi, D. C. (2004). *Cacti of eastern Brazil*. Royal Botanic Gardens, *Kew. UK*.

Taylor, N. P., Albuquerque-Lima, S., Telhe, M., & Zappi, D. C. (2022). Further additions and corrections to Cacti of Eastern Brazil. *Bradleya*, *2022*(40), 61-92.

Taylor, N. P., Zappi, D. C., Romeiro-Brito, M., Telhe, M. C., Franco, F. F., & Moraes, E. M. (2023). A phylogeny of *Cereus* (Cactaceae) and the placement and description of two new species. *Taxon, 72*(6), 1321-1333.

Taylor, N., and Zappi, D. (2018). Additions and corrections to 'Cacti of Eastern Brazil'. *Bradleya*, 2018 (36), 2-21.

Toby Pennington, R., Prado, D. E., & Pendry, C. A. (2000). Neotropical seasonally dry forests and Quaternary vegetation changes. *Journal of Biogeography, 27*(2), 261-273.

Töpel, M., Zizka, A., Calió, M. F., Scharn, R., Silvestro, D., & Antonelli, A. (2017). SpeciesGeoCoder: Fast categorization of species occurrences for analyses of biodiversity, biogeography, ecology, and evolution. *Systematic Biology*, *66*(2), 145-151.

Tripati, A. K., Roberts, C. D., & Eagle, R. A. (2009). Coupling of CO2 and ice sheet stability over major climate transitions of the last 20 million years. *Science*, *326*(5958), 1394-1397.

Vasconcelos, T., O'Meara, B. C., & Beaulieu, J. M. (2022). A flexible method for estimating tip diversification rates across a range of speciation and extinction scenarios. *Evolution*, *76*(7), 1420-1433.

Werneck, F. P. (2011). The diversification of eastern South American open vegetation biomes: historical biogeography and perspectives. *Quaternary Science Reviews*, *30*(13-14), 1630-1648.

Werneck, F. P., Costa, G. C., Colli, G. R., Prado, D. E., & Sites Jr, J. W. (2011). Revisiting the historical distribution of Seasonally Dry Tropical Forests: new insights based on palaeodistribution modelling and palynological evidencegeb. *Global Ecology and Biogeography*, *20*(2), 272-288.

Wiens, J. J. (2011). The niche, biogeography and species interactions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1576), 2336-2350.

Winner, K., Noonan, M. J., Fleming, C. H., Olson, K. A., Mueller, T., Sheldon, D., & Calabrese, J. M. (2018). Statistical inference for home range overlap. *Methods in Ecology and Evolution*, *9*(7), 1679-1691.

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

Zhang, C., Rabiee, M., Sayyari, E., & Mirarab, S. (2018). ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. *BMC bioinformatics*, *19*(6), 15-30.

Zizka, A. (2019). Big data suggest migration and bioregion connectivity as crucial for the evolution of Neotropical biodiversity. *Frontiers of Biogeography*, *11*(2).

Zizka, A., Silvestro, D., Andermann, T., Azevedo, J., Duarte Ritter, C., Edler, D., ... & Antonelli, A. (2019). CoordinateCleaner: Standardized cleaning of occurrence records from biological collection databases. *Methods in Ecology and Evolution*, *10*(5), 744-751.

4.6 Supplementary Material

Table S1. Species classification regarding the presence of fire-adapted traits, used to investigate the association between trait states, biome occurrence, and speciation rates using phylogenetic ANOVA.

Taxon	Fire Adapted (globose-discoid habit)
D. bahiensis	yes
D. boliviensis	yes
D. buenekeri	yes
D. catingicola	yes
D. diersianus	yes
D. fariae-peresii	yes
D. ferricola	yes
D. hartmannii	yes
D. hartmannii_B	yes
D. heptacanthus	yes
D. horstii	yes
D. placentiformis	yes
D. pseudoinsignis	yes
D. boomianus	yes
D. petr-halfari	yes
D. zehntneri	yes

M. azureus	no
M. amethystinus	no
M. bahiensis	no
M. braunii	no
M. brederooianus	no
M. broadwayi	no
M. caroli-linnaei	no
M. concinnus	yes
M. conoideus	no
M. deinacanthus	no
M. ernestii	no
M. estevesii	no
M. ferreophilus	no
M. glaucescens	no
M. harlowii	no
M. inconcinnus	no
M. intortus	no
M. levitestatus	no
M. macracanthos	no
M. mazelianus	no

M. neoviridescens	no
M. oreas_oreas	no
M. pachyacanthus	no
M. paucispinus	yes
M. salvadorensis	no
M. sergipensis	no
M. violaceus_violaceus	no
M. depressus	no
M. zehntneri	no

Table S2. Divergence time estimated (CI 95%) for Cactoideae subfamily, with a sampling focus on Cereinae subtribe, using the *RelTime* method on *MEGA11*. Record number from *Laboratório de Diversidade Genética e Evolução* (LAGEVol) for each specimen is subsequent to species name.

Taxon	**CI_Lower (Ma)**	**CI_Upper (Ma)**
Arrojadoa bahiensis S174A1	1,008	6,665
Arrojadoa eriocaulis S174A2	0,786	6,594
Arrojadoa marylanae S170A2	1,248	8,453
Arrojadoa multiflora S174A3	0,989	7,278
Arrojadoa penicillata S174A4	0,839	6,858
Arrojadoa rhodantha S142A6	0,839	6,858
Arrojadoa olsthoorniana S165A40	0,786	6,594
Arthrocereus melanurus magnus S182A2	1,019	4,975
Arthrocereus melanurus odorus S182A8	0,525	3,245
Arthrocereus glaziovii S162A2	1,019	4,975
Arthrocereus rondonianus S170A17	0,525	3,245
Arthrocereus spinosissimus S170A20	2,347	8,945
Astrophytum myriostigma S170A23	7,413	15,291
Brasilicereus estevesii S162A3	2,351	10,053
Brasilicereus markgrafii S168A2	1,243	6,453
Brasiliecereus phaeacanthus S165A1	1,243	6,453
Browningia hertlingiana S152A7	4,141	11,293
Cereus gerardi S162A24	0,769	4,981
Cereus lepidotus S184A1	0,890	5,118
Cereus mirabella S162A26	1,033	5,420
Cereus mortensenii S184A3	1,292	5,920

Cereus spegazzini S180A14	0,916	4,469
Cereus trigonodendron S184A4	1,510	7,056
Cereus vargasianus S184A5	0,916	4,469
Cereus bicolor S103B3	0,874	5,375
Cereus fernambucensis S107V3	0,326	2,713
Cereus hexagonus S162A28	0,769	4,981
Cereus hildmannianus S103C3	1,011	5,526
Cereus insularis S115C51	0,326	2,713
Cereus jamacaru S44V3	0,942	5,469
Cereus kroenleinii S149V10	0,821	4,095
Cereus pierre-braunianus S145F1	0,942	5,469
Cereus repandus S162VA22	1,292	5,920
Cereus saddianus S99A6	0,821	4,095
Cereus ingens S169A2	1,433	6,864
Cereus stenogonus S149V11	0,874	5,375
Cipocereus bradei S137A5	0,942	5,469
Cipocereus crassisepalus S174A5	1,513	6,856
Cipocereus laniflorus S174A6	1,249	6,103
Cipocereus minensis S153A1	0,904	5,469
Cipocereus pleurocarpus S174A7	1,033	5,647
Cipocereus pusilliflorus L40A4	1,858	9,117
Cleistocactus winteri colademono S170A25	1,595	7,042
Coleocephalocereus aureus S183A2	0,603	4,056
Coleocephalocereus goebelianus S162A4	1,598	8,229

Coleocephalocereus braunii S174A8	0,692	4,475
Coleocephalocereus fluminensis S79C1	0,692	4,475
Coleocephalocereus purpureus S174A10	0,603	4,056
Discocactus bahiensis L25A9	0,831	5,280
Discocactus hartmanni L26A47	0,531	4,224
Discocactus horstii L26A15 L25A15	0,480	3,921
Discocactus placentiformis L26A10	0,669	4,705
Discocactus zehntneri L26A12	1,002	5,349
Discocactus boomianus L26A13	1,225	6,019
Discocactus boliviensis L26A2	0,299	3,130
Discocactus catingicola L26A3	0,511	4,139
Discocactus diersianus L25A1	0,602	4,548
Discocactus fariae-peresii L25A2	0,564	4,211
Discocactus ferricola S162A6	0,299	3,130
Discocactus hartmannii L26A5	0,663	4,667
Discocactus heptacanthus S162A32	0,511	4,139
Discocactus pseudoinsignis L26A11	0,480	3,921
Discocactus petr-halfari S165A19	1,002	5,349
Discocactus buenekeri L26A17	1,226	6,019
Echinocereus pentalophus S170A29	5,621	13,785
Echinopsis calochlora S162A7	2,400	8,636
Espostoopsis dybowskii S152A13	2,150	9,638
Estevesia alex-bragae S162A35	1,033	5,420
Facheiroa squamosa S180A4	1,490	8,414

