Biogeographic History of *Chamaecrista*: Ancient Occupation Versus Recent Radiation In The Brazilian Mountains

Biogeographic History of *Chamaecrista*: the role of Brazilian savanna in the Leguminosae diversification in the Neotropics

Juliana Gastaldello Rando1, Matheus Martins Teixeira Cota2, Alexandre Zuntini3, Anselmo Nogueira4, Suzana Alcantara5, Gwilin Peter Lewis6, Luciano Paganucci de Queiroz2 & José Rubens Pirani9

1. Universidade Federal do Oeste da Bahia, Programa de Pós Graduação em Ciências Ambientais, Rua Professor José Seabra Lemos, 316, Recanto dos Pássaros, 47808-021, Barreiras, Bahia, Brazil.

2. Universidade Estadual de Feira de Santana, Departamento de Ciências Biológicas, Laboratório de Sistemática Molecular de Plantas, Km 03 BR 116, Campus, 44031-460, Feira de Santana, Bahia, Brazil.

3.

4.

5.

6.

7. Instituto de Biociências, Departamento de Botânica, Universidade de São Paulo, Rua do Matão 277, 05508-900, São Paulo, Brazil.

*Corresponding author:* [*ju.gastaldello.rando@*](mailto:ju_rando@hotmail.com)*gmail (author for correspondence)*

Journal of Biogeography

ABSTRACT

Aim:

Location: Brazilian Cerrado

Taxon: *Chamaecrista* (Leguminosae)

Methods:

Results:

Main Conclusions:

KEYWORDS:

INTRODUCTION (em torno de 900 palavras, mas está com 1423)

The Neotropical flora is marked by its high diversity, harboring the most-species ecoregions in the world, and with an exceptional number of endemic species (Gentry 1988, Myers et al. 2000, Antonelli and Sanmartín 2011). This richness of species in Neotropical region has been the targeted in several studies of biologists, and with the increment of methods incorporating tempo in the phylogenies in the last decades, they have associated geological events to evolutionary history of specific lineages to understand the causes that could explain this patterns of diversity (e.g Hughes & Eastwood 2006, Antonelli et al. 2009, 2011, Bacon et al. 2013, Egbert et al. 2014, Fine et al. 2014, Vasconcelos et al. 2020). Thus, these studies have reinforced that specific geological events in South America, as the isolation as an “island continent” for a long geological period, the Andes uplift, the landbridge formation through the raise of the Panama isthmus, and by the paleoclimatic fluctuations, may have lead, in large scale, changes in plant cover and in its diversification (see also Burnham & Graham 1999). Besides those, several other factors listed as biotic and abiotic forces leading to diversification have been also discussed in an attempt to explain the high number of species existing in the Neotropical region (Antonelli & Sanmartín 2011, Lagomarsino et al. 2016).

The huge diversity found throughout the vegetation coverage in the Amazonian, Andes, and Atlantic Forest regions explains the emphasis given to these systems in evolutionary and biogeographic studies in South America (eg. Wesselingh et al. 2010, Antonelli et al. 2009, Antonelli and Sanmartín 2011, Fine et al. 2014, Lagomarsino et al. 2016, falta de mata Atlântica). And, few studies focusing on the biogeographic history and diversification of plants from species-rich ecosystems such as Neotropical savanna and *campo rupestre* vegetation (but see Simon et al. 2009, Simon & Pennington 2012, Vasconcelos 2020).

Cerrado domainis a term applied to the prevailing formation of Neotropical savanna in central South America, with predominance of grassy ground cover, with variable density of trees that extends from Northeastern Brazil into Paraguay and Bolivia, reaching its extensive area in central Brazil (Ratter 2003, Gottesberger & Gottesberger 2006, Azevedo et al. 2020). After the Amazon, the Cerradois the largest formation in the Neotropics, covering more than 2 million km2, and harbours the most diverse flora among the other savannas in the world (Murphy et al. 2016). Due to this high diversity, with high levels of endemic species strongly threatened, it is also considered a global biodiversity hotspot area (Myers et al. 2000). Currently, only 19,8% of its original extension is covered by the original vegetation (Strassburg et al. 2017). The high levels of endemism are not uniformly distributed in the Cerrado landscape. Representative numbers are found along highland areas, where a specific vegetation occurs on the rocky outcrops, known as *campo rupestre* (rupestrian grasslands), and it exhibits outstanding levels of species richness and endemism (Prance 1994, Vasconcelos 2008, 2009, Alves, Silva, Oliveira, & Medeiros, 2014, Colli et al. 2019). Traditionally, the *campo rupestre* has been considered as part of the Cerrado by most botanists and plant ecologists (Ribeiro & Walter 1998, Gottesberger & Gottesberger 2006, Simon & Proença 2000, Simon et al. 2009). More recently, it has been emphasized the necessity to recognize this area as a distinct bioregion in order to contribute to a better understanding of biogeographical patterns in the Neotropics (Colli et al. 2019).

