

Review

Evolutionary Genetics of Cacti: Research Biases, Advances and Prospects

Fernando Faria Franco ^{1,†}, Danilo Trabuco Amaral ^{1,2,†}, Isabel A. S. Bonatelli ³, Monique Romeiro-Brito ¹, Milena Cardoso Telhe ¹ and Evandro Marsola Moraes ^{1,*}

¹ Departamento de Biologia, Centro de Ciências Humanas e Biológicas, Universidade Federal de São Carlos (UFSCar), Sorocaba 18052-780, Brazil; franco@ufscar.br (F.F.F.); danilo.trabuco@gmail.com (D.T.A.); romeiro.monique@gmail.com (M.R.-B.); milena.telhe@gmail.com (M.C.T.)

² Programa de Pós-graduação em Biologia Comparada, Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Universidade de São Paulo (USP), Ribeirão Preto 14040-901, Brazil

³ Departamento de Ecologia e Biologia Evolutiva, Universidade Federal de São Paulo (UNIFESP), Diadema, São Paulo 04021-001, Brazil; belbonatelli@gmail.com

* Correspondence: emarsola@ufscar.br

† These authors contributed equally to this work.

Abstract: Here, we present a review of the studies of evolutionary genetics (phylogenetics, population genetics, and phylogeography) using genetic data as well as genome scale assemblies in Cactaceae (Caryophyllales, Angiosperms), a major lineage of succulent plants with astonishing diversity on the American continent. To this end, we performed a literature survey (1992–2021) to obtain detailed information regarding key aspects of studies investigating cactus evolution. Specifically, we summarize the advances in the following aspects: molecular markers, species delimitation, phylogenetics, hybridization, biogeography, and genome assemblies. In brief, we observed substantial growth in the studies conducted with molecular markers in the past two decades. However, we found biases in taxonomic/geographic sampling and the use of traditional markers and statistical approaches. We discuss some methodological and social challenges for engaging the cactus community in genomic research. We also stressed the importance of integrative approaches, coalescent methods, and international collaboration to advance the understanding of cactus evolution.

Keywords: Cactaceae; literature survey; evolutionary genetics; genomics



Citation: Franco, F.F.; Amaral, D.T.; Bonatelli, I.A.S.; Romeiro-Brito, M.; Telhe, M.C.; Moraes, E.M.

Evolutionary Genetics of Cacti: Research Biases, Advances and Prospects. *Genes* **2022**, *13*, 452.

<https://doi.org/10.3390/genes13030452>

Academic Editor: Zengjian Jeffrey Chen

Received: 25 January 2022

Accepted: 25 February 2022

Published: 1 March 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Cactaceae (Caryophyllales, Angiosperms) is the major lineage of succulent plants, originating during the Eocene-Oligocene transition [1–3], and it is recognized by its remarkable diversity [4,5]. It is one of the most conspicuous examples of species radiation in the Americas, with intense diversification in the last 10 million years [1]. Cactus species have been traditionally studied by morphologists, and several taxonomic reorganizations have been proposed in the last century (e.g., [4,6–8]). As an emblematic and ecologically relevant group, many evolutionary studies using genetic-based markers have been published in recent decades (e.g., [9–15]). However, in several examples, classical molecular markers often lack sufficient information to resolve phylogenetic relationships (e.g., [16–19]) and genetic variation at intraspecific level (e.g., [20–22]). Considering the efforts undertaken so far in sampling traditional molecular markers for the family Cactaceae, and that the increasing accessibility to new sequencing technologies [23,24] has fostered genomic sampling in Cactaceae, we believe that a turning point has been reached and that this is the appropriate context for a review of current knowledge achieved using those molecular markers.

Evolutionary studies using molecular markers have been routinely performed to address questions on the natural history of living organisms, such as life histories, population genetics, gene flow, taxonomy, and evolutionary relationships [25]. Moreover, target taxa

may be useful as a biological model to study broader aspects of their habitats and their geography, such as the environmental fluctuations and evolution [26]. In this regard, cacti have relevant idiosyncrasies to be considered biological models for evolutionary studies. For example, given that it is a recent radiated group [1], many clades might be useful to study recent and explosive diversification. Moreover, as many species may hybridize in nature (e.g., [13,27,28]), cacti may provide many insights into hybrid zones, species cohesion, and species delimitation. Finally, due to its intrinsic association with drought stress, cacti are useful models for biogeographic approaches aiming to investigate the effects of Quaternary climate oscillations on xeric biomes (e.g., [21,29–32]). As a result, many studies with an evolutionary focus have been conducted on cacti in recent decades.

Here, we performed a literature survey to obtain detailed information and summarize important aspects of molecular marker-based studies investigating cactus evolution. Specifically, we address historical tendencies and the combined use of molecular markers, phylogenetics, biogeography, species delimitation, hybridization, and genome assemblies. Many advances, limitations, and biases in cactus research are presented. We emphasize the importance of an integrative approach in the future to properly understand family evolution and alternative methods for data analysis. Finally, we discuss the methodological challenges involved in genomic research as well as the need to integrate the cactologist community, especially those from low-income countries with high cactus diversity.

2. Materials and Methods

The core of this review focused on evolutionary genetics works, mostly related to biodiversity descriptions, such as phylogenetic, phylogeography, population genetics, species delimitation, and hybridization studies. For this purpose, studies of introduced/nonnative greenhouse cacti and those that did not use molecular markers were not considered. The literature survey was carried out in the Web of Science ISI database (til 31 December 2021) using the following terms: “phylogeograph*” + “Cact*”; “phylogenet*” + “Cact*”; “delimit*” + “Cact*”; “hybridization” + “Cact*”. Specific searches for microsatellite markers and their transferability were applied using the terms “SSR” + “transferability” + “Cact*” and “microsatellite” + “transferability” + “Cact*”. This non-exhaustive survey was followed by the inspection of the papers’ titles, abstracts, and keywords. Based on our interpretation of the main authors’ aims, we classified the papers into three main topics: phylogeny, population genetics, and phylogeography. From each paper, we recorded the following information, partitioned in (i.) publication year, DOI, first/last author country; (ii.) taxonomy (taxon/clade), species geographic occurrence (based on dominion proposed by [33]); (iii.) molecular marker type; (iv.) main topics; (v.) methodological approach; (vi.) taxon-specific goals; (vii.) biogeographic hypothesis(es); and (viii.) main conclusions (detailed in Tables S1 and S2 in the Supplementary Information).

An additional survey (til 31 December 2021) was performed to recover studies applying functional genomic approaches into the study of Cactaceae (genome and transcriptome datasets). For that, we used the following terms: “genom*” + “Cact*”, “Transcriptom*” + “Cact*”, and “RNA-Seq” + “Cact*”. We checked the papers’ titles, abstracts, and keywords as mentioned above (Table S3). The recorded information was as follows: (i.) publication year, DOI, first/last author country; (ii.) taxonomy (taxon/clade); (iii.) genomic source (nuclear, plastid, or mitochondrial); and (iv.) main conclusions (detailed in Table S3 in the Supplementary Information). We generated two flow diagrams based on Page et al. [34] to explore our survey method (Figure S1).

3. Results

We retrieved 181 articles published from 1992 to 2021, of which 166 papers were mainly associated with phylogeny, population genetics, and phylogeography topics (Tables S1 and S2), while 27 papers were published since 2015 and were related to genomic and transcriptome assemblies (Table S3). Ninety papers were categorized in the phylogeny, 67 in the population genetics, and nine in the phylogeography topics (Tables S1 and S2). Most of

these papers were led by researchers from the United States, Mexico, and Brazil (Figure S2). The number of papers in these fields has grown in the past two decades (Figure 1a). As expected, PCR-based markers are historically prevalent in evolutionary studies of cacti, but the number of studies using genome-wide markers has been constantly growing in the last five years (Figure 1b). Most studies are concentrated in two cactus clades: BCT (Browningieae–Cereae–Trichocereae) and Opuntiae (Figure 2). The Nearctic, the Mexican Transition zone, the Mesoamerican, and the Chacoan Domains (*sensu* [33]) are the most represented biogeographic domains in our literature survey. However, these studies covered only a small fraction of all the species richness and endemism of the Neotropical region (Figure 3).

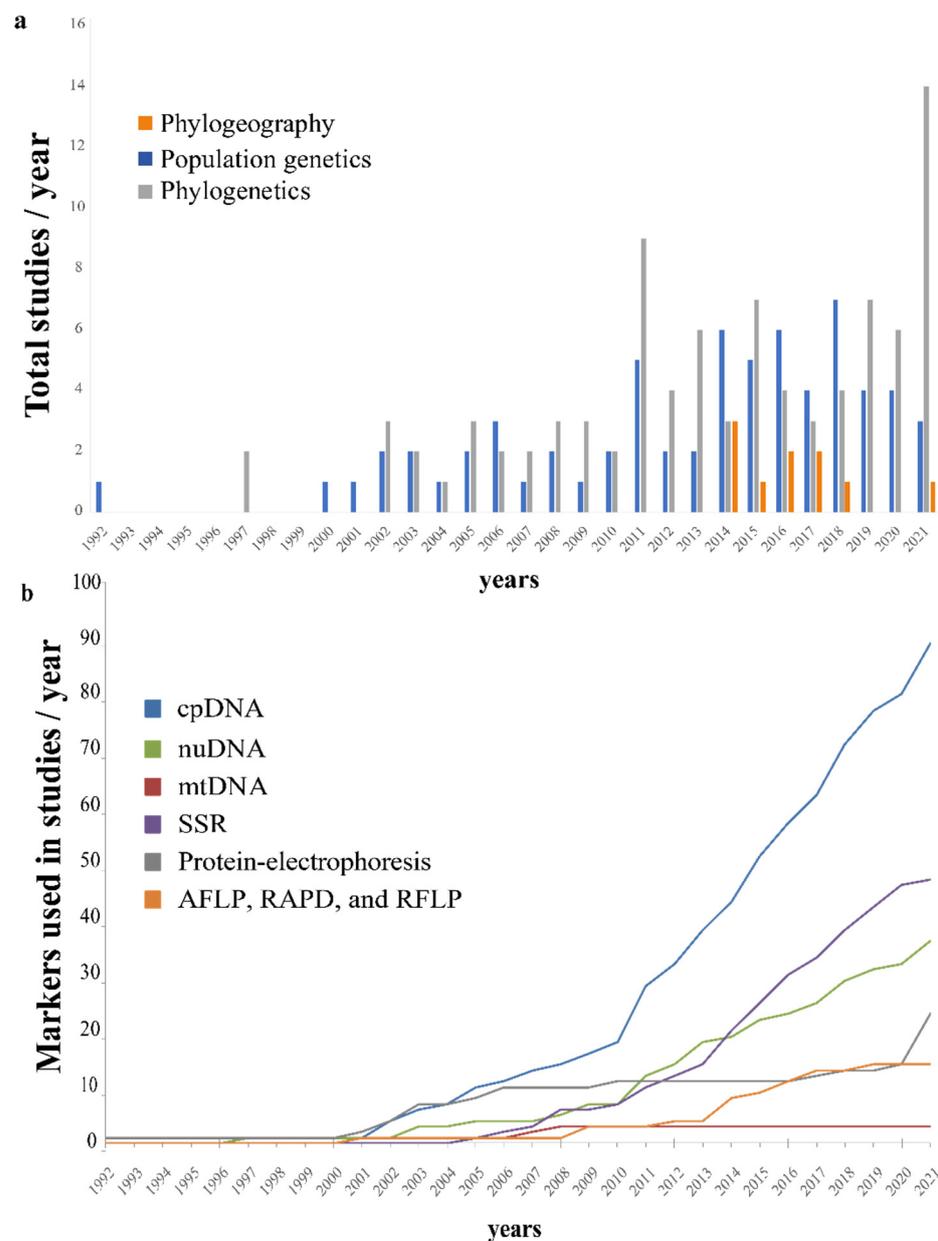


Figure 1. The number of evolutionary studies in Cactaceae, tabulated by main topics described in the text (a), and cumulative curve of major molecular markers used in the studies (b), conditioned on survey data spanning 1992–2021.