Facheiroa ulei S168A7	1,490	8,414
Ferocactus herrarea S170A34	3,876	13,059
Ferocactus glaucescens S170A33	3,876	13,059
Gymnocalycium denudatum S152A3	1,076	5,190
Gymnocalycium horstii S170A35	1,076	5,190
Hagaaecereus decumbens S152A2	1,264	6,307
Harrisia martinii S152A16	0,374	2,673
Harrisia adscendens S142A7	1,379	5,552
Harrisia balansae S162A12	0,374	2,673
Leocereusbahiensis S168A8	2,060	10,053
Matucana intertexta S152A1	1,264	6,307
Melocactus azureus L26A52	0,483	3,644
Melocactus braunii L26A20	0,651	3,962
Melocactus amethystinus L26A21	0,655	4,189
Melocactus caroli-linnaei L42A8	0,080	1,693
Melocactus concinnus L26A25	0,746	4,114
Melocactus deinacanthus S174A19	0,860	4,795
Melocactus glaucescens L8G9	0,989	5,647
Melocactus intortus L42A9	0,292	2,626
Melocactus salvadorensis S165A3	0,534	3,686
Melocactus violaceus L26A55	0,989	5,647
Melocactus bahiensis L26A19	0,609	4,045
Melocactus brederooianus L25A12	0,351	3,034
Melocactus broadwayi L42A5	0,202	2,338

Melocactus conoideus S174A18	0,609	4,045
Melocactus ernestii L26A27	0,655	4,189
Melocactus estevesii L42A2	0,202	2,338
Melocactus ferreophilus L26A30	0,620	4,092
Melocactus harlowii L42A10	0,395	3,018
Melocactus inconcinnus L26A56	0,716	4,189
Melocactus levitestatus L26A33	0,696	4,318
Melocactus macracanthos S162A15	0,292	2,626
Melocactus mazelianus L42A6	0,080	1,693
Melocactus oreas L26A53	0,351	3,034
Melocactus neoviridescens L26A36	0,620	4,092
Melocactus pachyacanthus L26A50	0,483	3,644
Melocactus paucispinus L26A37	0,860	4,795
Melocactus sergipensis L25A11	0,873	4,710
Melocactus violaceus L28O1	0,989	5,647
Melocactus zehntneri L25A10	0,534	3,686
Micranthocereus estevesii S32V1	1,602	8,201
Micranthocereus polyanthus S162A20	1,141	6,383
Micranthocereus auriazureus S168A10	1,710	8,916
Micranthocereus violaciflorus S162A21	1,605	8,766
Micranthocereus albicephalus S165A37	1,797	8,916
Micranthocereus densiflorus S174A11	0,728	4,455
Micranthocereus dolichospermaticus S174A12	1,602	8,201
Micranthocereus flaviflorus S165A26	0,728	4,455

Micranthocereus purpureus S165A9	1,516	7,684
Micranthocereus streckeri S174A13	1,141	6,383
Oreocereus celsianus S152A5	1,598	7,126
Pachycereus marginatus S170A47	5,621	13,785
Parodia leninghausii S119A5	5,533	14,212
Parodia magnifica S170A49	5,533	14,212
Pilocoereus pachycladus schoebelli S45B6	0,841	6,741
Pilosocereus aurisetus S97A2	0,683	5,546
Pilosocereus brasiliensis L30A8	0,814	6,137
Pilosocereus catimbauensis S180A13	0,746	5,816
Pilosocereus densiareolatus S43A6	0,846	6,575
Pilosocereus glaucochrous S165A27	0,652	5,495
Pilosocereus machrisi L40A1	0,663	5,375
Pilosocereus magnificus S37A5	0,793	6,503
Pilosocereus mollispinus S162A39	0,520	4,935
Pilosocereus multicostatus S39A3	0,599	5,362
Pilosocereus parvus S138A21	0,743	5,718
Pilosocereus pentaedrophorous S113V1	0,652	5,495
Pilosocereus tuberculatus S131A1	1,317	6,471
Pilosocereus vilaboensis S149V3	0,384	3,805
Pilosocereus albisummus S141B9	0,846	6,575
Pilosocereus arrabidae S79H2	0,814	6,137
Pilosocereus aureispinus S21V2	1,405	8,593
Pilosocereus aurisetus aurilanatus S7A9	0,626	5,484

Pilosocereus azulensis S77A9	0,857	6,532
Pilosocereus catingicola salvadorensis L28E12	0,450	3,927
Pilosocereus chrysostele S52V1C	0,964	6,934
Pilosocereus diersianus L18B1	0,569	5,043
Pilosocereus flavipulvinatus S120A3	0,450	3,927
Pilosocereus flexibilispinus L16A6	0,599	5,362
Pilosocereus floccosus S38A1	0,793	6,503
Pilosocereus frewenii S120V11	0,277	2,544
Pilosocereus fulvilanatus S77V36	0,626	5,484
Pilosocereus gounellei S142A3	0,277	2,544
Pilosocereus jauruensis S23V1	0,384	3,805
Pilosocereus lanuginosus S148A6	0,548	4,491
Pilosocereus leucocephalus S148A1	0,316	3,363
Pilosocereus piahyensis S151V1	0,787	6,133
Pilosocereus polygonus S148A4	0,472	4,234
Pilosocereus purpusii S148A2	0,316	3,363
Pilosocereus pussilibaccatus L15A5	0,743	5,718
Pilosocereus rizzoanus L21A	0,520	4,935
Pilosocereus royenii S148A8	0,472	4,234
Pilosocereus splendidus S139AV1	0,884	6,741
Pilosocereus ulei L39A10	1,208	8,006
Praecereus saxicola S149V16	0,583	3,214
Praecereus euchlorus S150F3	0,583	3,214
Samaipaticereus corroanus S152A14	1,596	7,042

Selenicereus anthonyanus S170A54	3,752	11,534
Selenicereus chrysocardium S170A55	3,752	11,534
Stephanocereus leucostele S165A5	1,391	8,626
Stephanocereus luetzelburgii S180A19	1,009	6,666
Uebelmannia buiningii S124A2	0,730	3,304
Uebelmannia gummifera S121A5	0,730	3,304
Uebelmannia pectinifera S62A11	1,985	6,112
Gymnocalycium anisitsii S162V10	1,696	6,241

Table S3. Diversification rate within the subtribe Cereinae estimated with the *hisse* R package using *MISSE* method.

Taxon	**speciation rate**
Arrojadoa bahiensis S174A1	0.199
Arrojadoa bahiensis S174A1	0.199
Arrojadoa eriocaulis S174A2	0.198
Arrojadoa marylanae S170A2	0.197
Arrojadoa multiflora S174A3	0.197
Arrojadoa penicillata S174A4	0.198
Arrojadoa penicillata S174A4	0.198
Arrojadoa rhodantha S142A6	0.198
Arrojadoa rhodantha S142A6	0.198
Arrojadoa olsthoorniana S165A40	0.198
Brasilicereus estevesii S162A3	0.198
Brasilicereus markgrafii S168A2	0.200
Brasilicereus phaeacanthus S165A1	0.200
Cereus lepidotus S184A1	0.200
Cereus mirabella S162A26	0.200
Cereus mortensenii S184A3	0.200
Cereus gerardi S162A24	0.202
Cereus spegazzinii S180A14	0.200
Cereus trigonodendron S184A4	0.197
Cereus vargasianus S184A5	0.200
Cereus bicolor S103B3	0.199
Cereus fernambucensis S107V3	0.203
Cereus hexagonus S162A28	0.202
Cereus hildmannianus S103C3	0.198
Cereus insularis S115C51	0.203
Cereus jamacaru S44V3	0.198
Cereus kroenleinii S149V10	0.200
Cereus pierre-braunianus S145F1	0.198
Cereus repandus S162VA22	0.200
Cereus saddianus S99A6	0.200
Cereussp ingens S169A2	0.197