Although the floristic distinctiveness between Cerrado and *campo rupestre*, they also had distinct geological formations. The initial development of Neotropical savannas is estimated to be ancient, between 25 to 28 million years (Graham 1991, Burnham & Graham 1999, Azevedo et al. 2020), however the few studies performed with plants has indicated the establishment of the vegetation with typical elements as found today to be more recent, around four million years, during the late Quaternary (Simon et al. 2009). The origin of Neotropical savanna is likely related to the retraction of forest vegetation and the global increase of plant biomass with C4 photosynthesis (Cerling et al. 1997, Cerling et al. 1998, Bouchenak-Khelladi et al. 2010, Arakaki et al. 2011). Several works have enfazy that the fire regime is an important eco-evolutionary force in this biome (Oliveira-Filho & Ratter 1995, Ratter et al. 1997, Pennington et al. 2000, Simon et al. 2009, Simon & Pennington 2012), where the natural accumulation of plant biomass becomes the fuel material for occasional fires. Evidence for plant adaptations to the fire regime were observed in several Cerrado species, such as thick leaves, corky bark, xylopodia and the larger ability of plants to survive to this disturbance (Pennington et al. 2000, Simon & Pennington 2012).

Already the *campo rupestre* occurs on a much older formation, the Brazilian highlands are part of most ancient landscapes on earth (Shaefer et al. 2016), in the case of Espinhaço Range, it is a relict of an ancient sea floor and deserts (Barbosa e Sad, 1973), rising in the pre-Cambrian, nearly 640 Ma (Alkmin, 2012). Despite its old geological age, likewise to Cerrado, the phylogenetic studies have indicated a recent diversification of most of groups sampled (e.g. Souza et al. 2013, Rando et al. 2016, Vasconcelos et al. 2020) with few exceptions (Alcantara et al. 2018, Vasconcelos et al. 2020). In general, it is suggested the climatic fluctuations of the Pleistocene had an important role in the diversification of plants, considering that the tempo of diversification of the most studied lineages coincide with this geological event (Vasconcelos et al. 2020). In the absence of fossil records, molecular phylogenetic trees have been the best tool to explore the diversification of plants in the Cerrado and in the *campo rupestre*. Despite this gap causing limitations in the investigations about the past of these floras, it is possible that the increment of phylogenies associating with the particularities of each group can bring more evidence for the possible processes that lead to diversification in these species-rich ecosystems (Simon et al. 2009, Vasconcelos et al. 2020). In relation to shared abiotic features, as well as in the Cerrado, the *campo rupestre* is a fire-prone vegetation, and the plants can also respond to this disturbance with morphological, physiological and demographic specializations (Conceição & Pirani, 2005, Figueira et al. 2016). The evolutionary history of *campo rupestre* vegetation is strongly connected with that of Cerrado vegetation, where the fire is a very important and evolutionary factor for both (Figueira et al. 2016).

The Leguminosae is the most species-rich family of plants in the Neotropics (Gentry 1988). In Brazil, it is also the richest group, with 3.033 species being more than 50% endemic to the country (Flora do Brasil 2021). Although the well known diversity of subfamily Papilionoideae, with around 14.000 species (71,5% of Leguminosae) widely distributed in the world, while the other five subfamilies reaching to approximately 5.600 species (28,5%) (LPWG 2017), in the Neotropics is observed an inversion of these values of diversity. For example, in Brazil, the diversity of Legumes family is divided in 25% of Papilionoideae and 75% of the other occurents subfamilies (Flora do Brasil 2020). Interestingly, this inversion of values can be mainly attributed to the high diversity of two genera found in the Cerrado and *campos rupestres*, *Chamaecrista* (L.) Moench and *Mimosa* L., both belonging to the Caesalpinioideae subfamily. *Mimosa* and *Chamaecrista* comprise together more than 1/3 of the Leguminosae diversity in the Cerrado (Dutra et al. 2020, Rando et al. 2020). The diversity of *Mimosa* is more centered in the Cerrado, while *Chamaecrista* shows a great diversity in the Cerrado and also in the *campo rupestre*. In both ecosystems, *Chamaecrista* is the second most diverse genus of the flora (Silveira et al. 2016, Rando et al. 2020).