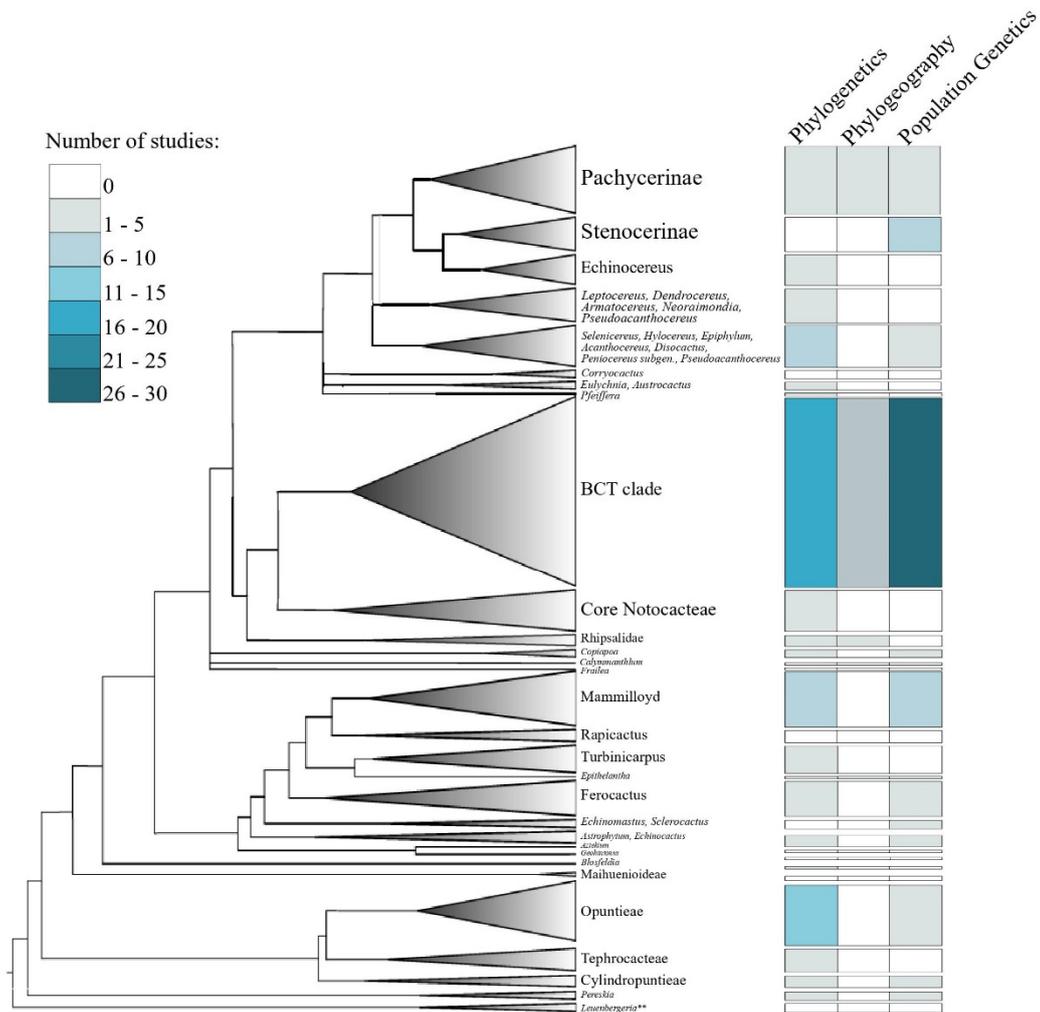


Figure 2. Distribution of evolutionary studies on the cactus backbone phylogeny (Adapted from [19] permission has been obtained).

The phylogenetic category is the most multidisciplinary topic, connected to the other two topics (i.e., population genetics and phylogeography), and with all approaches identified in the recorded papers (Figure 4). Many phylogenetic studies have been conducted to answer questions on divergence time, ancestral reconstruction, taxonomy, and species circumscription. Conversely, phylogeography and population genetic categories are less associated among them, and among study approaches. The population genetics’ studies were most interested in conservation purposes, although they also displayed fewer interactions with other topics.

Many of the recovered papers (>50%) relating to genomic and transcriptome assemblies were published in the past two years, most of which were related to plastome or transcriptome assembly reports (Table S3). This suggests a growing interest in cactus genomics as sequencing technologies become more feasible to researchers.

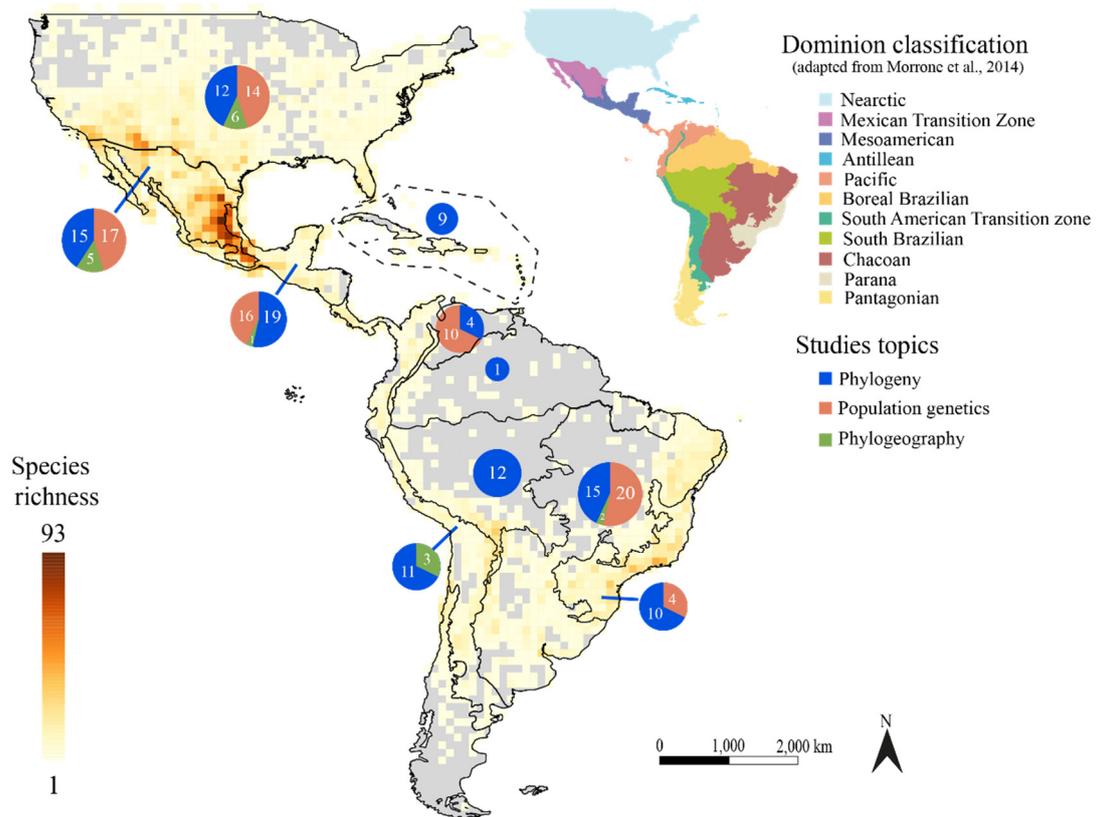


Figure 3. Distribution of species richness within the Cactaceae family according to the dominion classification of Morrone et al. [33]. The number of studies in the main regions is shown in blue circles.

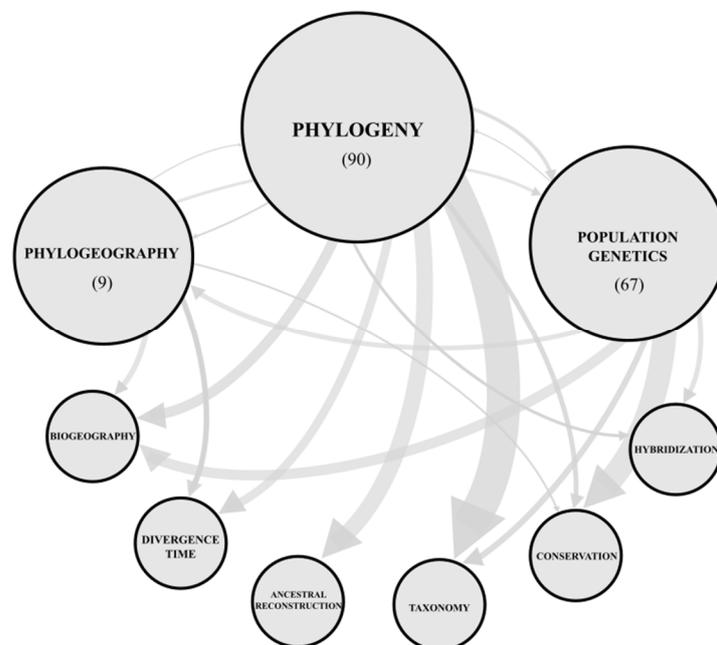


Figure 4. Network showing connections among the main topics addressed in evolutionary studies of Cactaceae. Circle sizes are proportional to the number of studies in each category (see Table S1), and line thickness is proportional to the number of connections between categories.

4. Discussion

4.1. Molecular Markers

The use of molecular markers for phylogenetic, population genetics, and phylogeographic purposes has followed historical tendencies [35]. Protein electrophoresis (allozymes and isozymes), restriction fragment length polymorphism (RFLP), random amplified polymorphic DNA (RAPD), and amplified fragment length polymorphism (AFLP) were the primary markers used in evolutionary studies in cacti (26 studies), but they have lost their importance in the past decade. Since the 2000s, there has been a transition to more polymorphic PCR-based markers (Figure 1b). As a result, most of the studies started to use plastid DNA sequences (cpDNA; 90 studies), followed by Simple-sequence repeats (SSRs; 48 studies) and nuclear DNA (nrDNA; 37 studies) markers (Figure S3). The use of high-throughput genomic data has been applied in the past decade, comprising 10 studies.

Approximately 28% of the studies (45 studies) combined the use of multiple markers. Most of them employed cpDNA plus nrDNA (31 studies), followed by the combination of cpDNA plus SSR markers (seven studies) (Figure S3). The mtDNA (mitochondrial DNA) marker was used in only three phylogenetic studies, probably due to the slower evolution, unstable structure, and higher recombination rates of this genome in plants [36,37]. On the other hand, plastid markers have been traditionally used for evolutionary studies within Cactaceae, favored by the availability of universal primers to segments potentially informative for multiple phylogenetic scales across angiosperms [38]. Although plastid regions presented highly informative regions to resolve phylogenetic relationships within Cactaceae tribes (e.g., [39]), these regions generally showed low potential to resolve deep internodes at the species level (e.g., [18,40]). Furthermore, plastid markers are prone to lineage-specific variations even among closely related species [41]. For cacti, this issue was tested for the genus *Cereus*, agreeing with the non-correspondence between the variation in plastid regions at the inter- and intraspecific levels [22,42]. The practical consequence is the need for a preliminary essay for screening molecular variation in candidate regions to be applied in the target group, a step that may be laborious. Nonetheless, it is important to highlight the valuable phylogenetic information enclosed in the whole plastome data. In Cactaceae, the use of a complete chloroplast genome has been recently explored [43,44], highlighting the potential of plastomes for phylogenetic purposes in this family.

A large proportion of studies using PCR-based technologies combine markers from plastid and nuclear genomes (Figure S3). Although the ribosomal internal transcribed region (ITS) presents extensive paralogy due to the absence of concerted evolution in some cactus clades, such as *Mammillaria* [45,46], it was one of the most common nrDNA marker adopted in combined studies, in addition to the phytochrome C gene (*PhyC*). The *trnK*, *matK*, *trnL-F*, and *trnL* were the most used cpDNA markers. The combination likely occurred due to a low mutation rate of cpDNA markers compared to nrDNA regions [36,47] and low sequence variation of cpDNA among closely related species or among conspecific populations (e.g., [21,48,49]). Even so, the use of multiple independent molecular markers has been a successful strategy to resolve phylogenetic relationships within rapid and recent divergent groups of Cactaceae [21,50–52]. Furthermore, the integrative use of plastid and nuclear genome information allows us to understand many biological issues in plants, such as patterns of plastid-nuclear discordance caused by differential dispersal between pollen and seeds.

The SSR is the second most commonly used type of molecular marker (Figure 1b). Microsatellites show recognized advantages to population genetics' studies [53,54]. Moreover, the possibility of effectively transferring these markers across taxa may overcome the challenges, cost, and time-consuming development of new SSR primers [53,55]. In cacti, this technique has been applied with success (Table S1) at both the cross-species and cross-genera levels [28,56–60]. In addition to the possibility of cross-amplification using heterologous primers, the emergence of NGS technologies has allowed the description of new SSR loci from bioinformatic surveys (e.g., [51,61]). In this context, SSRs remain a

marker of choice for exploring intraspecific genetic variation in cactus studies, even in the genomic era.

Although relatively few genomic-scale studies were recovered in our literature survey, we noticed an increasing use of genome-wide markers in the past five years (Figure 1b). For example, different methods to sequence a reduced representation of genomes have been used in evolutionary studies on cacti, resolving rapid and recent diversification in this group (e.g., [50,62,63]). In addition, de novo assembled genomes, transcriptomes, and plastomes have been successfully used in some cacti groups (e.g., [32]). These recently available genomics data are a valuable source to assist target enrichment sequencing approaches, enabling the identification of hundreds of orthologous regions (e.g., [64]). This technique may enable obtaining sequences from multiple and unlinked nuclear orthologous genes to perform evolutionary studies at multiple phylogenetic scales on this group.

Representative sampling is a critical factor when using molecular markers for intraspecific inferences. For instance, the sample size is an outstanding predictor of population structure inference [65], evidencing a bias caused by inadequate sampling on within- and among-population variation. In another way, when thousands of loci are sampled by high-throughput sequencing, very small sample sizes (i.e., two individuals) seem to be enough to properly sample within-population variation [66]. Cactaceae diversity has several attributes that demand more sampling efforts than less diverse plant groups. For example, despite the small range size of many cacti species, which would demand lower sample sizes, the extreme genetic differentiation observed even among close but physically isolated populations (e.g., [48,67]) requires a spatially more accurate sampling of the populations. Ultimately, both the robustness and accuracy of phylogenetic and population genetic inferences depend on the appropriate sampling of the genetic diversity within the taxon.