Cereus stenogonus S149V11	0.199
Cipocereus bradei S137A5	0.199
Cipocereus crassisepalus S174A5	0.198
Cipocereus laniflorus S174A6	0.198
Cipocereus minensis S153A1	0.199
Cipocereus pleurocarpus S174A7	0.199
Cipocereus pusilliflorus L40A4	0.198
Coleocephalocereus aureus S183A2	0.202
Coleocephalocereus goebelianus S162A4	0.198
Coleocephalocereus goebelianus S162A4	0.198
Coleocephalocereus braunii S174A8	0.202
Coleocephalocereus fluminensis S79C1	0.202
Coleocephalocereus purpureus S174A10	0.202
Discocactus bahiensis L25A9	0.200
Discocactus hartmanni L26A47	0.235
Discocactus horstii L44A1	0.257
Discocactus pseudoinsignis L26A10	0.209
Discocactus zehntneri L26A12	0.198
Discocactus boomianus L26A13	0.197
Discocactus boliviensis L26A2	0.430
Discocactus catingicola L26A3	0.242
Discocactus diersianus L25A1	0.217
Discocactus fariae-peresii L25A2	0.229
Discocactus ferricola S162A6	0.430
Discocactus hartmannii L26A5	0.210
Discocactus heptacanthus S162A32	0.242
Discocactus pseudoinsignis L26A11	0.257
Discocactus petr-halfari S165A19	0.198
Discocactus buenekeri L26A17	0.197
Espostoopsis dybowskii S152A13	0.198
Estevesia alex-bragae S162A35	0.200
Facheiroa squamosa S180A4	0.199
Facheiroa ulei S168A7	0.199
Leocereus bahiensis S168A8	0.198
Melocactus azureus L26A52	0.258

Melocactus braunii L26A20	0.221
Melocactus amethystinus L26A21	0.215
Melocactus caroli-linnaei L42A8	5.368
Melocactus concinnus L26A25	0.210
Melocactus deinacanthus S174A19	0.201
Melocactus glaucescens L8G9	0.198
Melocactus intortus L42A9	0.967
Melocactus salvadorensis S165A3	0.243
Melocactus violaceus L26A55	0.197
Melocactus bahiensis L26A19	0.224
Melocactus brederooianus L25A12	0.360
Melocactus broadwayi L42A5	1.577
Melocactus conoideus S174A18	0.224
Melocactus ernestii L26A27	0.215
Melocactus estevesii L42A2	1.577
Melocactus ferreophilus L26A30	0.220
Melocactus harlowii L42A10	0.410
Melocactus inconcinnus L26A56	0.212
Melocactus levitestatus L26A33	0.211
Melocactus macracanthos S162A15	0.760
Melocactus mazelianus L42A6	5.368
Melocactus oreas L26A53	0.360
Melocactus neoviridescen L26A36	0.220
Melocactus pachyacanthus L26A50	0.258
Melocactus paucispinus L26A37	0.201
Melocactus sergipensis L25A11	0.201
Melocactus depressus L28O1	0.198
Melocactus zehntneri L25A10	0.243
Micanthocereus estevesii S32V1	0.198
Micranthocereus polyanthus S162A20	0.198
Micranthocereus auriazureus S168A10	0.198
Micranthocereus violaciflorus S162A21	0.197
Micranthocereus albicephalus S165A37	0.198
Micranthocereus densiflorus S174A11	0.202
Micranthocereus dolichospermaticus S174A12	0.198

Micranthocereus flaviflorus S165A26	0.202
Micranthocereus purpureus S165A9	0.198
Micranthocereus streckeri S174A13	0.198
Pilocoereus pachycladus schoebelli S45B6	0.198
Pilosocereus aurisetus S97A2	0.203
Pilosocereus brasiliensis L30A8	0.199
Pilosocereus catimbauensis S180A13	0.200
Pilosocereus densiareolatus_S43A6	0.198
Pilosocereus glaucochrous S165A27	0.204
Pilosocereus machrisi L40A1	0.205
Pilosocereus magnificus S37A5	0.198
Pilosocereus mollispinus S162A39	0.221
Pilosocereus multicostatus S39A3	0.205
Pilosocereus parvus S138A21	0.200
Pilosocereus pentaedrophorous S113V1	0.204
Pilosocereus pentaedrophorous S113V1	0.204
Pilosocereus tuberculatus S131A1	0.200
Pilosocereus vilaboensis S149V3	0.269
Pilosocereus albisummus S141B9	0.198
Pilosocereus arrabidae S79H2	0.199
Pilosocereus aureispinus S21V2	0.197
Pilosocereus aurisetus aurilanatus S7A9	0.206
Pilosocereus azulensis S77A9	0.198
Pilosocereus catingicola salvadorensis L28E12	0.221
Pilosocereus chrysostele S52V1C	0.197
Pilosocereus diersianus L18B1	0.214
Pilosocereus flavipulvinatus S120A3	0.221
Pilosocereus flexibilispinus L16A6	0.205
Pilosocereus floccosus S38A1	0.198
Pilosocereus frewenii S120V11	0.206
Pilosocereus fulvilanatus S77V36	0.206
Pilosocereus gounellei S142A3	0.206
Pilosocereus jauruensis S23V1	0.269
Pilosocereus lanuginosus S148A6	0.226
Pilosocereus leucocephalus S148A1	0.395

Pilosocereus piahyensis S151V1	0.199
Pilosocereus polygonus S148A4	0.250
Pilosocereus purpusii S148A2	0.395
Pilosocereus pussilibaccatus L15A5	0.200
Pilosocereus rizzoanus L21A	0.221
Pilosocereus royenii S148A8	0.250
Pilosocereus splendidus S139AV1	0.198
Pilosocereus ulei L39A10	0.197
Praecereus saxicola S149V16	0.201
Praecereus euchlorus S150F3	0.201
Stephanocereus leucostele S165A5	0.197
Stephanocereus luetzelburgii S180A19	0.199
Stephanocereus luetzelburgii S180A19	0.199

Table S4. Number of range shifts (from root to tip in the phylogenetic tree) estimated for each lineage using *BioGeoBEARS*, and the number of biomes shifts (range shifts corresponding to phylogenetic niche shifts estimated with *machuruku* package.

Taxon	**Range shifts**	**Biome shifts**
D. zehntneri	0	0
D. petr-halfari	0	0
D. buenekeri	0	0
D. bahiensis	0	0
D. placentiformis	1	0
D. fariae-peresii	1	0
D. pseudoinsignis	1	0
D. horstii	1	0
D. hartmannii	1	0
D. diersianus	1	0
D. heptacanthus	1	0
D. catingicola	2	0
D. hartmannii (PY)	1	0
D. boliviensis	2	1
D. ferricola	1	0
D. boomianus	0	0
M. glaucescens	0	0
M. depressus	0	0
M. violaceus	0	0
M. harlowii	0	0
M. macracanthos	1	1

M. caroli-linnaei	0	0
M. mazelianus	1	1
M. broadwayi	0	0
M. estevesii	1	1
M. intortus	0	0
M. deinacanthus	0	0
M. paucispinus	1	1
M. ernestii	1	0
M. amethystinus	1	1
M. inconcinnus	0	0
M. oreas_oreas	0	0
M. brederooianus	0	0
M. braunii	0	0
M. conoideus	1	0
M. bahiensis	2	0
M. concinnus	1	0
M. zehntneri	1	0
M. salvadorensis	0	0
M. sergipensis	0	0
M. neoviridescens	0	0
M. ferreophilus	0	0
M. levitestatus	1	1
M. pachyacanthus	0	0
M. azureus	0	0
Total n. of shifts:	25	7

Table S5. Bhattacharyya-coefficient estimated for species pairs that experienced range shift along their climatic response curves. Bhattacharyya-coefficient lower than 0.3 is highlighted in bold.

Lineages Ovelap	**bio_1**	**bio_10**	**bio_15**	**bio_19**	**bio_4**	**bio_8**
Node43 X *M. levitestatus*	0,6741	0,6482	**0,0000**	**0,0000**	0,9423	0,6128
Node32 X M. ernestii	0,7794	0,7696	0,5936	0,7394	0,7135	0,8012
Node32 X *M. amethystinus*	0,8871	0,9302	0,4698	**0,0000**	0,3604	0,9195
Node23 X *M. macracanthos*	**0,0000**	**0,0000**	0,7996	**0,0000**	0,7949	**0,0002**
Node36 X Node33	0,3300	0,9088	0,9761	0,5086	0,9931	0,9328
Node7 X Node6	0,9703	0,9920	0,7834	0,9910	0,9178	0,8648
Node38 X *M. concinnus*	0,8374	0,8442	0,7399	0,7084	0,9493	0,8648
Node39-M. zehntneri X Node39-M. salvadorensis	0,6225	0,6907	0,8114	0,9539	0,8007	0,8379
Node36-M. bahiensis X Node36-M. conoideus	0,4759	0,6286	0,6073	0,4307	0,6408	0,6548
Node29-M. paucispinus X Node29-M. deinacanthus	**0,0141**	**0,0178**	**0,1347**	**0,2093**	**0,0000**	**0,2093**
Node27-M. broadwayi X Node27-M. estevesii	0,5214	0,4637	**0,0000**	**0,0000**	**0,0000**	0,5504
Node26-M. caroli-linnaei X Node26-M. mazelianus	0,7328	0,7487	**0,1586**	**0,0000**	0,3700	0,8307
Node16-D_boliviensis X Node16-D_ferricola	0,7265	0,7386	0,5518	**0,2955**	0,3858	0,7986
Node14-D_heptacanthus X Node14-D_catingicola	0,9440	0,9383	0,8103	0,7084	0,9184	0,8199

Figure S1. Diversification rates estimated for Cereinae subtribe using the*MiSSE* method.