Indeed, with all expressive diversity and consequently importance in the Cerrado and *Campos rupestres* vegetation, *Chamaecrista* is a remarkable group for biogeographical and evolutionary studies in the Neotropical Flora. Thus, the main objective of this study was to investigate the biogeographical history of a pantropical genus, a plant group with a high diversity in the Neotropical region centred in the Brazilian savanna and in the *campo rupestre*, using two approaches: a global scale investigation, including a well-sampled phylogeny of closely related genera (Caesalpinia + Cassiae groups); and a more detailed scale for Neotropical region. In these approaches we addressed three main questions: (1) How biogeographical processes (dispersion, vicariance and extinction) have interacted during its diversification? (2) When, where and how was the diversification of *Chamaecrista* in Brazilian savanna and in the *campo rupestre*? (3) Is there evidence of habitat shift associated with morphological transitions during the *Chamaecrista* diversification?

**Material and Methods**

**Taxon sampling**

*Chamaecrista* is a pantropical genus with 366 species, 295 are restricted to Neotropical region (LPWG 2021), however the centre of diversity is in the *Cerrado* and in the *campo rupestre* vegetation, where are found 225 species, 174 endemics (Irwin & Barneby 1982, Rando et al. 2020). This richness also makes the *Chamaecrista* the eighth most diverse genus in the Brazilian Flora (Rando et al. 2020). *Chamaecrista* appeared within the clade classified as Cassiae group (7 genera/ ~750 spp), that is sister of Caesalpinia group (26 genera/ ~217 spp) (Gagnon et al. 2016, LPWG 2017).

In order to investigate the history of biogeographic events in Caesalpinia and Cassinae groups and in more detail *Chamaecrista*, we first selected all genera and species with available sequences of these groups. Out of the 1318 DNA sequences used for the phylogenetic and biogeographic analyses, 891 were retrieved from GenBank, and 427 were new. From GenBank, we preferably selected the specimens collected by experts in Leguminosae, and used by them in other works of Leguminosae taxonomy (Marazzi et al. 2006, Conceição et al. 2009, Rando et al. 2016, Gagnon et al. 2016, LPWG 2017, Cota et al. 2021). This sampling covered all 33 genera recognized in the Caesalpinia and Cassiae groups, and 43 % of all diversity of species, covering mostly four regions *(trnK/MatK*, *TrnD-T*, *TrnL-F* and ITS). Our focal group *Chamaecrista* is represented by 231 species, and in addition to the cited regions, we included one more nuclear region (ETS). An effort has been taken to sample species of *Chamaecrista* from all Brazilian floristic regions, and a dense sampling was performed in the Cerrado and *campo rupestre* vegetation, which is the diversity center of the genus (Fig. X). The sampling of *Chamaecrista* represented 64% of species in the world, and 80% of Brazilian diversity. Genbank accession numbers for DNA sequences generated in this study are associated with voucher specimens (Appendix Y). The procedures in the laboratory for the extraction, amplification and sequencing followed the protocols already used in Conceição et al. (2009), Rando et al. (2016) and Queiroz & Snak (2020). The detailed laboratory procedures are described in the Supplementary S1.

**Time calibration**

Because of the great number of species only with *trnK/MatK* region (mainly for Caesalpinia group and *Senna*), and ETS region basically for *Chamaecrista* species. Tem alguma referência que fala que missing data pode ser ruim nas análises no BEAST???? e falar um pouco sobre o TreePL. Thus, we performed these two methods of time calibration tree in order to compare the age obtained in these analyses and use that one with better sampling of species and regions.

**Bayesian divergence age estimation**

Forward and reverse sequencing reads were assembled into contigs and edited in Geneious Prime software (Biomatters 2020). Sequences of all loci were aligned by Muscle (Edgar 2004a, 2004b), with manual edition in Geneious software to correct obvious alignment errors and to remove sections of dubious quality. The total alignment comprises 3068 base pairs, partitioned in five different regions (Table 2).

Due to the great number of species only with the *trnK/MatK* we firstly perform the Bayesian divergence time using exclusively this region, and another analysis more restrictive in the number of species with all regions sampled (*trnK/MatK*, *TrnD-T*, *TrnL-F*, ETS and ITS). Both analyses were implemented on CIPRES Science Gateway (Miller et al. 2010) in BEAST2 with parameter files set up in Beauti (Drummond & Rambaut, 2012; Bouckaert et al. 2014, 2019) in the CIPRES Science Gateway (Miller et al. 2010).