4.2. Phylogenies

The phylogenetic studies on Cactaceae are concentrated on two main clades, Browningeae-Cereae-Trichocereae (BCT) and Opuntiae (Figure 2; Table S2). These clades are among the most diverse groups in this family (Figure 2) and are mainly located in two centers of the highest cacti diversity (Figure 3; [19]). However, some clades are poorly explored in a phylogenetic context, such as the subtribes Stenocereinae and *Corryocactinae*, in addition to the *incertae sedis* genera *Calymmanthium* and *Frailea*. The poor representation of these clades in phylogenetic studies might be because most of the diversity in these taxa occurs outside of countries where cactus phylogenetics studies are concentrated (USA, Mexico, and Brazil; Figure S1). In the same way, some hotspots of cacti diversity, such as the Andes/Chaco region, are rarely represented in cacti phylogenetic studies, probably due to the low number of researchers interested in the molecular phylogeny of cacti in the countries within these regions.

In Cactaceae, poor resolution and the presence of recalcitrant nodes are common observations [19]. The recent and rapid diversification of major clades within this family [1,2] coupled with the use of few molecular markers are likely behind the lack of phylogenetic resolution and node support in phylogenies encompassing different taxonomic levels. Single- or few-locus phylogenetic approaches provide insufficient genetic information to overcome the phylogenetic bias produced by natural phenomena in the genome of these taxa, such as incomplete lineage sorting (ILS), hemiplasy, and hybridization events [15,68]. An alternative approach to overcome these challenges is to use multiple loci data under coalescent phylogenetic methods, increasing the chance of obtaining more robust and accurate phylogenies [69,70]. Indeed, this is a rising trend observed in evolutionary studies within Cactaceae (see Molecular Markers topic). It is worth noting that the rapid differentiation experienced by many cacti lineages [1] may lead to short-deep internodes characteristics of “anomaly zones” in the tree [71], imposing additional challenges to phylogenetic reconstruction, such as those observed in the *Cereus-Cipocereus* branch [62]. Many cactus

phylogeneticists will need to deal with these theoretical/methodological issues, regardless of the amount of dataset used.

The majority of recovered phylogenetic studies in Cactaceae implement the concatenation approach. By equalizing conflicting genealogical histories of independent genes in a super matrix, this approach may produce well-supported misleading phylogenetic relationships [72]. A way to overcome this challenge is to use methods that incorporate the different coalescent time and genealogical history of each locus independently (e.g., ASTRAL-III [73]; SVDquartets [74]; STARBEAST2 [75]; SNAPP [76]). Therefore, the use of coalescent phylogenetic methods is extremely important to be applied in groups that lead to intensive gene tree discordance, as occurs in cacti [15].

4.3. Biogeography

Biogeographic studies using molecular markers in Cactaceae have become an area of interest only since the beginning of the 2000s (Table S1). In the last two decades, we observed 33 studies that applied biogeographic reconstruction methods within the three main topics proposed here; approximately 50% are based on phylogenetic inference within genera or higher taxonomic levels (16 studies; Figure 4). Downstream analyses related to phylogenetic approaches, such as divergence time estimates, are mandatory for well-calibrated biogeography reconstruction [77]. The association between the reconstruction and the test of postulated biogeographic hypotheses is focused on understanding the patterns and processes that drive species diversification [78,79]. However, even using biogeographic/phylogeographic approaches, some of the papers do not frame their questions as particular or general biogeographical hypotheses (Tables S1 and S2) to be tested, restricting their findings to a description of ancestral geographic ranges of the target clade. It is desirable that further biogeographic studies should be more explicit regarding the biogeographic conjecture target of testing, to improve the comparisons and to identify drivers of diversification into the family.

Based on the recovered papers, we were able to identify four main associated biogeographic theories: (i.) “Geomorphology”; (ii.) “Riverine”; (iii.) “Quaternary climatic changes”; and (iv.) “Neogene aridification” (Table S4). Most of the studies suggested “Quaternary climatic changes” (ca. 72%) and/or “Geomorphology” (ca. 27%) as the most relevant hypotheses to explain range shifts, distributions, and diversification of cactus species in the Neotropics. For instance, the climatic oscillation involving the glacial cycles of the Pleistocene is thought to have promoted the expansion and contraction (fragmentation) range of populations from distinct genera, such as *Pilosocereus*, *Cephalocereus*, *Rhipsalis*, and *Echinopsis* [21,29,31,50,80]. Moreover, paleoclimates drove marine transgressions/regressions, which have been inferred to have impacted the population dynamics of coastal and island populations [3,60,81]. Geomorphologic events, such as the uplift of the Andes, Trans Mexican Volcanic Belt structures, and geomorphological compartmentalization of the Brazilian Central Plateau, also influenced the distribution and expansion of cactus species [16,21,81–83]. The “Neogene aridification” theory advocates the expansion and recent radiation of the Cactaceae in the last 35 Ma (Miocene to Pliocene) [1,2,13,32,84]. The “Riverine” theory was recovered only once, associated with the genus *Pilosocereus* and the eastern Brazilian rivers, such as the Sao Francisco and Doce rivers [85]. One important consideration is that alternative biogeographic hypotheses should not be considered as mutually exclusive, as most biogeographic patterns can be explained by the combined and continuous effects of different historical events [86]. Most studies accessed here show a continuous diversification trend in cactus clades from the Neogene to the Pleistocene (Table S4 and Figure S4).

The study’s topics in Cactaceae are restricted to a few biogeographic domains (Table S2), such as the Nearctic, Mexican Transition Zone, Mesoamerican, and the Chacoan sensu [33], which are still far from covering all the species richness and endemism of the Neotropics (Figure 3). This bias is intensified by a knowledge trend: political boundaries are barriers for researchers to explore large-scale diversity patterns, as most researchers concentrate

their works within the limits of their countries. An additional challenge in both empirical and theoretical biogeographic studies relate to the definition of geographic operational units, which are generally set as the biomes in the study region. The delineation of biogeographic units can become quite a challenging task given the high diversity and complexity of Neotropical biomes, for example, those exhibiting a mosaic-like distribution [87,88]. Therefore, to acknowledge the limits and definitions of a biome, several and distinct particularities must be considered, such as the distribution, potential natural vegetation, and general abiotic environmental variables [89–91].

A biogeographical synthesis of Cactaceae is lacking as well as a more complete knowledge of diversification drivers in drylands [92]. Xeric landscapes display a complex evolutionary history, including several hotspots of diversity and centers of diversification [1,93,94]. Drylands also display great functional diversity and morphological/physiological adaptation, distinct interactions between plants (i.e., mutualism), plants plus soil (i.e., inorganic compounds), and plants plus microbiota (i.e., nematodes, bacteria, insects, and viruses [92,95,96]). Why do we have so much species richness in environments with such low water availability and nutrient-poor soil? The biogeographical theory recovered in these studies, as we cited, brought mechanisms to determine and explain part of this question. However, many other mechanisms still need to be evaluated and tested, exploring new theories, distinct cactus taxa, and new methodological and statistical approaches.

Like many other areas of biodiversity sciences, biogeography requires efforts toward an integrative approach by using organismal traits, environmental variables, and genetic data [97–99]. Recent studies within the field of phylogeography, for instance, have developed predictive frameworks to aggregate this information and predict patterns of diversity along with species distribution [100,101]. These approaches have taken advantage of the predictive power of machine learning and the availability of distributional, climate, and trait data housed in public databases such as Global Biodiversity Information Facility GBIF [102], Worldclim [103], and TRY [104]. Considering the importance of life-history traits to predict patterns of diversity and distribution within taxonomic groups, we argue that the development of trait-based approaches, which have already started in the phylogeography of other taxonomic groups [97–99,105,106], could be extended to broad applications in the biogeography of Cactaceae.

4.4. Species Delimitation

There are some reasons to use molecular data to assist species recognition in cacti. First, cacti experience a diverse range of threats [93,94], and molecular data may contribute to revealing unprotected cryptic diversity or identifying smuggled living cactus. Second, the family has an astonishing diversity of species and growing forms [4], in addition to a substantial interpopulation variation in morphology, which may cause taxonomic uncertainty. Altogether, these factors contribute to the origin of many “species complexes” in Cactaceae, which are systems that may impose major challenges for species delimitation [26]. Consistently, molecular markers have been used to assist species recognition and species circumscription in the cactus family (Figure 4), most of which are biased to PCR-based markers within disparate phylogenetics and population genetics approaches. However, incorporating modern sequencing technologies using coalescent-based methods is of great promise for this purpose [107].

Distance-based methods, for example, were useful to identify approximately 77% of the 528 species of Mexican cacti using sequences of the plastid gene *matK* (“*matK* barcode”) [108]. This strategy may have many practical applications, although its use is taxonomically limited due to challenges that preclude the adoption of a universal barcode strategy in plants [109]. A limiting factor, for instance, is the low genetic variability found in most plastid DNA sequences at the shallowest taxonomic level, leading to the need to combine several sequence markers. Moreover, another major issue is the lineage-specific variation as a consequence of heterogeneity in the evolutionary rate among plant lineages,

such as those found in cactus [42], which limits a universal application of the same marker across broad taxonomic spectra.

The integrative framework assessing both molecular and phenotypic/ecological traits has been recently adopted to shed light on taxonomically species-level puzzling groups (e.g., [110–113]). Although this is not always explicit, the underlying species concept in most of these approaches is the practical phylogenetic concept [114,115] or the genotypic cluster concept [116]. For example, Aquino et al. [117] combined molecular (four cpDNA regions) and morphological data (quantitative and qualitative traits) to delimit species from the North American genus *Epithelantha* (subfamily Cactoideae, tribe Cacteeae). In addition to performing species recognition, the authors were able to propose new species and a taxonomic key for species diagnosis in this genus. More recently, the same research group used environmental variables and soil characteristics to delimit the *Epithelantha* species [84]. In another study, Alvarado-Sizzo et al. [118] used congruent results obtained from genetic structure, inferred using SSR markers, ecological niche modeling, and morphological data to delimit four species of the *Stenocereus griseus* species complex (subfamily Cactoideae, tribe Cacteeae).

Apart from the formal description, genetic data may also be useful to test alternative hypotheses in taxonomic unstable groups. For example, the *Pilosocereus aurisetus* group (Subfamily Cactoideae, Tribe Cereae) comprises seven recognized species [4] but with a history of repeated taxonomic re-evaluations [119]. Khan et al. [48] used a set of molecular data (SSR and cpDNA regions) and a population structure approach to test taxonomic hypotheses regarding *Pilosocereus jauruensis*, a species exhibiting a sky-island distribution. The authors found populations of *P. jauruensis* showing species-level genetic divergence and geographically coincident with the occurrence of the heterotypic synonymous *P. densivillovus* [120]. Rather than resuscitating the obsolete name *P. densivillovus*, Khan et al. [48] highlighted the level of evolutionary independence of these *P. jauruensis* populations, which should be treated as independent units for conservation purposes. Recently, these results from Khan et al. [48] were used to support the proposition of a new subspecies, *P. jauruensis* subsp. *cinnopetalus* [121].

The treatment given by Khan et al. [48] for their results is consistent with the General Lineage Concept of Species (GLCS), [122]. Briefly, in this concept, the incipient species are evolving metapopulation lineages in which the species' attributes (e.g., phenotypic divergence; monophyly, reproductive isolation) are complementary stages of a process of continuous divergence and not a mandatory characteristic as in the traditional species concepts. The GLCS concept embraces the bridge between microevolutionary processes generating branching patterns, the core of Darwinism theory, and, more specifically, of the phylogeography discipline [123]. In this sense, the multispecies coalescent model (MSC) is a candidate approach to consistently delimit species according to the GLCS by the statistical modeling of the relationship between the gene trees and the lineage history [107,124]. There are several MSC methods available, including those based on full likelihood (e.g., GMYC [125]), Bayesian posterior probabilities (BPP) (e.g., [126]), and supervised [127] and unsupervised machine learning approaches [128].

In our literature survey, we found only one paper using MSC for species delimitation in a cactus taxon. Perez et al. [107] introduced a new framework coupling the MSC with deep learning algorithms (convolutional neural networks) to test alternative species delimitation hypotheses within the *P. aurisetus* species complex. They were able to select a taxonomic 'splitter' hypothesis circumscribing five species instead of the single one currently recognized. The *P. aurisetus* species complex shows characteristics that impose many additional challenges to species delimitation, such as recent divergence, a highly fragmented sky-island distribution, and several instances of local morphological differentiation. These characteristics are shared with many other cactus systems, suggesting that there is an avenue to be explored on species delimitation of cactus using coalescent theory. We stress that the methods for species delimitation are additional tools to assist taxonomists rather

than operational criteria for species diagnosis. They surely may help in the identification of candidate lineages as a “species hypothesis” for the evaluation of specialists.