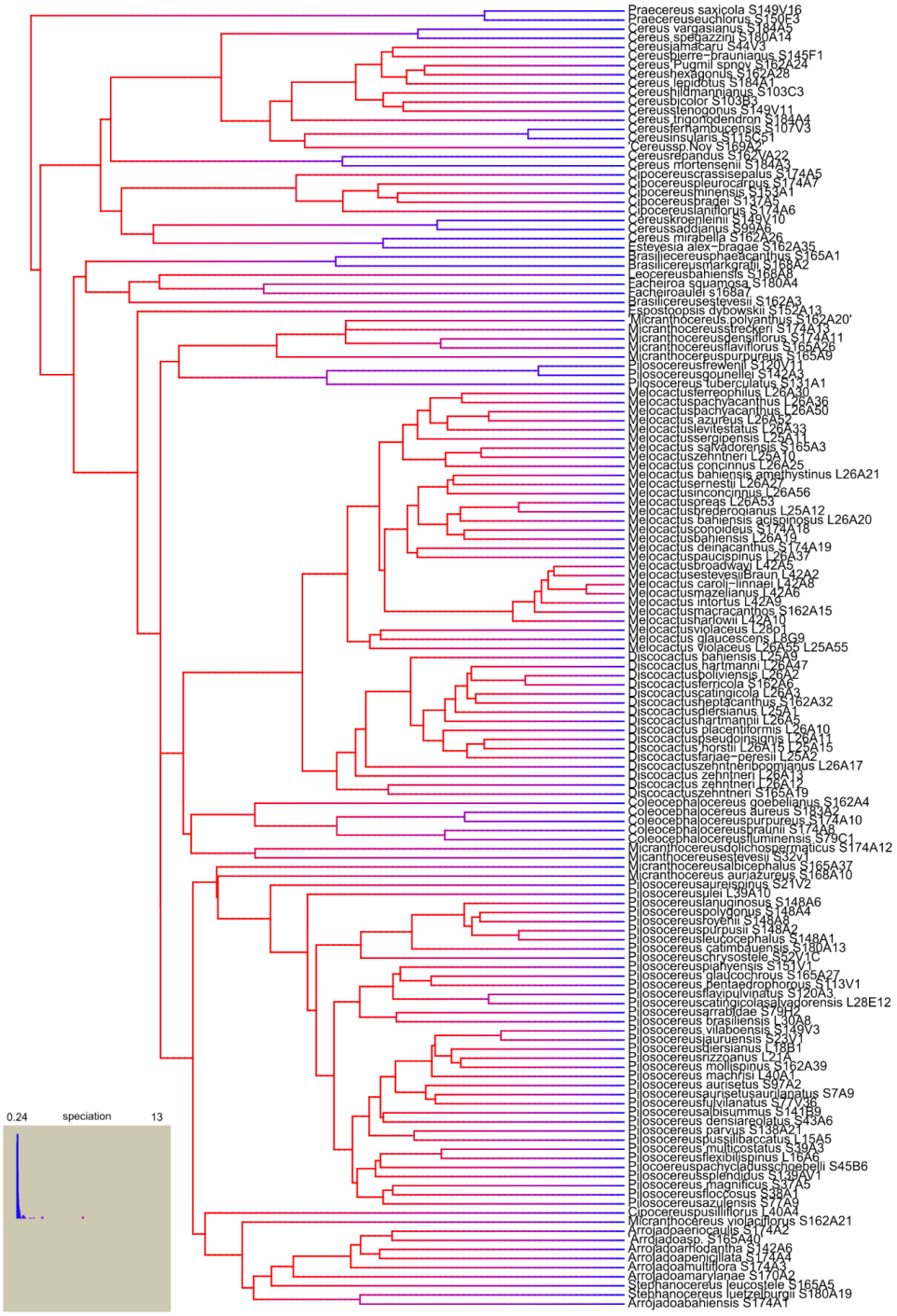


Figure S2. Response curves estimated along the phylogeny using the *machuruku* R package. A) Phylogeny used as input in the *machuruku* R package, showing node numbers and tip names. B) Response curves estimated for nodes 6 and 7 (highlighted in green), which correspond to range shifts identified on the same nodes by BioGeoBEARS.

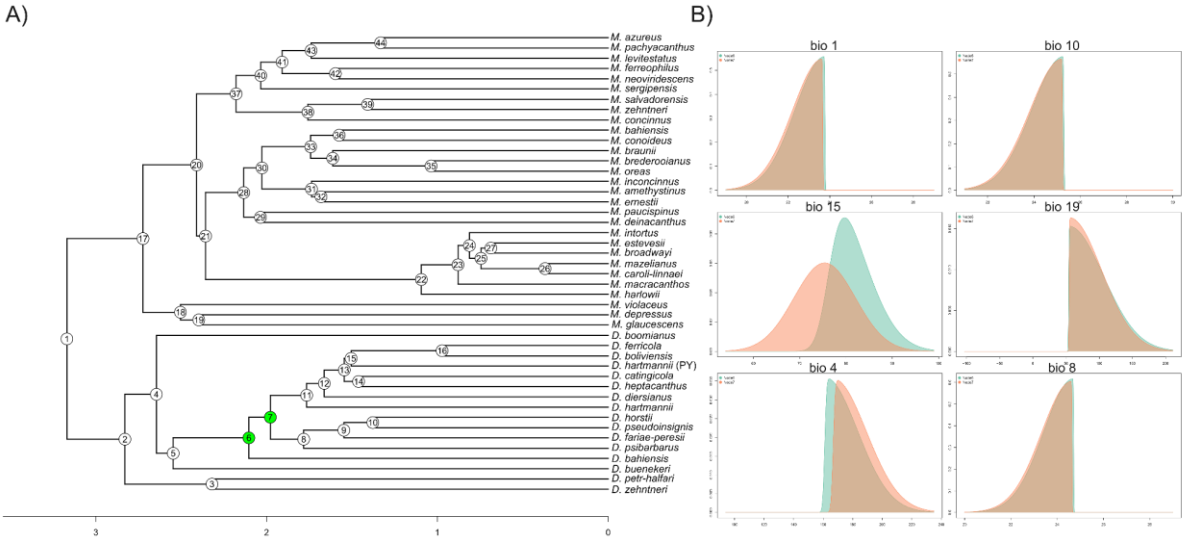


Figure S3. Response curves estimated along the phylogeny using *machuruku* R package. A) Phylogeny used as input on *machuruku* R package, showing node numbers and tip names. B) Response curves estimated for nodes 33 and 36 (highlighted in green), which corresponded to range shifts identified by *BioGeoBEARS*.

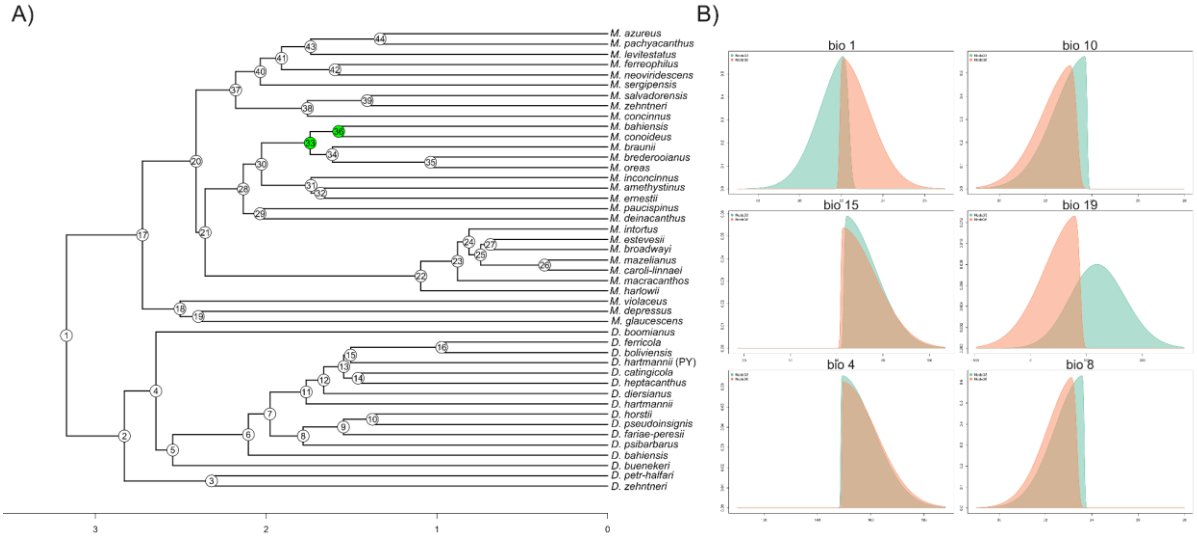


Figure S4. Response curves estimated along the phylogeny using the *machuruku* R package. A) Phylogeny used as input in the *machuruku* R package, showing node numbers and tip names. B) Response curves estimated for node 14 and their derived lineages (highlighted in green), which correspond to range shift identified on the same node by *BioGeoBEARS*.