The molecular clocks were unlinked across partitions, in both analyses we adopted the RB (Reversible Jump) parameter (Reference). The analysis was performed using an uncorrelated-rates relaxed molecular clock model (UCLD), and the Birth Death tree model. For the *trnK/MatK* analysis, we constrained the ages of four nodes. For the root node (*Cassia* clade + Caesalpinia clade) the age estimate was taken from Bruneau et al. (2008), 58.6 mya and standard deviation XXX, assuming a normal prior of distribution rate. Three fossils were used for calibration: *Mezoneuron*, *Caesalpinia* s.s., and *Senna* (table 3). All these fossils were used by Bruneau et al. (2008) and all were assigned to the stem node of the groups (figure XX) using a log-normal prior with a value of 45.0 of mean and standard deviation of 0.005. For the more restrictive analysis, we constrained the ages of two nodes. For the root node (Cassia clade), the age estimate was also taken from Bruneau et al. (2008), 53 mya and standard deviation 1.567, assuming a normal prior of distribution rate. And, only the *Senna* fossil was assigned in this case. The estimated ages taken from Bruneau et al. (2008) in both cases were obtained with 19 points of calibration using 18 fossils of Legumes of different genera.

Independent runs were performed with 50,000,000 generations for *trnK/MatK* (three runs)and 100,000,000 generations for the five regions in the more restrictive analyses (XX runs). The convergence of individual runs was assessed using Tracer v.1.6 (Drummond & Rambaut, 2007b). All analyses were combined in the LogCombiner v1.8.0 program, the burning was established for 10%. All trees were summarized in the TreeAnnotator v1.8.0 program.

**Maximum likelihood divergence age estimation**

| Node | Taxonomic identify and phylogenetic position | Age | SD | Fossil organ, locality | References |
| --- | --- | --- | --- | --- | --- |
| Stem node | *Mezoneuron* | 45 Ma. |  | Fruits several sites SE and W USA | Herendeen and Dilcher (1991) |
| Stem node | *Caesalpinia* s.s | 45 Ma. |  | Fruits and Leaflets, SE USA, | Herendeen (1992) |
| Stem node | *Senna* | 45 Ma. |  | Fruits, SE USA and Mexico | Herendeen (1992), Calvillo Canadell and Cevalloz Ferriz (2005) |

Table 3.

we applied the constraint age using a log-normal prior with a value of 45.0 of mean and standard deviation of 0.005.

**Geographical distribution and Bioregionalization** (289 words)

In order to define the bioregions we adopted the network approach as the base for the delimitation (Vilhena & Antonelli 2015). The occurrences were extracted from Global Biodiversity Information Facility (GBIF, [www.gbif.org](http://www.gbif.org)), through the R package ‘rgbif’ (Chamberlain et al. 2020) (GBIF Occurrence Download for 317 species of Caesalpinia + part of Cassiae clade https://doi.org/10.15468/dl.h8bsbs, and for 324 species of *Chamaecrista* <https://doi.org/10.15468/dl.5gxx8t>). The two datasets were filtered in order to flag records without coordinates, or with coordinate displacements, and historical collections (here adopted before 1950). All these filtering steps were performed in R (R Core Team 2021) through the ‘CoordinateCleaner’ package (Zizka et al. 2019). After the filtering, we checked the range of distribution for each species, and when some occurrence arose out of the known distribution, the specimens were checked in order to certify the determination (when images of vouchers are available) or were excluded. The known distribution was obtained from specific literature, herbarium data, and our own collections. The consulted references for each species, as well as its distribution and corresponding biogeographic areas are detailed in the Supplementary S2. The datasets were implemented in the Infomap Bioregions web application that identifies taxon-specific bioregions from species distribution data (Edler et al. 2016, https://www.mapequation.org/bioregions/). Our results were compared with Brazilian Domains (IBGE 2004), biogeographical provinces of Morrone (2001), and ecoregions of Olson et al. (2001), and the new provinces of Colli et al. (2019), looking for correspondent regions and which one could be more appropriate to include in the broader and in the detailed biogeographic analyses. More details of the bioregionalization, and the parameters used in the network analyses, see Supplementary S3. The bioregions delimited were: wider analyses A=Paleotropic, B=Indo-Malay, C=Austrolasia, D=Afrotropic, E= Neartic (excluding Baja California, Sonora, Altiplano Mexicano and Tamaulipas), F= Neotropical region I, Caribbean Subregion including Baja California, Sonora, Altiplano Mexicano and Tamaulipas, G = Neotropical region II, H = Andina Region. And for detailed analyses A= Amazonia, B= Cerrado, C=Caatinga, D= Campos rupestres, E=Mata Atlantica, F=Central America Central, G=North America, H= Africa, I=Paleotropic + Indo-Malay and J = Austrolasia.