4.5. Hybridization

Natural hybridization plays an important role in plant evolution [129–131]. Introgressive hybridization, reinforcement, local extinction, or hybrid speciation are the most relevant outcomes of hybridization, but interspecific gene flow may lead to a variety of results depending on the level of divergence between species, geographical context, and the balance with selective forces [132]. The role of hybridization in cactus evolution has long been highlighted [10,13,15,27,28,133,134]. Registers of interspecific and intergeneric cactus hybrids are common in both nature and cultivation. The constancy in basic chromosome number and the presence of self-incompatible flowers are factors that would improve the chance for hybridization in many cacti [28], although the occurrence of hybridization seems to be more frequent in some clades in the family (e.g., genus *Opuntia*). Molecular markers are valuable allies to study the frequency and extension of hybridization [25]. Moreover, many statistical methods are available to investigate gene flow, depending on the depth of hybridization that one aims to detect [132]. For instance, these methods may involve the detection of genetic admixtures and F1 hybrids (e.g., *STRUCTURE* [135]; *NEWHYBRIDS* [136]), symmetric and asymmetric gene flow (e.g., *MIGRATE* [137]; *G-PhoCS* [138]), introgressive hybridization (e.g., *INTROGRESS* [139]), and gene tree inconsistencies (e.g., *SNaQ* [140]). Despite this, based on our survey criteria and categorization, we found relatively few studies exploring hybridization with molecular markers (Figure 4), most of which were related to our population genetics category (nine studies) followed by phylogenetics (five studies).

The genetic evidence for hybridization in the cactus family was generated using different approaches and molecular markers, such as analysis of RAPD banding pattern data [10,141], levels of admixture in codominant markers (allozymes [142]; SSR [143]) and biallelic SNPs generated by RAD-Seq [27] as well as observation of reticulations in phylogenies (e.g., [13,144]). Overall, the molecular data corroborate the occurrence of interspecific hybrids in different genera, such as *Opuntia* [10,141], *Sclerocactus* [145], and *Melocactus* [27]. Furthermore, molecular markers have assisted studies on intergeneric hybrids in the cactus family, such as *Consolea* × *Opuntia* hybrids in the subfamily Opuntioideae [13] and *Espostoa* × *Haageocereus* in the subfamily Cactoideae [28].

Moreover, molecular markers have been useful to test taxon-specific hypotheses on species boundaries and hybridization (e.g., [143,144,146]). Interestingly, independent results have shown the maintenance of species genetic cohesion even when facing some level of genetic admixture [27,145]. For example, Khan et al. [27] investigated four hybrid zones hosting *Melocactus concinnus* and four congeneric species (*Melocactus ernestii*, *Melocactus glaucescens*, *Melocactus paucispinus*, and *Melocactus zehntneri*). Briefly, the authors found intense hybridization in sympatric areas (with a bi and tri-species admixture) but with a low amount of introgression (2–5%), indicating that the studied species present weak pre-mating but strong post-mating reproductive isolation. These results highlight the importance of hybridization and natural selection in the maintenance of the genetic integrity of the *Melocactus* species.

The major role of hybridization in cactus speciation is recurrently explored in genera with high amounts of polyploids [13,14,28]. This is the case of the genus *Opuntia* [13,144], which presents approximately 60% of polyploid species [147]. Recently, using an integrative approach with phylogenetic, cytogenetic, and taxonomic datasets, Köhler et al. [14] showed that the octoploid ($2n = 88$) accessions of *Opuntia* aff. *ficus-indica* found in Santa Fe, Argentina likely presents a putative hybrid origin between the native *O. rioplatensis* and the introduced *O. aff. ficus-indica*. This hybrid form was described as *O. x cristalensis* [14].

Some hybridization studies in our survey have been conducted using SSR markers amplified with heterologous primers. Arakaki et al. [28] used plastid DNA and a set of nuclear SSR markers, formerly described for *Haageocereus* [148], to test hypotheses of hybridization involving species of *Espostoa* and *Haageocereus*. The SSR data showed evidence that hybrids between *Espostoa* and *Haageocereus* are viable, and that hybridization has gone beyond the F1 generation, suggesting the presence of hybrid swarms in some of the studied populations. However, although SSR data are useful to study hybridization, some statistical frameworks demand larger datasets.

Here, we recovered only one paper using genome-wide markers to explore hybrid zones [27]. It is worth mentioning that the use of RAD-Seq markers by Khan et al. [27] greatly improves the resolution of the genetic architecture of hybridization in the genus *Melocactus* in comparison with previous studies based on traditional markers or phenotypes (e.g., [142]). Undoubtedly, the study of hybridization in many Cactaceae taxa will be more accessible by the use of new sequencing technologies. The possibility of generating genomic datasets will allow the investigation of genetic variation across the entire genome, providing deep insights into the hybridization process in Cactaceae [132,149]. Specifically, for the cactus family, we seek to understand many open questions, such as: How permeable to gene flow are species? What is the importance of hybrid speciation for the group? What is the role of reinforcement in improving pre-mating reproductive isolation?

4.6. Genome Assemblies

The studies of functional genomics within Cactaceae evaluated here (excluding studies restricted to structural or genome size estimation from flow cytometry) are clearly associated with the advance and availability of NGS methodologies in the past decade. The species' samplings in these studies were focused on a few tribes, such as Phyllocactaeae, Cereaceae, Cactaceae, Cyliandropuntiae, and Rhipsalideae (Table S3). The amount of nuclear genomic data is still underexplored in the family and focused on North American cacti diversity. Currently, we recovered only three papers describing genomic sequencing and assembly [15,30,150]. From them, we found a chromosome-scale genome assembly of *Hylocereus undatus*, a draft of the whole genome of *Carnegiea gigantea*, four low-coverage genomes of species from North American cacti (*Pachycereus springlei*, *Lophocereus schottii*, *Stenocereus thurberi*, and *Pereskia humboldtii*), and a low-coverage genome from a South American cactus species (*Cereus fernambucensis*). The number of plastomes (ten studies) and transcriptomes (six studies) available in public repositories has been increasing in the last few years (from 2015 to August 2021). This is probably due to the small size of the plastome organelle and gene length, lower need for bioinformatics skills, and lower hardware/software requirements applied for both plastome and transcriptome assembly approaches. This kind of information is of paramount importance, providing genomic resources and background for future studies of cacti lineages, including phylogenetics and phylogeographic studies.

The RNA-Seq method has been used in the family for different purposes (Table S3). For instance, Walker et al. [151] and Wang et al. [152] used this kind of dataset for phylogenomic and evolutionary studies, while other authors tested differential gene expression in cactus species under distinct ecological stress, drought event conditions, RNase-based self-incompatibility, gene expression in root development, and genes under positive selection into specific clades [30,153–158] (Table S3). Thus, the application of transcriptome methodologies to study ecological and environmental stresses may provide new evidence to understand cacti adaptations [30,159].

It is worth noting that several of these papers are from corresponding authors located in the USA or China (Table S3), the latter with non-occurrence of native cacti species. This observation suggests that the availability of NGS technology with a low cost, transcriptome/genome assembly expertise, and lack of functional genomic information within the Cactaceae family seems to be attractive for different research teams around the world. On the other hand, it is widely known that the use of genomic technologies is biased around the

world, especially concerning inequalities in countries' expenditures for research and development, lack of skilled personnel, and low quality/quantity of outsourcing services [160]. Regarding the costs, an interesting review published by McKain et al. [161] reports the approximate values for distinctive genomic strategies commonly used in plant studies.

Although genomics technology has advanced in the last decade, as well as the number of plant genomes available, the number of Cactaceae nuclear genomes does not follow the same pattern. Some methodological challenges might be associated with this observation, such as extraction of high-quality DNA/RNA solutions or genome assembly troubles associated with the level of ploidy, heterozygosity, and high repetitive sequence content, which seems common in Cactaceae [30]. Poor DNA/RNA quality and purity are problems commonly observed in succulent plants since they are rich in mucilage and secondary metabolites [162]. Distinct protocols are currently available for succulent plants, combining the use of commercial kits or the adaptation of existing ones [163]. Actually, in our survey, we observed the use of distinct protocols and distinctive tissues for nucleic acid extractions. Thus, it is important to test different methods and standardize them for particular taxon/tissue/laboratory conditions.

Challenges faced during genome assembly are most related to the use of bioinformatic knowledge, including software and hardware characteristics, to solve biological issues [164]. These features cause ambiguous mapping of sequences, promoting misassembled false genomes and erroneous arrangements [165,166]. The unanchored assembly sequences may create allelic variations that do not necessarily exist, which may be the cause of most of the problems [167]. There are no easy answers for these issues. A good genome assembly methodology practice requires the application of a distinct software assembler and the evaluation of the datasets to be merged into a final genome [168]. An additional challenge for plant genomics is provided for the growing evidence-based on comparative genomes suggesting large structural differences and presence–absence variances of gene content even between individuals of the same species [169]. This observation culminates in the concept of “pan-genomes”, associated with the investigation of multiple individuals to have a reliable reference genome [170], an issue relatively far from the current reality of cactus genomics but which will need to be addressed in the future. New sequencing methods/platforms and bioinformatic algorithms are still in development and in continuous improvement, which may overcome many of these challenges in the near future

5. Concluding Remarks and Future Directions

We report considerable growth in publications associated with evolutionary genetics (Figure 1a) and genome assemblies (Table S3) after the 2000s. These studies have added substantial knowledge on many issues. Despite this, many biases and challenges still remain in our field. For example: How can the challenges involved in new sequencing technologies be overcome? How can large datasets and phenotypic and environmental traits be better integrated and analyzed? How can the taxonomic (Figure 2) and geographical (Figure 3) biases observed in our survey be mitigated?

We believe that many collaborative efforts at different levels should be made with regards to these challenges. First, the results based on new markers have been promoting fundamental advances in both the ability to generate informative data and to reach theoretical advances in evolutionary genetics. Even though PCR-based markers remain informative for cacti research, the application of new methodologies seems to be appropriate to improve the statistical power in the estimation of evolutionary parameters, such as by using coalescent-based methods. Restriction digest-based methods (RAD-Seq and related technologies) are among the candidates to obtain a genome-wide set of loci shared in a sample of individuals [161]. As RAD-Seq requires no prior genomic resources whatsoever, it has become the most widely used genomic approach for high-throughput SNP genotyping in ecological and evolutionary studies of non-model organisms [171] and a possibility of choice for cactus researchers for both phylogenetics [62] and population-level analyses [27]. Reference-guided assemblies to process genome skimming data have also been successfully

applied to resolve phylogenies in cactus species (e.g., [44]). Moreover, the recent description of two panels of orthologous genes specific for the cactus family [172,173] will certainly encourage researchers to generate multi-locus data in the short term. The advantage of such panels is better reproducibility and that they allow a better comparison of different cacti groups and plants.

The challenges to broadly implement NGS technologies are not restricted to technical issues (high-quality DNA/RNA, assemblies, etc.) but also involve other dimensions, such as funding and personal skills. For example, most cactus species naturally grow in developing countries with heterogeneous challenges to research and development (R&D). Indeed, in low-income countries the investments in R&D are in general lower than 1% of gross domestic product (GPD) while this percentage is usually higher than 2% in nations with a global leadership in science [174]. Thus, despite the complexity and many variables of these issues, the cactologist community should be integrated and engaged in pushing governments to find solutions to at least a fraction of these challenges, since many species are under threat [93,94]. Fostering international collaborations may contribute to the transfer of technology and knowledge among researchers of different countries and regions around the globe. For example, the Earth Bio Genome Project (EBP) is an ambitious project aiming to characterize the genomes of all of Earth's eukaryotic biodiversity over a period of 10 years, involving a network of researchers from all continents [175]. Certainly, this type of collaborative project should be encouraged in the microcosmos of cactus research through partnerships between researchers (or funding agencies) of high-, middle-, and low-income countries. In 2021, during the COVID-19 pandemic outbreak, the International virtual mini-symposium "Cactaceae: Phylogenetics, Evolution and Conservation in the Genomic Era" was realized [176]. This event put together an international community of experts allowing a holistic view of the research programs of many research groups. We are confident that this event might be a landmark toward the establishment of an international network of cactus research focused on evolutionary genetics.

International collaborations seem to also be necessary to obtain permits and improve sampling in a few explored areas, with a high diversity of species and endemism, as we have shown substantial taxonomic and geographic biases (Figures 2 and 3). Several clades were poorly explored, including groups that are not recovered in any work, such as the genera *Corryocactus*, *Calymmanthium*, and *Frailea*. The bias in the number of studied species and domains seems to be associated with the number of publications in the countries/regions where these species occur or are distributed. The formation of human resources in such areas would also be imperative to mitigate the historical biases throughout a new generation of researchers interested in cactus diversity.

Finally, to better integrate different source data in predictive analysis [97,98], initiatives toward the establishment of common databases and data availability are necessary. Moreover, it is also important to improve the incentive in the generation of phenotypic, environmental, and life history trait datasets, which are still lacking for most cactus clades. In this sense, collaborative efforts to integrate geneticists, ecologists, botanists, and evolutionary biologists seem to be imperative to understand different dimensions of cactus diversity. Cacti are relevant models to understand evolutionary principles and biogeographic processes, and their study has inspired generations of researchers over time. Mitigating the taxonomic/geographic biases and the implementation of NGS technologies associated with multi-locus and integrative statistical frameworks seems to be necessary steps to better understand this emblematic group of plants.