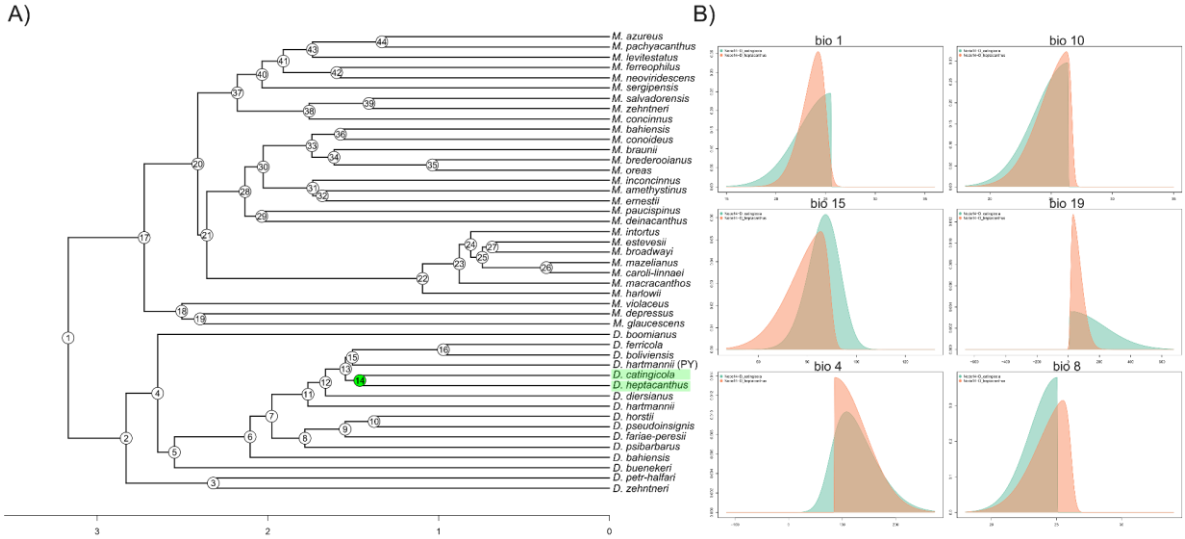


Figure S5. Response curves estimated along the phylogeny using *machuruku* R package. A) Phylogeny used as input in the *machuruku* R package, showing node numbers and tip names. B) Response curves estimated for node 16 and their derived lineages (highlighted in green), which correspond to a range shift identified in the same node by BioGeoBEARS.

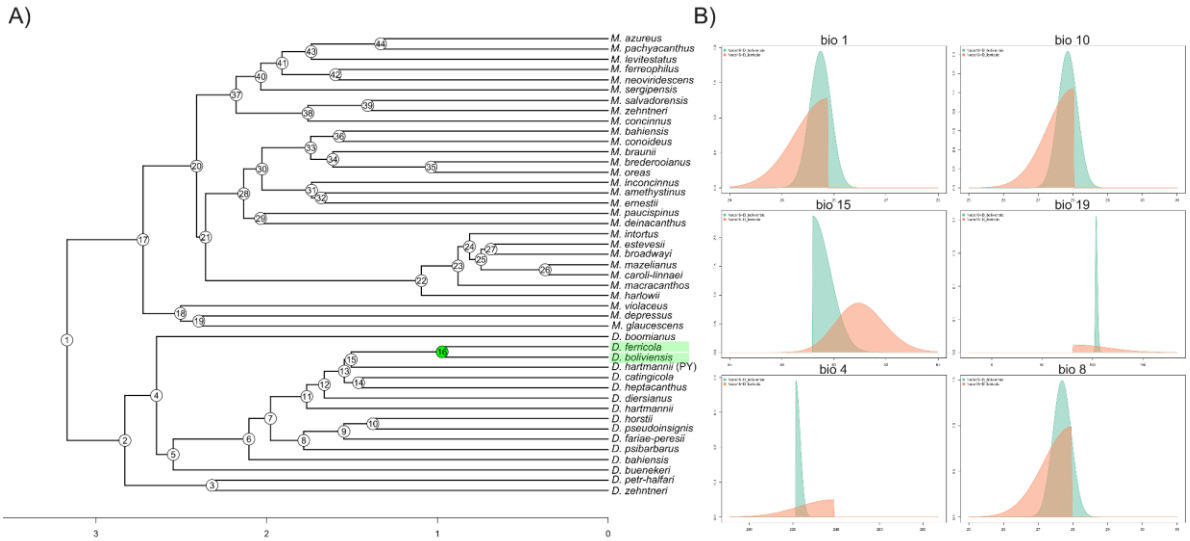


Figure S6. Response curves estimated along the phylogeny using *machuruku* R package. A) Phylogeny used as input in the *machuruku* R package, showing node numbers and tip names. B) Response curves estimated for node 23 and their derived lineages (highlighted in green), which correspond to a range shift identified in the same node by *BioGeoBEARS*.

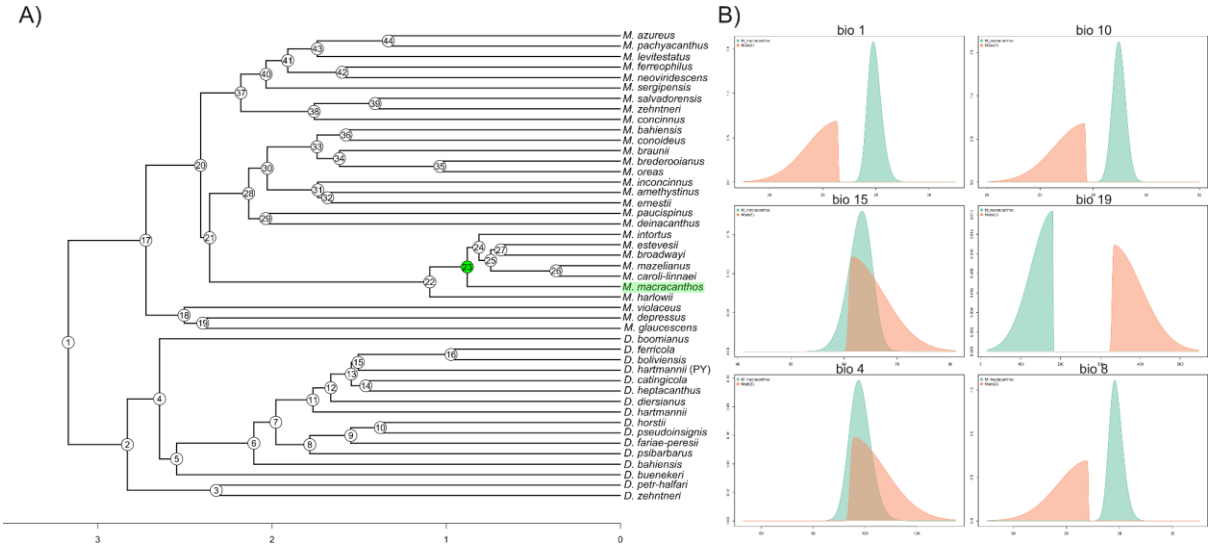


Figure S7. Response curves estimated along the phylogeny using *machuruku* R package. A) Phylogeny used as input on *machuruku* R package, showing node numbers and tip names. B) Response curves estimated for node 26 and their derived lineages (highlighted in green), which correspond to a range shift identified in the same node by *BioGeoBEARS*.