**Biogeographic Analyses**

To evaluate the biogeographic history of Caesalpinia + Cassiae group, and in more details for *Chamaecrista*, we employed maximum-likelihood inference of geographic range using the Dispersal, Extinction, and Cladogenesis model (DEC) (Ree et al. 2005; Ree and Smith 2008) implemented in R using the "BiogeoBears" package (Matzke, XXXX)

To study geographic range evolution through time, we evaluated alternative biogeographical hypotheses by modeling area connectivity in three ways (Figure 2). *Model A*: We defined a simple model with no dispersal constraint (with no information based on paleogeography). According to this model, the rate of dispersal had a value of 1 (the highest dispersal rate) between all areas along the whole diversification period of *Chamaecrista*. Additionally, we defined two alternative hypotheses (*Models B* and *C*) with spatio-temporal dispersal constraints reflecting the likely paleogeographic history of tropical America from the Eocene onward (Gottesberger & Gottesberger 2006, Pennington et al. 2006, Antonelli et al. 2009) adopting the time-stratified parameters. *Model B*: First, we defined three time frames: time frame 1 (53.9 – 28 mya, Eocene and early Oligocene), time frame 2 (28 – 10 mya, Oligocene and late Miocene), and time frame 3 (10 mya – until to present, from Oligocene and all Quaternary) (Figure 2). For each time frame we applied different rates of dispersal between areas. For the time frame 1 we considered a low dispersal rate from Amazon Region (area A) to the other biogeographical areas (rate=0.1), based mainly on the putative lack of savannas (areas B and C), restingas (area E) and of a connection between South America and the rest of Americas (Panama isthmus not yet established, area G). On the other hand, dispersal between Amazon (area A) and Paranaense Subregion (area D) was probably facilitated by the existence of larger extensions of forests during that period, so we accept the value 1 only between these two areas. In the time frame 2, we considered the arising of Neotropical savannas defined as a higher dispersal rate between Amazon and savannas (rate=1). At the same time dispersal from the Amazon to the Paranaense Subregion would be more difficult in a putative scenario of retraction of rainforests and expansion of savannas (rate=0.1). For time frame 3 we considered a high dispersal rate between almost all areas due to the effective connection between the Americas and the establishment of the biogeographical regions within South America known today. In this case the dispersal to Central and North America from the South could be easier and the dispersal among Cerrado (area B), Caatinga (area C), Campos rupestres (area F), and Restingas (area E) could also be facilitated (rate =1). *Model C*: In this model we used the same time frame from *Model B*. The differences between *Model C* and *B* are in the time frame 1 and 2. In the time frame 1 we hypothesized that ancestral species could disperse from the Amazon region to Paranaense subregion (similar *Model B*) and also to Campos rupestres (rate=1). This hypothesis was implemented for considering that the geological history of Espinhaço Range (the main area presenting campo rupestre vegetation) is very old, probably arising in the proterozoic (Marshak et al. 2006). And in the time frame 2 we considered just dispersion from campos rupestres to savanna areas (rate=1) using a lower dispersal rate (rate=0.1) between Amazon subregion and savannas areas. The best biogeographic model was chosen using the likelihood ratio test (LRT) comparing the pairwise models.Detailed dispersal rates for each model and the scripts are available in DRYAD?. Parâmetro J.

REFERENCES

Alves, R. J. V., Silva, N. G., Oliveira, J. A., & Medeiros, D. (2014). Circumscribing *campos rupestres* – egadiverse Brazilian rocky montane savannas. *Brazilian Journal of Botany*, *74*, 355–362. https://doi. org/10.1590/1519‐6984.23212

Amorim, D. & Pires, M.R.S. 1996. Neotropical biogeography and a method for maximum biodiversity estimation. *In* Bicudo, C.E.M. & Menezes, N.A. (eds.) *Biodiversity in Brazil. A first approach*. CNPq, São Paulo, p 183-219.

Andrade-Lima, D. 1982. Present-day forest refuges in Northeastern Brazil. In Prance, G.T. (ed.) *Biological diversification in the tropics*. Columbia University Press, New York, p. 245-251.

Antonelli, A., Nylander, J.A.A., Persson, C. & Sanmartín, I. 2009. Tracing the impact of the Andean uplift on Neotropical plant evolution. PNAS 106 (24) 9749-9754.

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

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

Azevedo, J.A.R.; R.G. Collevatti, C.A. Jaramillo, C. A. E. Strömberg, T. B. Guedes, P. Matos-Maraví, C. D. Bacon, J. D. Carillo, S. Faurby and A. Antonelli. 2020. On the Young Savannas in the Land of Ancient Forests *In* Neotropical Diversification: Patterns & Process. Rull, V. & Carnaval, A.C. (eds). Springer Nature Switzerland, pag. 271- 298.