Supplementary Materials: The supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/genes13030452/s1>. Figure S1: Flow diagrams of the systematic review performed in this review; Figure S2: Distribution of cactus studies by the first author's country; Figure S3: Number of studies using each type of molecular marker; Figure S4: The main biogeographic theories addressed in the evolutionary studies of Cactaceae; Table S1: Summary of reviewed studies on evolutionary genetics of cacti classified as the main topics "Populations genetics" and "Phylogeography"; Table S2: Summary of reviewed studies on evolutionary genetics of cacti classified as the main topic "Phylogeny"; Table S3: Summary of reviewed studies on evolutionary genetics of cacti classified as the main topics "Genome/Transcriptome"; Table S4: Frequently referenced theories in biogeographic studies of cacti and the number of studies associated with each one.

Author Contributions: F.F.F. and E.M.M. conceived the preliminary idea; F.F.F., D.T.A., I.A.S.B., M.R.-B., M.C.T. and E.M.M. equally contributed to performing data collection and writing. All authors made multiple revisions. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP): grant numbers 2020/15161-3, 2018/06937-8; and 2019/11233-2; Coordenação de Aperfeiçoamento de Pessoal de Nível Superior-Brasil (CAPES): Finance Code 001; and National Council for Scientific and Technological Development (CNPq): grant number 303940/2019-0.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Acknowledgments: We thank the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) for a grant (2020/15161-3 to F.F.F.) and two fellowships (2018/06937-8 to M.R. and 2019/11233-2 to M.C.T.). This study was also financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior-Brasil (CAPES)-Finance Code 001 (fellowship to D.T.A.). E.M.M. thanks the National Council for Scientific and Technological Development (CNPq, grant 303940/2019-0) for financial support.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Arakaki, M.; Christin, P.-A.; Nyffeler, R.; Lendel, A.; Eggli, U.; Ogburn, R.M.; Spriggs, E.; Moore, M.J.; Edwards, E.J. Contemporaneous and recent radiations of the world's major succulent plant lineages. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 8379–8384. [[CrossRef](#)] [[PubMed](#)]
2. Hernández-Hernández, T.; Brown, J.; Schlumpberger, B.O.; Eguiarte, L.E.; Magallón, S. Beyond aridification: Multiple explanations for the elevated diversification of cacti in the New World Succulent Biome. *New Phytol.* **2014**, *202*, 1382–1397. [[CrossRef](#)] [[PubMed](#)]
3. Silva, G.A.R.; Antonelli, A.; Lendel, A.; Moraes, E.D.M.; Manfrin, M.H. The impact of early Quaternary climate change on the diversification and population dynamics of a South American cactus species. *J. Biogeogr.* **2017**, *45*, 76–88. [[CrossRef](#)]
4. Hunt, D.R.; Taylor, N.P.; Charles, G. *New Cactus Lexicon*; DH Books: New York, NY, USA, 2006; 900p.
5. Korotkova, N.; Aquino, D.; Arias, S.; Eggli, U.; Franck, A.; Gómez-Hinostrosa, C.; Guerrero, P.C.; Hernández, H.M.; Kohlbecker, A.; Köhler, M.; et al. Cactaceae at Caryo-phyllales. org—a dynamic online species-level taxonomic backbone for the family. *Willdenowia* **2021**, *51*, 251–270. [[CrossRef](#)]
6. Britton, N.L.; Rose, J.N. *The Cactaceae*; Press of Gibson Brothers: Washington, DC, USA, 1920; Volume II, 227p.
7. Gibson, A.C.; Nobel, P.S. *The Cactus Primer*; Harvard University Press: Cambridge, MA, USA, 1986; 286p.
8. Anderson, E.F. *The Cactus Family*; Timber Press: Portland, OR, USA, 2001; 776p.
9. Parker, K.C.; Hamrick, J.L. Genetic diversity and clonal structure in a columnar cactus, *Lophocereus schottii*. *Am. J. Bot.* **1992**, *79*, 86–96. [[CrossRef](#)]
10. Mayer, M.S.; Williams, L.M.; Rebman, J.P. Molecular Evidence for the hybrid origin of *Opuntia prolifera* (Cactaceae). *Madroño* **2000**, *47*, 109–115.
11. Nyffeler, R. Phylogenetic relationships in the cactus family (Cactaceae) based on evidence from trnK/matK and trnL-trnF sequences. *Am. J. Bot.* **2002**, *89*, 312–326. [[CrossRef](#)]

12. Solórzano, S.; Cortés-Palomec, A.C.; Ibarra, A.; Dávila, P.; Oyama, K. Isolation, characterization and cross-amplification of polymorphic microsatellite loci in the threatened endemic *Mammillaria crucigera* (Cactaceae). *Mol. Ecol. Resour.* **2009**, *9*, 156–158. [[CrossRef](#)]
13. Majure, L.C.; Puente, R.; Griffith, M.P.; Judd, W.S.; Soltis, P.S.; Soltis, D.E. Phylogeny of *Opuntia* s.s. (Cactaceae): Clade de-lineation; geographic origins; and reticulate evolution. *Am. J. Bot.* **2012**, *99*, 847–864. [[CrossRef](#)]
14. Köhler, M.; Oakley, L.J.; Font, F.; Las Peñas, M.L.; Majure, L.C. On the continuum of evolution: A putative new hybrid speciation event in *Opuntia* (Cactaceae) between a native and an introduced species in southern South America. *Syst. Biodivers.* **2021**, *19*, 1026–1039. [[CrossRef](#)]
15. Copetti, D.; Búrquez, A.; Bustamante, E.; Charboneau, J.L.M.; Childs, K.; Eguiarte, L.E.; Lee, S.; Liu, T.L.; McMahon, M.M.; Whiteman, N.; et al. Extensive gene tree discordance and hemiplasy shaped the genomes of North American columnar cacti. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 12003–12008. [[CrossRef](#)]
16. Ritz, C.M.; Martins, L.; Mecklenburg, R.; Goremykin, V.; Hellwig, F.H. The molecular phylogeny of *Rebutia* (Cactaceae) and its allies demonstrates the influence of paleogeography on the evolution of South American mountain cacti. *Am. J. Bot.* **2007**, *94*, 1321–1332. [[CrossRef](#)]
17. Realini, M.F.; González, G.E.; Font, F.; Picca, P.I.; Poggio, L.; Gottlieb, A.M. Phylogenetic relationships in *Opuntia* (Cactaceae, Opuntioideae) from southern South America. *Oesterreichische Bot. Z.* **2014**, *301*, 1123–1134. [[CrossRef](#)]
18. Franco, F.F.; Silva, G.A.R.; Moraes, E.M.; Taylor, N.; Zappi, D.C.; Jojima, C.L.; Machado, M.C. Plio-Pleistocene diversification of *Cereus* (Cactaceae, Cereaceae) and closely allied genera. *Bot. J. Linn. Soc.* **2017**, *183*, 199–210. [[CrossRef](#)]
19. Guerrero, P.; Majure, L.C.; Cornejo-Romero, A.; Hernández-Hernández, T. Phylogenetic Relationships and Evolutionary Trends in the Cactus Family. *J. Hered.* **2018**, *110*, 4–21. [[CrossRef](#)] [[PubMed](#)]
20. Cruse-Sanders, J.M.; Parker, K.C.; Friar, E.A.; Huang, D.I.; Mashayekhi, S.; Prince, L.M.; Otero-Arnaiz, A.; Casas, A. Managing diversity: Domestication and gene flow in *Stenocereus stellatus* Riccob. (Cactaceae) in Mexico. *Ecol. Evol.* **2013**, *3*, 1340–1355. [[CrossRef](#)] [[PubMed](#)]
21. Bonatelli, I.A.; Perez, M.F.; Peterson, A.T.; Taylor, N.P.; Zappi, D.C.; Machado, M.C.; Koch, I.; Pires, A.H.C.; Moraes, E.M. Inter-glacial microrefugia and diversification of a cactus species complex: Phylogeography and palaeodistributional reconstructions for *Pilosocereus aurisetus* and allies. *Mol. Ecol.* **2014**, *23*, 3044–3063. [[CrossRef](#)]
22. Silva, G.A.; Jojima, C.L.; Moraes, E.M.; Antonelli, A.; Manfrin, M.H.; Franco, F.F. Intra and interspecific sequence variation in closely related species of *Cereus* (CACTACEAE). *Biochem. Syst. Ecol.* **2016**, *65*, 137–142. [[CrossRef](#)]
23. van Dijk, E.L.; Auger, H.; Jaszczyszyn, Y.; Thermes, C. Ten years of next-generation sequencing technology. *Trends Genet.* **2014**, *30*, 418–426. [[CrossRef](#)]
24. Wetterstrand, K.A. DNA Sequencing Costs: Data from the NHGRI Genome Sequencing Program (GSP). Available online: www.genome.gov/sequencingcostsdata (accessed on 24 November 2021).
25. Avise, J.C. *Molecular Markers; Natural History and Evolution*; Chapman & Hall: Chapel Hill, NC, USA, 1994; 398p.
26. Pinheiro, F.; Dantas-Queiroz, M.V.; Palma-Silva, C. Plant Species Complexes as Models to Understand Speciation and Evolution: A Review of South American Studies. *Crit. Rev. Plant Sci.* **2018**, *37*, 54–80. [[CrossRef](#)]
27. Khan, G.; Franco, F.F.; Silva, G.A.; Bombonato, J.R.; Machado, M.; Alonso, D.P.; Ribolla, P.E.; Albach, D.C.; Moraes, E.M. Maintaining genetic integrity with high promiscuity: Frequent hybridization with low introgression in multiple hybrid zones of *Melocactus* (Cactaceae). *Mol. Phylogenet. Evol.* **2019**, *142*, 106642. [[CrossRef](#)] [[PubMed](#)]
28. Arakaki, M.; Speranza, P.; Soltis, P.S.; Soltis, D.E. Examination of Reticulate Evolution Involving *Haageocereus* and *Espostoa*. *Haseltonia* **2020**, *27*, 102–112. [[CrossRef](#)]
29. Ossa, C.G.; Montenegro, P.; Larridon, I.; Pérez, F. Response of xerophytic plants to glacial cycles in southern South America. *Ann. Bot.* **2019**, *124*, 15–26. [[CrossRef](#)] [[PubMed](#)]
30. Amaral, D.T.; Bombonato, J.R.; Andrade, S.C.D.S.; Moraes, E.M.; Franco, F.F. The genome of a thorny species: Comparative genomic analysis among South and North American Cactaceae. *Planta* **2021**, *254*, 44. [[CrossRef](#)]
31. Juárez-Miranda, A.I.; Cornejo-Romero, A.; Vargas-Mendoza, C.F. Population expansion and genetic structure in *Cephalocereus nizandensis* (Cactaceae), a microendemic cactus of rocky outcrops of the Tehuantepec basin, Mexico. *Plant Ecol. Evol.* **2021**, *154*, 217–230. [[CrossRef](#)]
32. Majure, L.C.; Barrios, D.; Díaz, E.; Zumwalde, B.A.; Testo, W.; Negrón-Ortíz, V. Pleistocene aridification underlies the evolutionary history of the Caribbean endemic, insular, giant *Consolea* (Opuntioideae). *Am. J. Bot.* **2021**, *108*, 200–215. [[CrossRef](#)]
33. Morrone, J.J. Biogeographical regionalisation of the Neotropical region. *Zootaxa* **2014**, *3782*, 1–110. [[CrossRef](#)]
34. Page, M.J.; McKenzie, J.E.; Bossuyt, P.M.; Boutron, I.; Hoffmann, T.C.; Mulrow, C.D.; Shamseer, L.; Tetzlaff, J.M.; Akl, E.A.; Brennan, S.E. The PRISMA 2020 state-ment: An updated guideline for reporting systematic reviews. *BMJ* **2021**, *88*, 105906. [[CrossRef](#)]
35. Schlötterer, C. The evolution of molecular markers—just a matter of fashion? *Nat. Rev. Genet.* **2004**, *5*, 63–69. [[CrossRef](#)]
36. Wolfe, K.H.; Li, W.-H.; Sharp, P.M. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proc. Natl. Acad. Sci. USA* **1987**, *84*, 9054–9058. [[CrossRef](#)]
37. Zhang, D.-X.; Hewitt, G.M. Nuclear DNA analyses in genetic studies of populations: Practice, problems and prospects. *Mol. Ecol.* **2003**, *12*, 563–584. [[CrossRef](#)] [[PubMed](#)]