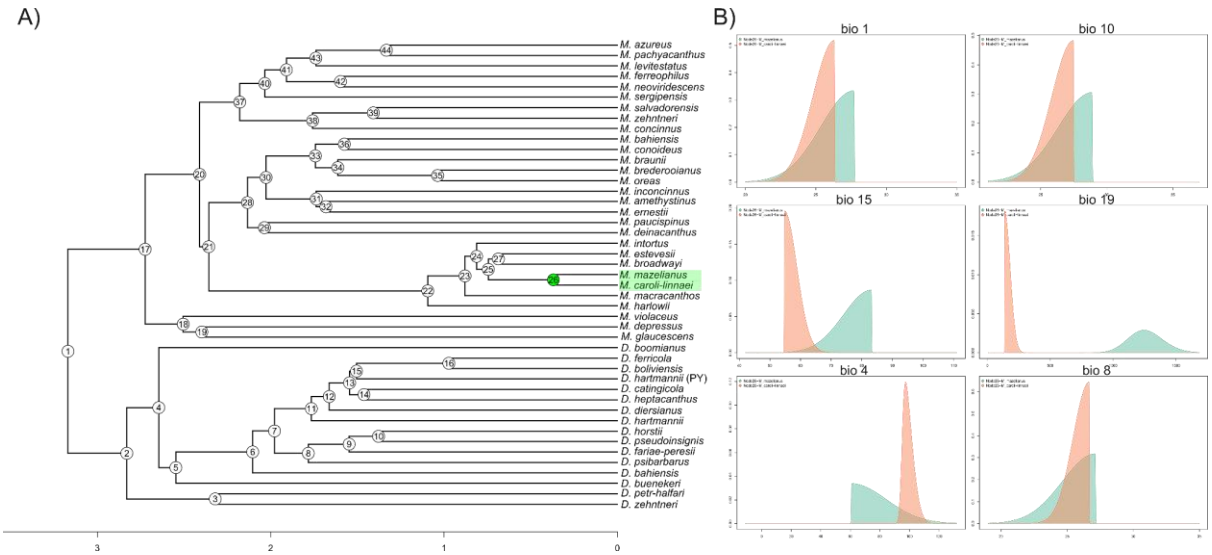


Figure S8. Response curves estimated along the phylogeny using *machuruku* R package. A) Phylogeny used as input on *machuruku* R package, showing node numbers and tip names. B) Response curves estimated for node 27 and their derived lineages (highlighted in green), which correspond to a range shift identified in the same node by *BioGeoBEARS*.

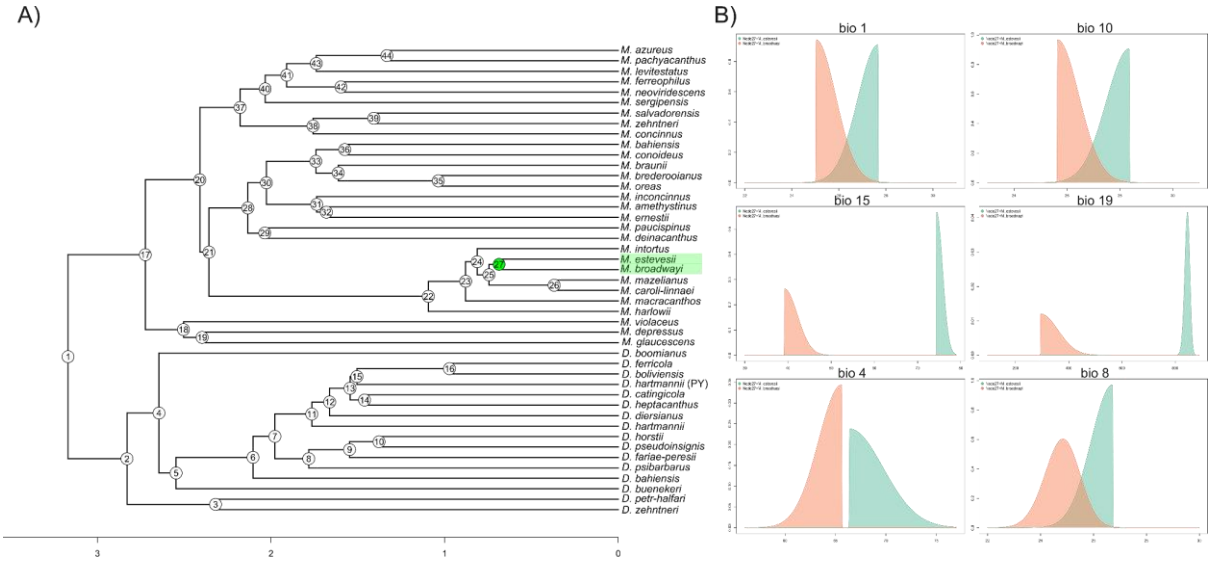


Figure S9. Response curves estimated along the phylogeny using *machurku* R package. A) Phylogeny used as input on *machurku* R package, showing node numbers and tip names. B) Response curves estimated for node 29 and their derived lineages (highlighted in green), which correspond to a range shift identified in the same node by *BioGeoBEARS*.

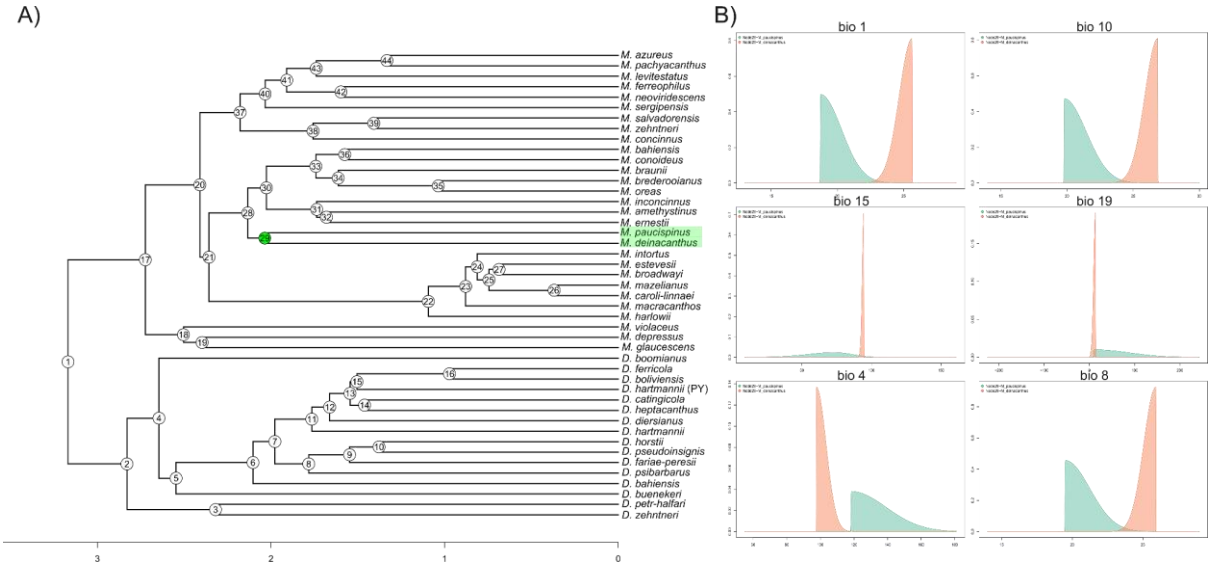


Figure S10. Response curves estimated along the phylogeny using *machuruku* R package. A) Phylogeny used as input on *machuruku* R package, showing node numbers and tip names. B) Response curves estimated for node 32 and their derived lineages (highlighted in green), which correspond to a range shift identified in the same node by *BioGeoBEARS*.