Barbosa ALM, Sad JHG (1973) Tectonic control of sedimentation and trace-element distribution in iron ores in central Minas Gerais (Brazil). In: Proceedings of kiev symposium 1970, Genesis of precambrian iron and manganese deposits, UNESCO, pp. 125–131

Batalha-Filho, H.; Fjeldsa, J.; Fabre, P.H. & Miyaki, C.Y. 2013. Connections between the Atlantic and the Amazonian forest avifaunas represent distinct historical events. *Journal of Ornithology* 154:41-50.

Bacon, C.D., Mora, A.; Wagner, W.L. & Jaramillo, C.A. 2013. Testing geological models of evolution of the Isthmus of Panama in phylogenetic framework. *Botanical Journal of the Linnean Society*, 171:287-300.

Bitencourt, C. & Rapini, A. 2013. Centres of Endemism in the Espinhaço Range: identifying cradles and museums of Asclepiadoideae (Apocynaceae). *Systematics and Biodiversity* 11:525-536.

Bouchenak-Khelladi, Y., Verboom, G.A., Savolainen, V. & Hodkinson, T.R. 2010. Biogeography of the grasses (Poaceae): a phylogenetic approach to reveal evolutionary history in geographical space and geological time. *Botanical Journal of the Linnean Society* 162: 543-557.

Bouckaert R, Heled J, Kühnert D, Vaughan T, Wu C-H, Xie D, Suchard MA, Rambaut A, Drummond AJ. 2014. BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Computational Biology* 10: e1003537.

Bouckaert R, Vaughan TG, Barido-Sottani J, Duchêne S, Fourment M, Gavryushkina A, Heled J, Jones G, Kühnert D, De Maio N, Matschiner M, Mendes FK, Müller NF, Ogilvie HA, du Plessis L, Popinga A, Rambaut A, Rasmussen D, Siveroni I, Suchard MA, Wu C-H, Xie D, Zhang C, Stadler T, Drummond AJ. 2019. BEAST 2.5: an advanced software platform for Bayesian evolutionary analysis. *PLoS Computational Biology* 15: e1006650.

Bruneau, A.; Mercure, M.; Lewis, G.P.; & Herendeen, P.S. 2008. Phylogenetic patterns and diversification in the Caesalpinioid legumes. *Botany* 86:697-718.

Burnham, R.J. & Graham, A. 1999. The history of Neotropical Vegetation: New Developmentes and status. *Annals of the Missouri Botanical Garden* 86:546-589.

Carvalho, C.J.B. & Couri, M.S. 2011. Biogeografia de Muscidae (Insecta, Diptera) da América do Sul. In Carvalho, C.J.B. & Almeida, E.A.B. (orgs.) *Biogeografia da América do Sul*. *Padrões e processo*s. ROCA, São Paulo, p. 277- 298.

Castro, A.A.J.F.; Martins, F.R.; Tamashiro, J.Y. & Shepherd, G.J. 1999. How Rich is the Flora of Brazilian Cerrados? *Annals of the Missouri Botanical Garden* 86: 192-224.

Cerling, T.E., Harris, J.M., MacFaddeb, B.J., Leakey, M.G., Quade, J., Eisenmann, V., Ehleringer, J.R. 1997. Global vegetation changes through the Miocene/Pliocene boundary. *Nature* 389:152-158.

Cerling, T.E., Harris, J.M., MacFaddeb, B.J., Quade, J., Leakey, M.G., Eisenmann, V., Ehleringer, J.R. 1998. Miocene/Pliocene shift: one step or several? *Nature* 393:126-127.

Chamberlain, S. et al. (2020). Rgbif: Interface to the Global Biodiversity Information Facility API. R package version 2.2.0.

Cody, S. Richardson, J.E., Rull, V., Ellis, C. & Pennington, R.T. 2010. The Great American Biotic Interchange revisited. *Ecography* 33: 326-332.

Colli-Silva, M. et al. (2019). Outstanding plant endemism levels strongly support the recognition of campo rupestre provinces in mountaintops of eastern South America. Journal of Biogeography, 46: 1723-1733.

Dasmman, R.F. 1974. Biotic Provinces of the World: Further Development of a System for Defining and Classifying Natural Regions for Purposes of Conservation. *IUCN Occasional Paper* No. 9. International Union for Conservation of Nature and Natural Resources, Morges, Switzerland.

Drummond, A.J., Ho, S.Y.W., Phillips, M.J. & Rambaut, A. 2006. Relaxed phylogenetics and dating with confidence. *Public Library of Science. Biology* 4:e88.

Drummond AJ, Suchard MA, Xie D, Rambaut A. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution* 29: 1969–1973.

Echternacht, L., Trovó, M., Oliveira, C.T. & Pirani, J.R. 2011. Areas of endemism in the Espinhaço Range in Minas Gerais, Brazil. *Flora* 206: 782–791.