38. Shaw, J.; Lickey, E.B.; Schilling, E.E.; Small, R.L. Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: The tortoise and the hare III. *Am. J. Bot.* **2007**, *94*, 275–288. [[CrossRef](#)] [[PubMed](#)]
39. Hernández- Hernández, T.; Hernández, H.M.; De-Nova, J.A.; Puente, R.; Eguiarte, L.E.; Magallón, S. Phylogenetic relationships and evolution of growth form in Cactaceae (Caryophyllales; Eudicotyledoneae). *Am. J. Bot.* **2011**, *98*, 44–61. [[CrossRef](#)] [[PubMed](#)]
40. Calvente, A.; Moraes, E.M.; Lavor, P.; Bonatelli, I.A.; Nacaguma, P.; Versieux, L.M.; Taylor, N.P.; Zappi, D.C. Phylogenetic analyses of *Pilosocereus* (Cactaceae) inferred from plastid and nuclear sequences. *Bot. J. Linn. Soc.* **2017**, *183*, 25–38.
41. Duchene, D.; Bromham, L. Rates of molecular evolution and diversification in plants: Chloroplast substitution rates correlate with species-richness in the Proteaceae. *BMC Evol. Biol.* **2013**, *13*, 65. [[CrossRef](#)]
42. Romeiro-Brito, M.; Moraes, E.M.; Taylor, N.P.; Zappi, D.C.; Franco, F.F. Lineage-specific evolutionary rate in plants: Contributions of a screening for *Cereus* (Cactaceae). *Appl. Plant Sci.* **2016**, *4*, 1500074. [[CrossRef](#)]
43. Majure, L.C.; Baker, M.A.; Cloud-Hughes, M.; Salywon, A.; Neubig, K.M. Phylogenomics in Cactaceae: A case study using the chollas sensu lato (Cylindropuntiaeae, Opuntioideae) reveals a common pattern out of the Chihuahuan and Sonoran deserts. *Am. J. Bot.* **2019**, *106*, 1327–1345. [[CrossRef](#)]
44. Köhler, M.; Reginato, M.; Souza-Chies, T.T.; Majure, L.C. Insights into Chloroplast Genome Evolution Across Opuntioideae (Cactaceae) Reveals Robust Yet Sometimes Conflicting Phylogenetic Topologies. *Front. Plant Sci.* **2020**, *11*, 729. [[CrossRef](#)]
45. Harpke, D.; Peterson, A. Non-concerted ITS evolution in *Mammillaria* (Cactaceae). *Mol. Phylogenet. Evol.* **2006**, *41*, 579–593. [[CrossRef](#)]
46. Harpke, D.; Peterson, A. Extensive 5.8S nrDNA polymorphism in *Mammillaria* (Cactaceae) with special reference to the identification of pseudogenic internal transcribed spacer regions. *J. Plant Res.* **2008**, *121*, 261–270. [[CrossRef](#)]
47. Small, R.L.; Cronn, R.C.; Wendel, J.F. Use of nuclear genes for phylogeny reconstruction in plants. *Aust. Syst. Bot.* **2004**, *17*, 145–170. [[CrossRef](#)]
48. Khan, G.; Godoy, M.O.; Franco, F.F.; Perez, M.; Taylor, N.P.; Zappi, D.; Machado, M.C.; Moraes, E.M. Extreme population subdivision or cryptic speciation in the cactus *Pilosocereus jauruensis*? A taxonomic challenge posed by a naturally fragmented system. *Syst. Biodivers.* **2017**, *16*, 188–199. [[CrossRef](#)]
49. Khan, G.; Ribeiro, P.M.; Bonatelli, I.A.S.; Perez, M.F.; Franco, F.F.; Moraes, E.M. Weak population structure and no genetic erosion in *Pilosocereus aureispinus*: A microendemic and threatened cactus species from eastern Brazil. *PLoS ONE* **2018**, *13*, e0195475. [[CrossRef](#)] [[PubMed](#)]
50. Perez, M.F.; Carstens, B.C.; Rodrigues, G.L.; Moraes, E.M. Anonymous nuclear markers reveal taxonomic incongruence and long-term disjunction in a cactus species complex with continental-island distribution in South America. *Mol. Phylogenet. Evol.* **2016**, *95*, 11–19. [[CrossRef](#)]
51. Fraga, D.A.; de Carvalho, A.F.; Santana, R.S.; Machado, M.C.; Lacorte, G.A. Development of microsatellite markers for the threatened species *Coleocephalocereus purpureus* (Cactaceae) using next-generation sequencing. *Mol. Biol. Rep.* **2019**, *47*, 1485–1489. [[CrossRef](#)]
52. Gutiérrez, A.V.; Filippi, C.V.; Aguirre, N.C.; Puebla, A.F.; Acuña, C.V.; Taboada, G.M.; Ortega-Baes, F.P. Development of novel SSR molecular markers using a Next-Generation Sequencing approach (ddRADseq) in *Stetsonia coryne* (Cactaceae). *An. Acad. Bras. Ciências* **2021**, *93*. [[CrossRef](#)]
53. Kalia, R.K.; Rai, M.K.; Kalia, S.; Singh, R.; Dhawan, A.K. Microsatellite markers: An overview of the recent progress in plants. *Euphytica* **2010**, *177*, 309–334. [[CrossRef](#)]
54. Idrees, M.; Irshad, M. Molecular markers in plants for analysis of genetic diversity: A review. *Eur. Acad. Res.* **2014**, *2*, 1513–1540.
55. Taheri, S.; Abdullah, T.L.; Yusop, M.R.; Hanafi, M.M.; Sahebi, M.; Azizi, P.; Shamshiri, R.R. Mining and Development of Novel SSR Markers Using Next Generation Sequencing (NGS) Data in Plants. *Molecules* **2018**, *23*, 399. [[CrossRef](#)]
56. Moraes, E.M.; Perez, M.F.; Téó, M.F.; Zappi, D.C.; Taylor, N.P.; Machado, M.C. Cross-species amplification of microsatellites reveals incongruence in the molecular variation and taxonomic limits of the *Pilosocereus aurisetus* group (Cactaceae). *Genetica* **2012**, *140*, 277–285. [[CrossRef](#)]
57. Fernandes, V.N.D.A.; das Neves, A.F.; Martin, P.G.; Mangolin, C.A.; Machado, M.D.F.P. Genetic structure and molecular divergence among samples of mandacaru (*Cereus* spp.; Cactaceae) as revealed by microsatellite markers. *Biochem. Syst. Ecol.* **2016**, *64*, 38–45. [[CrossRef](#)]
58. Bombonato, J.R.; Bonatelli, I.A.S.; Silva, G.A.R.; Moraes, E.M.; Zappi, D.C.; Taylor, N.P.; Franco, F.F. Cross-genera SSR transferability in cacti revealed by a case study using *Cereus* (Cereaceae; Cactaceae). *Genet. Mol. Biol.* **2019**, *42*, 87–94. [[CrossRef](#)] [[PubMed](#)]
59. Fava, W.S.; Gomes, V.G.N.; Lorenz, A.P.; Paggi, G.M. Cross-amplification of microsatellite loci in the cacti species from Brazilian Chaco. *Mol. Biol. Rep.* **2020**, *47*, 1535–1542. [[CrossRef](#)] [[PubMed](#)]
60. Franco, F.F.; Silva, F.A.; Khan, G.; Bonatelli, I.A.; Amaral, D.T.; Zappi, D.C.; Taylor, N.P.; Moraes, E.M. Generalizations of genetic conservation principles in islands are not always likely: A case study from a Neotropical insular cactus. *Bot. J. Linn. Soc.* **2021**, boab076. [[CrossRef](#)]
61. Bonatelli, I.A.D.S.; Carstens, B.C.; Moraes, E. Using Next Generation RAD Sequencing to Isolate Multispecies Microsatellites for *Pilosocereus* (Cactaceae). *PLoS ONE* **2015**, *10*, e0142602. [[CrossRef](#)] [[PubMed](#)]

62. Bombonato, J.R.; Amaral, D.T.; Silva, G.A.R.; Khan, G.; Moraes, E.M.; Andrade, S.C.; Eaton, D.A.R.; Alonso, D.P.; Ribolla, P.E.M.; Taylor, N.; et al. The potential of genome-wide RAD sequences for resolving rapid radiations: A case study in Cactaceae. *Mol. Phylogenet. Evol.* **2020**, *151*, 106896. [[CrossRef](#)]
63. Merklinger, F.F.; Böhnert, T.; Arakaki, M.; Weigend, M.; Quandt, D.; Luebert, F. Quaternary diversification of a columnar cactus in the driest place on earth. *Am. J. Bot.* **2021**, *108*, 184–199. [[CrossRef](#)]
64. Dodsworth, S.; Pokorný, L.; Johnson, M.G.; Kim, J.T.; Maurin, O.; Wickett, N.J.; Forest, F.; Baker, W.J. Hyb-Seq for Flowering Plant Systematics. *Trends Plant Sci.* **2019**, *24*, 887–891. [[CrossRef](#)]
65. Pelletier, T.A.; Carstens, B.C. Geographical range size and latitude predict population genetic structure in a global survey. *Biol. Lett.* **2018**, *14*, 20170566. [[CrossRef](#)]
66. Nazareno, A.G.; Bemmels, J.B.; Dick, C.W.; Lohmann, L.G. Minimum sample sizes for population genomics: An empirical study from an Amazonian plant species. *Mol. Ecol. Resour.* **2017**, *17*, 1136–1147. [[CrossRef](#)]
67. Silva, G.A.R.; Khan, G.; Ribeiro-Silva, S.; Aona, L.; Machado, M.C.; Bonatelli, I.A.S.; Moraes, E.M. Extreme genetic structure in a relict cactus genus from campo rupestre landscapes: Implications for conservation. *Biodivers. Conserv.* **2020**, *29*, 1263–1281. [[CrossRef](#)]
68. Moore, A.J.; de Vos, J.M.; Hancock, L.P.; Goolsby, E.; Edwards, E.J. Targeted Enrichment of Large Gene Families for Phylogenetic Inference: Phylogeny and Molecular Evolution of Photosynthesis Genes in the Portulugo Clade (Caryophyllales). *Syst. Biol.* **2017**, *67*, 367–383. [[CrossRef](#)] [[PubMed](#)]
69. McCormack, J.E.; Harvey, M.G.; Faircloth, B.C.; Crawford, N.G.; Glenn, T.C.; Brumfield, R.T. Data from: A phylogeny of birds based on over 1500 loci collected by target enrichment and high-throughput sequencing. *Dryad Dataset* **2013**, *8*, e54848. [[CrossRef](#)]
70. Bravo, G.A.; Antonelli, A.; Bacon, C.D.; Bartoszek, K.; Blom, M.P.K.; Huynh, S.; Jones, G.; Knowles, L.L.; Lamichhaney, S.; Marcussen, T.; et al. Embracing heterogeneity: Coalescing the Tree of Life and the future of phylogenomics. *PeerJ* **2019**, *7*, e6399. [[CrossRef](#)] [[PubMed](#)]
71. Degnan, J.H.; Rosenberg, N. Discordance of Species Trees with Their Most Likely Gene Trees. *PLoS Genet.* **2006**, *2*, e68. [[CrossRef](#)]
72. Edwards, S. Is a new and general theory of molecular systematics emerging? *Evolution* **2009**, *63*, 1–19. [[CrossRef](#)]
73. Zhang, C.; Rabiee, M.; Sayyari, E.; Mirarab, S. ASTRAL-III: Polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinform.* **2018**, *19*, 15–30. [[CrossRef](#)]
74. Chifman, J.; Kubatko, L. Quartet Inference from SNP Data Under the Coalescent Model. *Bioinformatics* **2014**, *30*, 3317–3324. [[CrossRef](#)]
75. Ogilvie, H.A.; Bouckaert, R.R.; Drummond, A. StarBEAST2 Brings Faster Species Tree Inference and Accurate Estimates of Substitution Rates. *Mol. Biol. Evol.* **2017**, *34*, 2101–2114. [[CrossRef](#)]
76. Bryant, D.; Bouckaert, R.; Felsenstein, J.; Rosenberg, N.A.; RoyChoudhury, A. Inferring Species Trees Directly from Biallelic Genetic Markers: Bypassing Gene Trees in a Full Coalescent Analysis. *Mol. Biol. Evol.* **2012**, *29*, 1917–1932. [[CrossRef](#)]
77. Mantooth, S.J.; Riddle, B.R. Molecular Biogeography: The Intersection between Geographic and Molecular Variation. *Geogr. Compass* **2011**, *5*, 1–20. [[CrossRef](#)]
78. Crisp, M.D.; Trewick, S.A.; Cook, L. Hypothesis testing in biogeography. *Trends Ecol. Evol.* **2011**, *26*, 66–72. [[CrossRef](#)] [[PubMed](#)]
79. Stevens, R.D.; Weber, M.M.; Villalobos, F. Contemporary Biogeography. In *Phyllostomid Bats*; Fleming, T.H., Dávalos, L.M., Mello, M.A.R., Eds.; University of Chicago Press: Chicago, IL, USA, 2020; pp. 391–410.
80. Ornelas, J.F.; Rodriguez-Gomez, F. Influence of Pleistocene Glacial/Interglacial Cycles on the Genetic Structure of the Mistletoe Cactus *Rhypsalis baccifera* (Cactaceae) in Mesoamerica. *J. Hered.* **2015**, *106*, 196–210. [[CrossRef](#)] [[PubMed](#)]
81. Franco, F.F.; Jojima, C.L.; Perez, M.F.; Zappi, D.C.; Taylor, N.; Moraes, E.M. The xeric side of the Brazilian Atlantic Forest: The forces shaping phylogeographic structure of cacti. *Ecol. Evol.* **2017**, *7*, 9281–9293. [[CrossRef](#)] [[PubMed](#)]
82. Nason, J.D.; Hamrick, J.L.; Fleming, T.H. Historical vicariance and postglacial colonization effects on the evolution of genetic structure in *Lophocereus*, a Sonoran Desert columnar cactus. *Evolution* **2002**, *56*, 2214–2226. [[CrossRef](#)]
83. Gutiérrez-Flores, C.; García-De-León, F.J.; Cota-Sánchez, J.H. Microsatellite genetic diversity and mating systems in the columnar cactus *Pachycereus pringlei* (Cactaceae). *Perspect. Plant Ecol. Evol. Syst.* **2016**, *22*, 1–10.
84. Aquino, D.; Moreno-Letelier, A.; González-Botello, M.A.; Arias, S. The importance of environmental conditions in maintaining lineage identity in *Epithelantha* (Cactaceae). *Ecol. Evol.* **2021**, *11*, 4520–4531. [[CrossRef](#)]
85. de Menezes, M.O.T.; Zappi, D.; Moraes, E.M.; Franco, F.F.; Taylor, N.P.; Costa, I.; Loiola, M.I. Pleistocene radiation of coastal species of *Pilosocereus* (Cactaceae) in eastern Brazil. *J. Arid Environ.* **2016**, *135*, 22–32. [[CrossRef](#)]
86. Rull, V. Neotropical Diversification: Historical Overview and Conceptual Insights. In *Neotropical Diversification: Patterns and Processes*; Springer: Cham, Switzerland, 2020; pp. 13–49. [[CrossRef](#)]
87. Zizka, A.; Carvalho-Sobrinho, J.G.; Pennington, R.T.; Queiroz, L.P.; Alcantara, S.; Baum, D.A.; Bacon, C.D.; Antonelli, A. Transitions between biomes are common and directional in Bombacoideae (Malvaceae). *J. Biogeogr.* **2020**, *47*, 1310–1321. [[CrossRef](#)]
88. 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.R.; Franco, F.F. Tracking the xeric biomes of South America: The spatiotemporal diversification of Mandacaru cactus. *J. Biogeogr.* **2021**, *48*, 3085–3103. [[CrossRef](#)]
89. Liu, Y.; Fu, B.; Wang, S.; Zhao, W. Global ecological regionalization: From biogeography to ecosystem services. *Curr. Opin. Environ. Sustain.* **2018**, *33*, 1–8. [[CrossRef](#)]