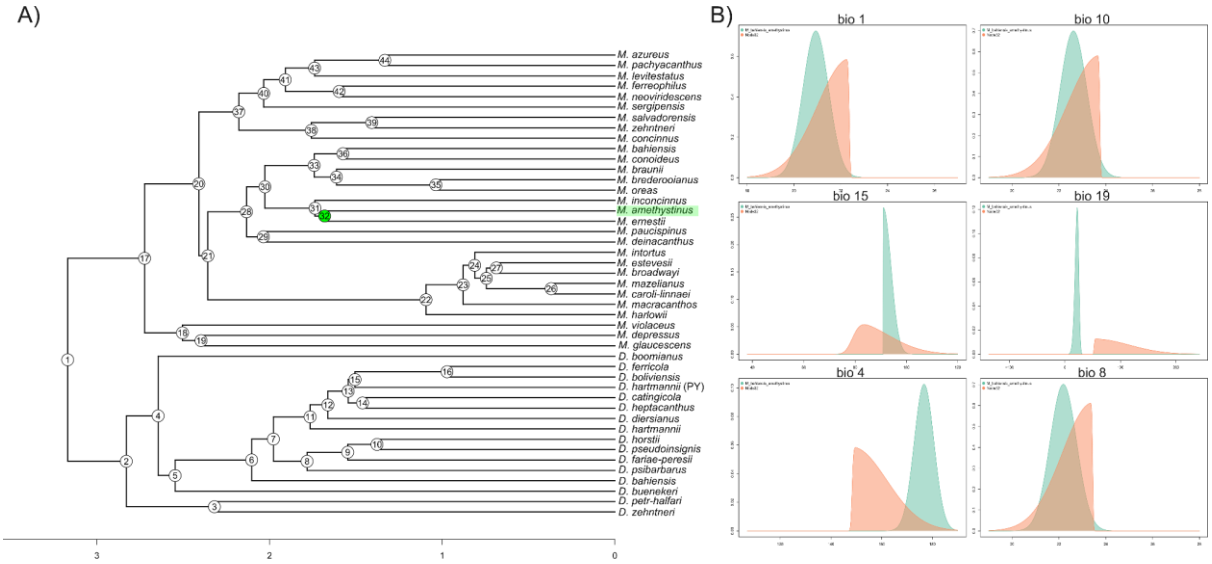


Figure S11. Response curves estimated along the phylogeny using *machuruku* R package. A) Phylogeny used as input on *machuruku* R package, showing node numbers and tip names. B) Response curves estimated for node 32 and their derived lineages (highlighted in green), which correspond to a range shift identified in the same node by *BioGeoBEARS*.

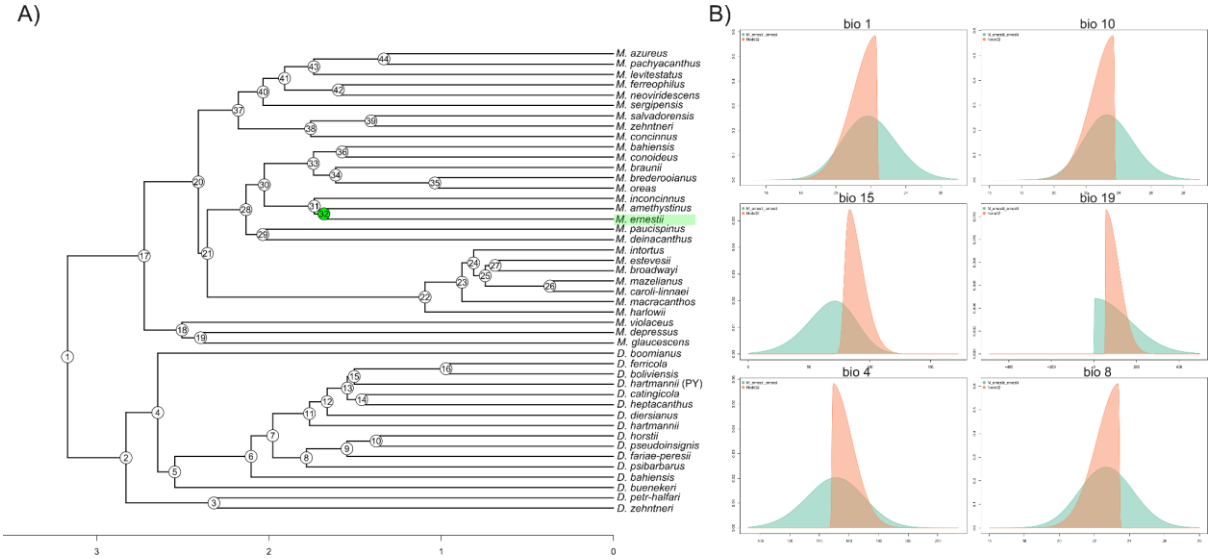


Figure S12. Response curves estimated along the phylogeny using *machuruku* R package. A) Phylogeny used as input on *machuruku* R package, showing node numbers and tip names. B) Response curves estimated for node 36 and their derived lineages (highlighted in green), which correspond to a range shift identified in the same node by *BioGeoBEARS*.

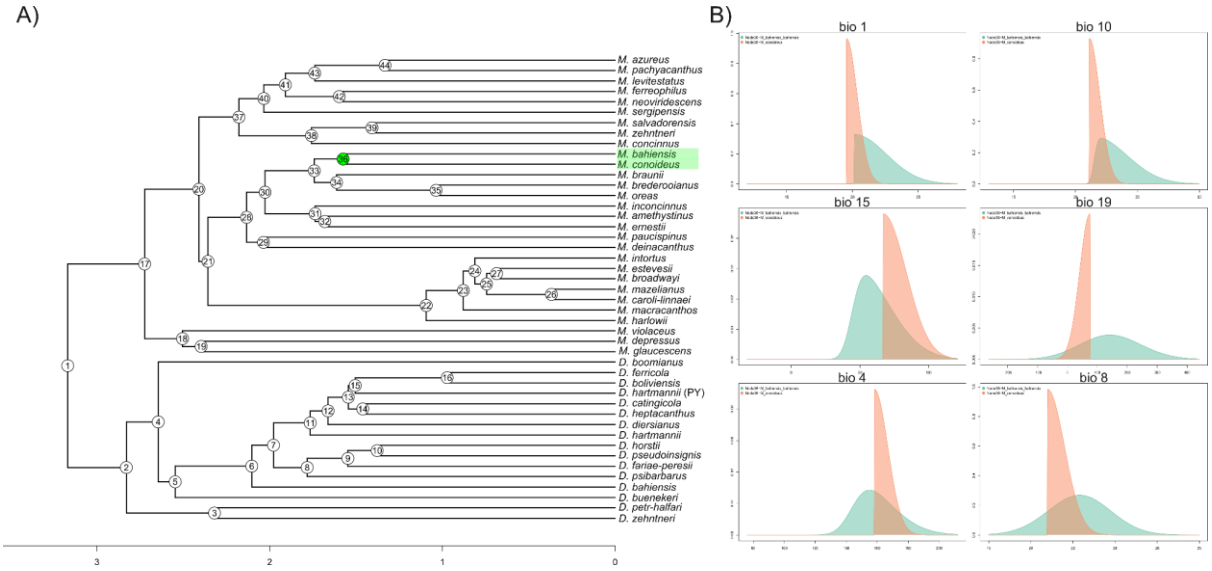


Figure S13. Response curves estimated along the phylogeny using *machuruku* R package. A) Phylogeny used as input on *machuruku* R package, showing node numbers and tip names. B) Response curves estimated for node 38and their derived lineages (highlighted in green), which correspond to a range shift identified in the same node by *BioGeoBEARS*.

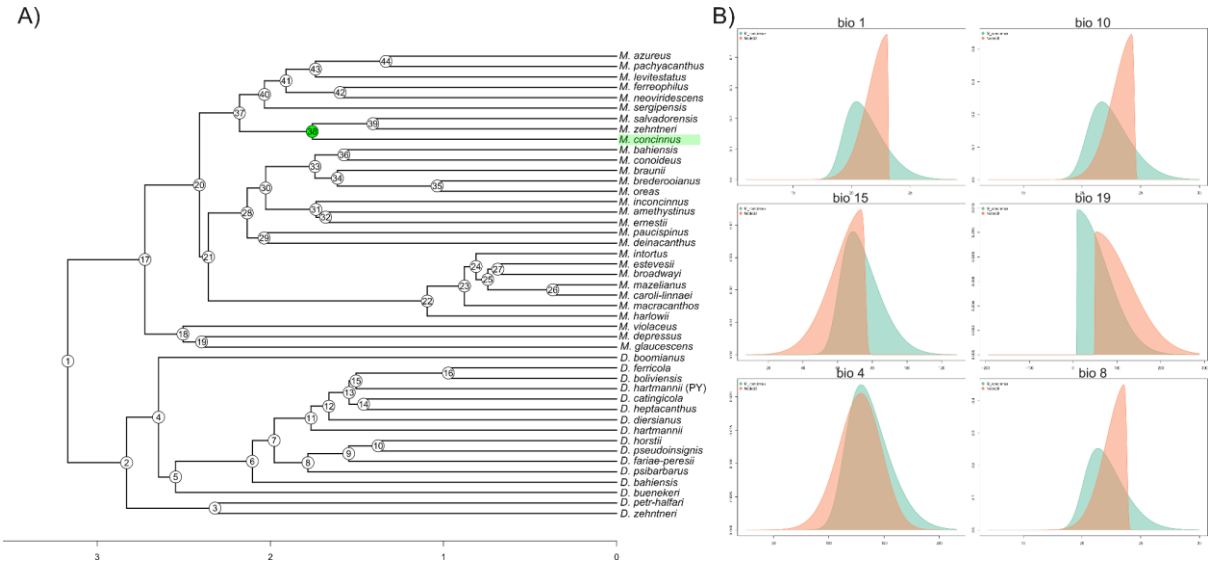


Figure S14. Response curves estimated along the phylogeny using *machuruku* R package. A) Phylogeny used as input on *machuruku* R package, showing node numbers and tip names. B) Response curves estimated for node 39 and their derived lineages (highlighted in green), which correspond to a range shift identified in the same node by *BioGeoBEARS*.