Edler, D., Guedes, T., Zizka, A. Rosvall, M., Antonelli, A. (2016). Infomap Bioregions: Interactive Mapping of Biogeographical Regions from Species Distributions. Systematic Biology, 66(2): 197-204.

Egbert, G. L., Aaron, O’Dea & Vermeij, G.J. 2014. Historical biogeography of the Isthmus of Panama. *Biological Reviews* 89:148-172.

Fine, P.V.A., Zapata, F. & Daly, D. 2014. By Examining the evolution and Historical Biogeography of the Protieae (Burseraceae). *Evolution*, 68-7: 1988-2004.

Gagnon, E., Bruneau, A., Hughes, C.E., Queiroz, L.P. & G.P. Lewis. (2016). A new generic system for the pantropical Caesalpinia group. PhytoKeys, 71: 1-160.

Gentry, A.H. 1988. Changes in plant community diversity and floristic composition

on environmental and geographical gradients. Ann. Mo. Bot. Gard. 75:1–34.

Giulietti, A.M. & Pirani, J.R. 1988. Patterns of geografic distribution of some plants species from the Espinhaço Range, Minas Gerais and Bahia, Brazil. *In Proceedings of a Workshop on Neotropical Distribution Patterns.* W.R. Heyer & P.E. Vanzolini, eds. Rio de Janeiro, Anais da Academia Brasileira de Ciências, p.39-69.

Giulietti, A.M., Pirani, J.R. & Harley, R.M. 1997. Espinhaço Range region, Eastern Brazil. *In The Americas centres of plant diversity: A guide and strategy for their conservation.* S.D. Davis, V.H. Heywood, O. Herrera- Macbryde, J. Villa-Lobos & A.C. Hamilton, eds. IUCN Publication Unity, Cambridge, v.3, p.397-404.

Gottsberger, G. & I. Silberbauer-Gottsberger. 2006. *Life in the Cerrado: A South American Tropical Seasonal Vegetation. Vol. I. Origin, Structure, Dynamics and Plant Use*. Reta Verlag, Ulm, Germany.

Harley, R.M. 1988. Evolution and distribution of *Eriope* (Labiatae), and its relatives, in Brazil. In: Vanzolini, P.E. & Heyer, W.R. (eds). *Proceedings of a workshop on Neotropical distribution patterns*, p. 71-120. Academia Brasileira de Ciências, Rio de Janeiro.

Herendeen, P.S., Crept, W.L.; & DilcherI, D.L. 1992. The fossil history of the Leguminosae: phylogenetic and biogeographic implications. In *Advances in legume systematics*. Part 4. P.S. Herendeen & D.L. Dilcher eds. Royal Botanic Gardens, Kew, Surrey, UK. pp. 303–316.

IBGE 2004

Irwin, H.S. & Barneby, R.C. 1978. Monographic studies in *Cassia* (Leguminosae - Caesalpinioideae) III. Sections Absus and Grimaldia. *Memoirs of the New York Botanical Garden* 30: 1-277.

Irwin, H.S. & Barneby, R.C. 1982. The American Cassinae: a synoptical revision of Leguminosae tribe Cassieae subtribe Cassinae in the New World. *Memoirs of the New York Botanical Garden* 35: 455-918.

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

Loeuille, B., Lohmann, L., Semir J,. & Pirani, J.R. IN PRESS. A phylogenetic analysis of Lychnophorinae (Asteraceae: Vernonieae) based on molecular and morphological data. *Systematic Botany*.

Lomolino, M. Riddle, Brown, J.H. 2006. Island Biogeography: Patterns in species richness in Biogeography 3o ed. Lomolino, M. Riddle, Brown, J.H eds. Sinauer Associates, Inc.

​​Legume Phylogeny Working Group (LPWG), Andrella, G. C., Atahuachi Burgos, M., Bagnatori Sartori, Â. L., Balan, A., Bandyopadhyay, S., Barrett, R. L., Barbosa Pinto, R., Brullo, S., Boatwright, J. S., Broich, S. L., Bruneau, A., Cardinal-McTeague, W., Cardoso, D., Castro Silva, I. C., Cervantes, A., Choo, L. M., Cobra e Monteiro, T., Compton, J., et al. (2021). *The World Checklist of Vascular Plants (WCVP): Fabaceae, vers. June 2021* (R. Govaerts, Ed.; ). http://sftp.kew.org/pub/data\_collaborations/Fabaceae/DwCA/

Marazzi, B., Endress, P.K., Queiroz, L.P., Conti, E. 2006. Phylogenetic relationships within *Senna* (Legumniosae, Cassinae) based on three chloroplast DNA regions: Patterns in the Evolution of Floral symmetry and extrafloral nectaries. American Journal of Botany, 93(2): 288-203.