90. Mucina, L. Biome: Evolution of a crucial ecological and biogeographical concept. *New Phytol.* **2018**, *222*, 97–114. [[CrossRef](#)] [[PubMed](#)]
91. Hunter, J.; Franklin, S.; Luxton, S.; Loidi, J. Terrestrial biomes: A conceptual review. *Veg. Classif. Surv.* **2021**, *2*, 73–85. [[CrossRef](#)]
92. Maestre, F.T.; Benito, B.M.; Berdugo, M.; Concostrina-Zubiri, L.; Delgado-Baquerizo, M.; Eldridge, D.J.; Guirado, E.; Gross, N.; Kéfi, S.; le Bagousse-Pinguet, Y.; et al. Biogeography of global drylands. *New Phytol.* **2021**, *231*. [[CrossRef](#)]
93. Goettsch, B.; Hilton-Taylor, C.; Cruz-Piñón, G.; Duffy, J.P.; Frances, A.; Hernández, H.M.; Inger, R.; Pollock, C.; Schipper, J.; Superina, M.; et al. High proportion of cactus species threatened with extinction. *Nat. Plants* **2015**, *1*, 15142. [[CrossRef](#)] [[PubMed](#)]
94. Goettsch, B.; Durán, A.P.; Gaston, K.J. Global gap analysis of cactus species and priority sites for their conservation. *Conserv. Biol.* **2018**, *33*, 369–376. [[CrossRef](#)] [[PubMed](#)]
95. Saco, P.M.; Heras, M.M.-D.L.; Keesstra, S.; Baartman, J.; Yetemen, O.; Rodríguez, J.F. Vegetation and soil degradation in drylands: Non linear feedbacks and early warning signals. *Curr. Opin. Environ. Sci. Health* **2018**, *5*, 67–72. [[CrossRef](#)]
96. Fontenele, R.S.; Köhler, M.; Majure, L.C.; Avalos-Calleros, J.A.; Argüello-Astorga, G.R.; Font, F.; Vidal, A.H.; Roumagnac, P.; Kraberger, S.; Martin, D.P.; et al. Novel circular DNA virus identified in *Opuntia discolor* (Cactaceae) that codes for proteins with similarity to those of geminiviruses. *J. Gen. Virol.* **2021**, *102*, 001671. [[CrossRef](#)]
97. Papadopoulou, A.; Knowles, L.L. Toward a paradigm shift in comparative phylogeography driven by trait-based hypotheses. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 8018–8024. [[CrossRef](#)]
98. Zamudio, K.R.; Bell, R.C.; Mason, N.A. Phenotypes in phylogeography: Species' traits, environmental variation, and vertebrate diversification. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 8041–8048. [[CrossRef](#)]
99. Sukumaran, J.; Knowles, L.L. Trait-Dependent Biogeography: (Re)Integrating Biology into Probabilistic Historical Biogeographical Models. *Trends Ecol. Evol.* **2018**, *33*, 390–398. [[CrossRef](#)]
100. Espíndola, A.; Ruffley, M.; Smith, M.L.; Carstens, B.C.; Tank, D.C.; Sullivan, J. Identifying cryptic diversity with predictive phylogeography. *Proc. R. Soc. B Biol. Sci.* **2016**, *283*, 20161529. [[CrossRef](#)] [[PubMed](#)]
101. Sullivan, J.; Smith, M.L.; Espíndola, A.; Ruffley, M.; Rankin, A.; Tank, D.; Carstens, B. Integrating life history traits into pre-dictive phylogeography. *Mol. Ecol.* **2019**, *28*, 2062–2073. [[CrossRef](#)] [[PubMed](#)]
102. GBIF: The Global Biodiversity Information Facility (2022) What Is GBIF? Available online: <https://www.gbif.org/what-is-gbif> (accessed on 13 January 2022).
103. Fick, S.E.; Hijmans, R.J. WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.* **2017**, *37*, 4302–4315. [[CrossRef](#)]
104. Kattge, J.; Diaz, S.; Lavorel, S.; Prentice, I.C.; Leadley, P.; Bönsch, G.; Garnier, E.; Westoby, M.; Reich, P.B.; Wright, I.J.; et al. TRY—A global database of plant traits. *Glob. Chang. Biol.* **2011**, *17*, 2905–2935. [[CrossRef](#)]
105. Paz, A.; Ibáñez, R.; Lips, K.R.; Crawford, A.J. Testing the role of ecology and life history in structuring genetic variation across a landscape: A trait-based phylogeographic approach. *Mol. Ecol.* **2015**, *24*, 3723–3737. [[CrossRef](#)]
106. Riddle, B.R. Comparative phylogeography clarifies the complexity and problems of continental distribution that drove AR Wallace to favor islands. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 7970–7977. [[CrossRef](#)]
107. Perez, M.F.; Bonatelli, I.A.S.; Romeiro-Brito, M.; Franco, F.F.; Taylor, N.P.; Zappi, D.C.; Moraes, E.M. Coalescent-based species delimitation meets deep learning: Insights from a highly fragmented cactus system. *Mol. Ecol. Resour.* **2021**. [[CrossRef](#)]
108. Yesson, C.; Bárcenas, R.T.; Hernandez, H.M.; Ruiz-Maqueda, M.D.L.L.; Prado, A.; Rodríguez, V.M.; Hawkins, J.A. DNA barcodes for Mexican Cactaceae, plants under pressure from wild collecting. *Mol. Ecol. Resour.* **2011**, *11*, 775–783. [[CrossRef](#)]
109. CBOL Plant Working Group. A DNA barcode for land plants. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 12794–12797. [[CrossRef](#)]
110. Baker, M.A.; Butterworth, C.A. Geographic distribution and taxonomic circumscription of populations within *Coryphantha section Robustispina* (Cactaceae). *Am. J. Bot.* **2013**, *100*, 984–997. [[CrossRef](#)]
111. Ritz, C.M.; Fickenscher, K.; Föller, J.; Herrmann, K.; Mecklenburg, R.; Wahl, R. Molecular phylogenetic relationships of the Andean genus *Aylostera* Speg. (Cactaceae; Trichocereae); a new classification and a morphological identification key. *Plant Syst. Evol.* **2016**, *302*, 763–780. [[CrossRef](#)]
112. Larridon, I.; Veltjen, E.; Semmouri, I.; Asselman, P.; Guerrero, P.; Duarte, M.; Walter, H.E.; Cisternas, M.A.; Samain, M.-S. Investigating taxon boundaries and extinction risk in endemic Chilean cacti (*Copiapoa subseccion Cinerei*, Cactaceae) using chloroplast DNA sequences, microsatellite data and 3D mapping. *Kew Bull.* **2018**, *73*, 55. [[CrossRef](#)]
113. Majure, L.C.; Encarnación, Y.; Clase, T.; Peguero, B.; Ho, K.; Barrios, D. Phylogenetics of *Leptocereus* (Cactaceae) on Hispaniola: Clarifying species limits in the *L. weingartianus* complex and a new species from the Sierra de Bahoruco. *PhytoKeys* **2021**, *172*, 17–37. [[CrossRef](#)]
114. Donoghue, M.J. A Critique of the Biological Species Concept and Recommendations for a Phylogenetic Alternative. *Bryologist* **1985**, *88*, 172. [[CrossRef](#)]
115. Cracraft, J. Species Concepts and Speciation Analysis. In *Current Ornithology*; Springer: New York, NY, USA, 1983; pp. 159–187. [[CrossRef](#)]
116. Mallet, J. A species definition for the Modern Synthesis. *Trends Ecol. Evol.* **1995**, *10*, 294–299. [[CrossRef](#)]
117. Aquino, D.; Cervantes, R.C.; Gernandt, D.S.; Arias, S. Species Delimitation and Phylogeny of *Epithelantha* (Cactaceae). *Syst. Bot.* **2019**, *44*, 600–615. [[CrossRef](#)]
118. Alvarado-Sizzo, H.; Casas, A.; Parra, F.; Arreola-Nava, H.J.; Terrazas, T.; Sánchez, C. Species delimitation in the *Stenocereus griseus* (Cactaceae) species complex reveals a new species, *S. huastecorum*. *PLoS ONE* **2018**, *13*, e0190385. [[CrossRef](#)]