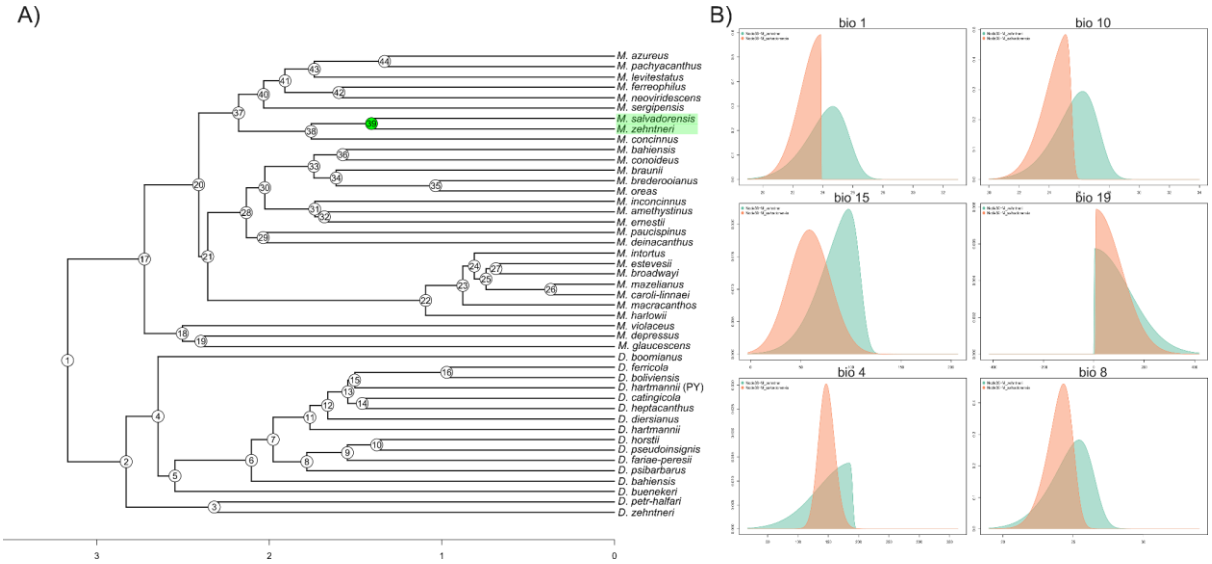
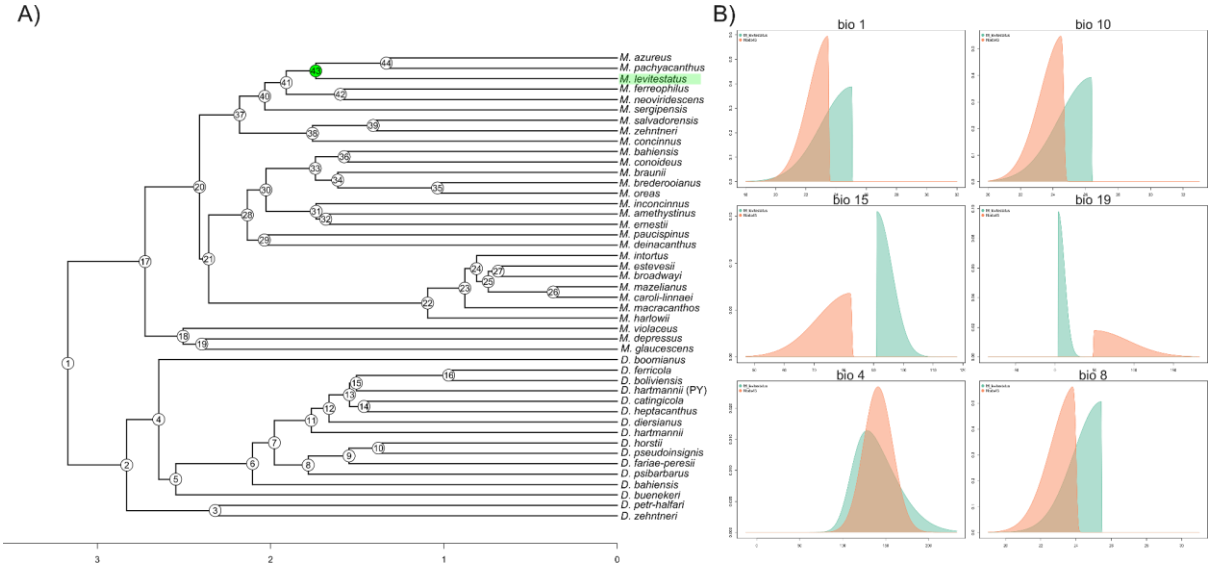


Figure S15. Response curves estimated along the phylogeny using *machuruku* R package. A) Phylogeny used as input on *machuruku* R package, showing node numbers and tip names. B) Response curves estimated for node 43 and their derived lineages (highlighted in green), which correspond to a range shift identified in the same node by *BioGeoBEARS*.



Concluding Remarks and Perspectives

The present thesis comprises three chapters delineating the use of genomic data to investigate the evolutionary patterns and the role of biome shifts in the diversification of lineages from Neotropical dry and open habitats, using the sister cacti genera *Melocactus* and *Discocactus* as a biological model. Although *Melocactus* and *Discocactus* are informative models for studying the processes of dispersion and connection between Neotropical dry and open biomes due to their sister relationship, geographical distribution, and the presence of a shared adaptive characteristic related to the colonization of a fire-prone habitat, neither genus had a robust phylogenetic hypothesis. Therefore, obtaining a robust phylogeny was the first and fundamental step to address questions related to lineage origins, biome shifts, trait evolution, and diversification rates, which were the key focal points of this study.

The first chapter elucidated the phylogenetic relationships within the species-rich genus *Melocactus*. We inferred the first robust phylogenetic hypothesis with a comprehensive Brazilian taxonomic sampling, which enabled us to propose some taxonomic realignments. Besides, it revealed a complex genus evolutionary history, intrinsically associated with climatic variations and soil specialization. In this chapter, we underscore the importance of integrating phylogenetic reconstruction with ecological variables.

In the second chapter, we proposed the first well-supported phylogenetic hypothesis of *Discocactus*, providing clearer insights into the species relationships. The use of genome-scale datasets and coalescent-based analysis was crucial for elucidating the number of *Discocactus* species, particularly in a scenario of disparate taxonomic classification. Considering the critical conservation status of most *Discocactus* species, clearly delineating its species numbers is crucial to a more accurate determination of the actual conservation status of each species, a critical step in developing conservation strategies. We also used machine learning methods to screen morphological traits that may have enabled the occupancy of different habitats by *Discocactus* species. We identified potential adaptive traits within *Discocactus* species and their correlations with different environmental conditions, offering new perspectives into *Discocactus* evolutionary history. Additionally, we demonstrated the value of integrating predictive machine learning with phylogenetic comparative methods to test hypotheses related to traits and their ecological and evolutionary aspects.

The third and last chapter investigates the biome origin of *Melocactus* and *Discocactus*, and the role of biome shifts within their diversification. The origin of both *Melocactus* and

Discocactus occurred within the Seasonally Dry Tropical Forest (SDTF), with subsequent colonization of the Savanna through dispersal events and in-situ diversification of *Discocactus*. Although range shifts were inferred, only a few corresponded with climatic niche shifts, and, therefore, were characterized as biome shifts. These biomes shifts, however, do not appear to have directly influenced the diversification of these genera, even when accounting for the presence of adaptive traits like fire adaptation. The origin and range shifts of both genera during the Pleistocene coincided with climatic and geological changes in northeastern Brazil, which contributed to the environmental heterogeneity of the region. The combination of climatic oscillations along with increasing spatial edaphic heterogeneity, likely acted as a trigger for the diversification of *Melocactus* and *Discocactus*. These results enabled us to draw attention to the need for more caution while discussing biome shifts and their role in lineage diversification. A more comprehensive approach, incorporating multiple and independent lines of evidence, particularly considering past environmental conditions, may provide more robust insights into the role of biome shifts in the evolutionary history of Neotropical plant lineages.