119. Zappi, D.C. *Pilosocereus (Cactaceae): The Genus in Brazil*; Royal Botanic Gardens: Kew, UK, 1994; 135p.
120. Braun, P.J.; Esteves, E.E. Cactaceae. *Pilosocereus densivillosus*. *Kakt. Sukk* **1994**, *45*, 108–114.
121. Braun, P.J.; Esteves, E.; Esteves, R. *Pilosocereus jauruensis* subsp. *cinnopetalus* (Cactaceae), eine neue Sippe aus dem mittleren Westen Brasiliens. *Kakt. Suk* **2020**, *71*, 73–77.
122. De Queiroz, K. The general lineage concept of species, species criteria, and the process of speciation: A conceptual unification and terminological recommendations. In *Endless Forms: Species Population Genetic Structure of Pilosocereus jauruensis 197 and Speciation*; Howard, D.J., Berlocher, S.H., Eds.; Oxford University Press: New York, NY, USA, 1998; pp. 57–75.
123. Avise, J.C.; Arnold, J.; Ball, R.M.; Bermingham, E.; Lamb, T.; Neigel, J.E.; A Reeb, C.; Saunders, N.C. INTRASPECIFIC PHYLOGEOGRAPHY: The Mitochondrial DNA Bridge Between Population Genetics and Systematics. *Annu. Rev. Ecol. Syst.* **1987**, *18*, 489–522. [[CrossRef](#)]
124. Carstens, B.C.; Pelletier, T.A.; Reid, N.M.; Satler, J.D. How to fail at species delimitation? *Mol. Ecol.* **2013**, *22*, 4369–4383. [[CrossRef](#)]
125. Pons, J.; Barraclough, T.; Gómez-Zurita, J.; Cardoso, A.; Duran, D.P.; Hazell, S.; Kamoun, S.; Sumlin, W.D.; Vogler, A.P. Sequence-Based Species Delimitation for the DNA Taxonomy of Undescribed Insects. *Syst. Biol.* **2006**, *55*, 595–609. [[CrossRef](#)]
126. Yang, Z.; Rannala, B. Unguided Species Delimitation Using DNA Sequence Data from Multiple Loci. *Mol. Biol. Evol.* **2014**, *31*, 3125–3135. [[CrossRef](#)]
127. Pei, J.; Chu, C.; Li, X.; Lu, B.; Wu, Y. CLADES: A classification-based machine learning method for species delimitation from population genetic data. *Mol. Ecol. Resour.* **2018**, *18*, 1144–1156. [[CrossRef](#)]
128. Derkarabetian, S.; Castillo, S.; Koo, P.K.; Ovchinnikov, S.; Hedin, M. A demonstration of unsupervised machine learning in species delimitation. *Mol. Phylogenet. Evol.* **2019**, *139*, 106562. [[CrossRef](#)]
129. Arnold, M.L. *Natural Hybridization and Evolution*; Oxford Series in Ecology and Evolution; Oxford University Press: New York, NY, USA, 1997; 232p.
130. Rieseberg, L.H. Hybrid origins of plant species. *Annu. Rev. Ecol. Syst.* **1997**, *28*, 359–389. [[CrossRef](#)]
131. Mallet, J.; Beltran, M.; Neukirchen, W.; Linares, M. Natural hybridization in heliconiine butterflies: The species boundary as a continuum. *BMC Evol. Biol.* **2007**, *7*, 28. [[CrossRef](#)]
132. Schley, R.J.; Twyford, A.D.; Pennington, R.T. Hybridization: A ‘double-edged sword’ for Neotropical plant diversity. *Bot. J. Linn. Soc.* **2021**, boab070. [[CrossRef](#)]
133. Friedrich, H. Hybridization as a factor in the evolution of the Cactaceae. *Cactus Succul. J.* **1974**, *46*, 213–214.
134. Machado, M.C. What is the role of hybridization in the evolution of the Cactaceae? *Bradleya* **2008**, *26*, 1–18. [[CrossRef](#)]
135. Pritchard, J.K.; Stephens, M.; Donnelly, P. Inference of population structure using multilocus genotype data. *Genetics* **2000**, *155*, 945–959. [[CrossRef](#)]
136. Anderson, E.C.; Thompson, E.A. A Model-Based Method for Identifying Species Hybrids Using Multilocus Genetic Data. *Genetics* **2002**, *160*, 1217–1229. [[CrossRef](#)]
137. Beerli, P.; Felsenstein, J. Maximum likelihood estimation of a migration matrix and effective population sizes in n subpopulations by using a coalescent approach. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 4563–4568. [[CrossRef](#)]
138. Gronau, I.; Hubisz, M.J.; Gulko, B.; Danko, C.G.; Siepel, A. Bayesian inference of ancient human demography from individual genome sequences. *Nat. Genet.* **2011**, *43*, 1031–1034. [[CrossRef](#)]
139. Gompert, Z.; Buerkle, C.A. introgress: A software package for mapping components of isolation in hybrids. *Mol. Ecol. Resour.* **2010**, *10*, 378–384. [[CrossRef](#)]
140. Solís-Lemus, C.; Bastide, P.; Ané, C. PhyloNetworks: A Package for Phylogenetic Networks. *Mol. Biol. Evol.* **2017**, *34*, 3292–3298. [[CrossRef](#)]
141. Griffith, M.P. Using molecular data to elucidate reticulate evolution in *Opuntia*. *Madroño* **2003**, *50*, 162–169.
142. Lambert, S.M.; Borba, E.L.; Machado, M.C.; Andrade, S.C.D.S. Allozyme Diversity and Morphometrics of *Melocactus paucispinus* (Cactaceae) and Evidence for Hybridization with *M. concinnus* in the Chapada Diamantina, North-eastern Brazil. *Ann. Bot.* **2006**, *97*, 389–403. [[CrossRef](#)]
143. Rayamajhi, N.; Sharma, J. Genetic diversity and structure of a rare endemic cactus and an assessment of its genetic relationship with a more common congener. *Genetica* **2018**, *146*, 329–340. [[CrossRef](#)]
144. Granados-Aguilar, X.; Mendoza, C.G.; Cervantes, C.R.; Montes, J.R.; Arias, S. Unraveling Reticulate Evolution in *Opuntia* (Cactaceae) From Southern Mexico. *Front. Plant Sci.* **2021**, *11*. [[CrossRef](#)]
145. Schwabe, A.L.; Neale, J.R.; McGlaughlin, M.E. Examining the genetic integrity of a rare endemic Colorado cactus (*Sclerocactus glaucus*) in the face of hybridization threats from a close and widespread congener (*Sclerocactus parviflorus*). *Conserv. Genet.* **2014**, *16*, 443–457. [[CrossRef](#)]
146. Plume, O.; Straub, S.C.K.; Tel-Zur, N.; Cisneros, A.; Schneider, B.; Doyle, J.J. Testing a Hypothesis of Intergeneric Allopolyploidy in Vine Cacti (Cactaceae: Hylocereeae). *Syst. Bot.* **2013**, *38*, 737–751. [[CrossRef](#)]
147. Majure, L.; Puente-Martinez, R. Phylogenetic relationships and morphological evolution in *Opuntia* s.str. and closely related members of tribe Opuntieae. *Succul. Plant Res.* **2014**, *8*, 9–30.
148. Arakaki, M.; Soltis, D.E.; Soltis, P.S.; Speranza, P.R. Characterization of polymorphic microsatellite loci in *Haageocereus* (Trichocereae, Cactaceae). *Am. J. Bot.* **2010**, *97*, e17–e19. [[CrossRef](#)]
149. Twyford, A.D.; Ennos, R.A. Next-generation hybridization and introgression. *Heredity* **2011**, *108*, 179–189. [[CrossRef](#)] [[PubMed](#)]

150. Chen, J.-Y.; Xie, F.-F.; Cui, Y.-Z.; Chen, C.-B.; Lu, W.-J.; Hu, X.-D.; Hua, Q.-Z.; Zhao, J.; Wu, Z.-J.; Gao, D.; et al. A chromosome-scale genome sequence of pitaya (*Hylocereus undatus*) provides novel insights into the genome evolution and regulation of betalain biosynthesis. *Hortic. Res.* **2021**, *8*, 1–15. [[CrossRef](#)] [[PubMed](#)]
151. Walker, J.F.; Yang, Y.; Feng, T.; Timoneda, A.; Mikenas, J.; Hutchison, V.; Edwards, C.; Wang, N.; Ahluwalia, S.; Olivieri, J.; et al. From cacti to carnivores: Improved phylotranscriptomic sampling and hierarchical homology inference provide further insight into the evolution of Caryophyllales. *Am. J. Bot.* **2018**, *105*, 446–462. [[CrossRef](#)] [[PubMed](#)]
152. Wang, N.; Yang, Y.; Moore, M.J.; Brockington, S.F.; Walker, J.F.; Brown, J.W.; Liang, B.; Feng, T.; Edwards, C.; Mikenas, J.; et al. Evolution of Portulacineae Marked by Gene Tree Conflict and Gene Family Expansion Associated with Adaptation to Harsh Environments. *Mol. Biol. Evol.* **2018**, *36*, 112–126. [[CrossRef](#)]
153. Rodríguez-Alonso, G.; Matvienko, M.; López-Valle, M.L.; Lázaro-Mixteco, P.E.; Napsucialy-Mendivil, S.; Dubrovsky, J.G.; Shishkova, S. Transcriptomics insights into the genetic regulation of root apical meristem exhaustion and determinate primary root growth in *Pachycereus pringlei* (Cactaceae). *Sci. Rep.* **2018**, *8*, 8529. [[CrossRef](#)]
154. Yang, Y.; Moore, M.J.; Brockington, S.F.; Mikenas, J.; Olivieri, J.; Walker, J.F.; Smith, S.A. Improved transcriptome sampling pinpoints 26 ancient and more recent polyploidy events in Caryophyllales, including two allopolyploidy events. *New Phytol.* **2017**, *217*, 855–870. [[CrossRef](#)] [[PubMed](#)]
155. Nong, Q.; Zhang, M.; Chen, J.; Zhang, M.; Cheng, H.; Jian, S.; Lu, H.; Xia, K. RNA-Seq De Novo Assembly of Red Pitaya (*Hylocereus polyrhizus*) Roots and Differential Transcriptome Analysis in Response to Salt Stress. *Trop. Plant Biol.* **2019**, *12*, 55–66. [[CrossRef](#)]
156. Zhou, Z.; Gao, H.; Ming, J.; Ding, Z.; Lin, X.; Zhan, R. Combined Transcriptome and Metabolome analysis of Pitaya fruit unveiled the mechanisms underlying Peel and pulp color formation. *BMC Genom.* **2020**, *21*, 734. [[CrossRef](#)] [[PubMed](#)]
157. Wu, Y.; Xu, J.; Han, X.; Qiao, G.; Yang, K.; Wen, Z.; Wen, X. Comparative transcriptome analysis combining SMRT-and Illumina-based RNA-Seq identifies potential candidate genes involved in betalain biosynthesis in pitaya fruit. *Int. J. Mol. Sci.* **2020**, *21*, 3288. [[CrossRef](#)]
158. Ramanauskas, K.; Igić, B. RNase-based self-incompatibility in cacti. *New Phytol.* **2021**, *231*, 2039–2049. [[CrossRef](#)]
159. Torres-Silva, G.; Correia, L.N.F.; Batista, D.S.; Koehler, A.D.; Resende, S.V.; Romanel, E.; Cassol, D.; Almeida, A.M.R.; Strickler, S.R.; Specht, C.D.; et al. Transcriptome Analysis of *Melocactus glaucescens* (Cactaceae) Reveals Metabolic Changes During in vitro Shoot Organogenesis Induction. *Front. Plant Sci.* **2021**, *12*. [[CrossRef](#)]
160. Helmy, M.; Awad, M.; Mosa, K.A. Limited resources of genome sequencing in developing countries: Challenges and solutions. *Appl. Transl. Genom.* **2016**, *9*, 15–19. [[CrossRef](#)]
161. McKain, M.R.; Johnson, M.G.; Uribe-Convers, S.; Eaton, D.; Yang, Y. Practical considerations for plant phylogenomics. *Appl. Plant Sci.* **2018**, *6*, e1038. [[CrossRef](#)]
162. Martínez-González, C.R.; Ramírez-Mendoza, R.; Jiménez-Ramírez, J.; Gallegos-Vázquez, C.; Luna-Vega, I. Improved method for genomic DNA extraction for *Opuntia* Mill. (Cactaceae). *Plant Methods* **2017**, *13*, 82. [[CrossRef](#)]
163. Yang, Y.; Moore, M.; Brockington, S.; Timoneda, A.; Feng, T.; Marx, H.E.; Walker, J.; Smith, S. An Efficient Field and Laboratory Workflow for Plant Phylotranscriptomic Projects. *Appl. Plant Sci.* **2017**, *5*. [[CrossRef](#)]
164. Tørresen, O.K.; Star, B.; Mier, P.; Andrade-Navarro, M.A.; Bateman, A.; Jarnot, P.; Gruca, A.; Grynberg, M.; Kajava, A.V.; Promponas, V.J.; et al. Tandem repeats lead to sequence assembly errors and impose multi-level challenges for genome and protein databases. *Nucleic Acids Res.* **2019**, *47*, 10994–11006. [[CrossRef](#)] [[PubMed](#)]
165. Michael, T.P.; VanBuren, R. Building near-complete plant genomes. *Curr. Opin. Plant Biol.* **2020**, *54*, 26–33. [[CrossRef](#)]
166. Sun, Y.; Shang, L.; Zhu, Q.-H.; Fan, L.; Guo, L. Twenty years of plant genome sequencing: Achievements and challenges. *Trends Plant Sci.* **2021**. [[CrossRef](#)] [[PubMed](#)]
167. Kyriakidou, M.; Tai, H.H.; Anglin, N.L.; Ellis, D.; Strömvik, M.V. Current Strategies of Polyploid Plant Genome Sequence Assembly. *Front. Plant Sci.* **2018**, *9*, 1660. [[CrossRef](#)]
168. Goerner-Potvin, P.; Bourque, G. Computational tools to unmask transposable elements. *Nat. Rev. Genet.* **2018**, *19*, 688–704. [[CrossRef](#)] [[PubMed](#)]
169. Tao, Y.; Zhao, X.; Mace, E.; Henry, R.; Jordan, D. Exploring and Exploiting Pan-genomics for Crop Improvement. *Mol. Plant* **2019**, *12*, 156–169. [[CrossRef](#)] [[PubMed](#)]
170. Bayer, P.E.; Golicz, A.A.; Scheben, A.; Batley, J.; Edwards, D. Plant pan-genomes are the new reference. *Nat. Plants* **2020**, *6*, 914–920. [[CrossRef](#)]
171. Andrews, K.R.; Good, J.M.; Miller, M.R.; Luikart, G.; Hohenlohe, P.A. Harnessing the power of RADseq for ecological and evolutionary genomics. *Nat. Rev. Genet.* **2016**, *17*, 81–92. [[CrossRef](#)]
172. Acha, S.; Majure, L.C. A New Approach Using Targeted Sequence Capture for Phylogenomic Studies across Cactaceae. *Genes* **2022**, *13*, 350. [[CrossRef](#)]
173. Romeiro-Brito, M.; Telhe, M.C.; Amaral, D.T.; Franco, F.F.; Moraes, E.M. *A Target Enrichment Probe Set for Deep and Shallow Phylogenetic Studies in Cactaceae*, Universidade Federal de Sao Carlos: Sorocaba, SP, Brazil, 2022; to be submitted.
174. UIS. Research and Development. UNESCO Institute of Statistics. Available online: <http://uis.unesco.org/en/topic/research-and-development> (accessed on 14 February 2022).

-
175. Lewin, H.A.; Robinson, G.E.; Kress, W.J.; Baker, W.J.; Coddington, J.A.; Crandall, K.A.; Durbin, R.; Edwards, S.V.; Forest, F.; Gilbert, M.; et al. Earth BioGenome Project: Sequencing life for the future of life. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 4325–4333. [[CrossRef](#)]
 176. International Virtual Mini-Symposium “Cactaceae: Phylogenetics, Evolution and Conservation in the Genomic Era”. Available online: <https://dbg.org/events/cactus-symposium-2021/> (accessed on 22 February 2022).