Maddison, W.P. & D.R. Maddison. 2010. Mesquite: A modular system for evolutionary analysis. Version 2.74. http://mesquiteproject.org

Martini, A.M.Z; Fiaschi, P.; Amorim, AM.; Paixão, J.M. et al. 2007. A hot-point within a hot-spot: a high diversity site in Brazil’s Atlantic Forest. *Biodiversity and Conservation* 16:3111-3128.

Martins, F.M., Templeton, A.R., Pavan, A.C.O., Kohlbach, B.C., Morgante, J.S. 2009. Phylogeography of the common vampire bat (*Desmodus rotundus*): marked population structure, Neotropical Pleistocene vivariance and incongruence between nuclear and mtDNA markers. *BMC Evolutionary Biology* 9:294.

Marshak, S.; Alkmim, F.F.; Whittington, A.; Pedrosa-Soares, A.C. 2006. Extensional collapse in the Neoproterozoic Araçuaí orogen, eastern Brazil: a setting for reactivation of asymmetric crenulation cleavage. *Journal of Structural Geology* 28, 129-147.

[Miller, M.A., Pfeiffer, W., and Schwartz, T. 2010. "Creating the CIPRES Science Gateway for inference of large phylogenetic trees" in Proceedings of the Gateway Computing Environments Workshop (GCE), 14 Nov. 2010, New Orleans, LA pp 1 - 8.](http://www.phylo.org/sub_sections/portal/sc2010_paper.pdf)

Morrone, J. J. (2001). Biogeografia de América Latino y el Caribe. M&T – Manuales & Tesis SEA, vol.3. Zaragoza, 148 pp.

Morton, J.K. 1972. Phytogeography of the West African mountains. *In* Valentine, D.H. (ed.) *Taxonomy, phytogeography and evolution*. Academic Press, London. p. 221-236.

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

Nogueira, A. Rey, P.J. & Lohmann, L.G. 2012. Evolution of exfloral nectaries: adaptative process and selective regime changes from forest to savanna. *Journal of Evolutionary Biology* 25:2325-2340.

Oliveira, P.E., Barreto, A.M.F. & Suguio, K. 1999. Late Pleistocene/Holecene climatic and vegetational history of the Brazilian caatinga: the fossil dunes of the middle São Francisco River. *Palaeogeography, Palaeoclimate, Palaeoecology* 152: 319-337.

Oliveira-Filho, A.T. & Ratter, J.A. 1995. A study of the origin of central Brazilian forests by the analysis of plant species distribution patterns. *Edinburg Journal of Botany* 52(2): 141-194.

Olson et al. (2001)

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

Pennington, R.T.; Gwilym P.L. & Ratter, J.A. 2006. An Overwiew of the Plant Diversity, Biogeography and Conservation of Neotropical Savannas and Seasonally Dry Forests in *Neotropical savannas and seasonally dry forests: Plant Diversity, Biogeography, and Conservation.* Pennington, R.T.; Gwilym P.L. & Ratter, J.A. eds. Taylor & Francis Group.

Prance, G.T. 1982. A review of the phytogeographic evidences for Pleistocene climate changes in the Neotropics. Annals of the Missouri Botanical Garden 69: 594-624.

Prance, G.T. 1994. A comparison of the efficacy of higher taxa and species numbers in the assessment of biodiversity in the Neotropics. *Philosophical Transactions of the Royal Society of London* 345:89-99.

Quantum Gis. A free and open-source software (FOSS) community. Avaliable in <http://qgis.org/en/site>.

R Core Team (2020). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Rando, J. G. & Pirani, J. R. 2011. Padrões de distribuição geográfica das espécies de *Chamaecrista* sect *Chamaecrista* ser. *Coriaceae* (Benth.) H.S. Irwin & Barneby (Leguminosae – Caesalpinioideae). *Revista Brasileira de Botânica* 34: 499-513.

Rando, J. G.; Loeuille, B. & Pirani, J. R.2013. Taxonomic novelties in *Chamaecrista* (Leguminosae: Caesalpinioideae) from Brazil. *Phytotaxa* 9: 17-25.

Ratter, J.A.; Bridgewater, S. & Ribeiro, J.F. 2003. Analysis of the floristic composition of the Brazilian Cerrado vegetation III: comparison of the woody vegetation of 376 areas. *Edinburgh Journal of Botany* 60(1): 57-109.

Raven, Peter H., Ray F. Evert, and Susan E. Eichhorn.1999. *Biology of Plants*, 6th ed. New York: W. H. Freeman and Company.

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

Ree, R.H. and Smith A.S. (2008). Maximum Likelihood Inference of Geographic Range Evolution by Dispersal, Local Extinction, and Cladogenesis. *Systematic Biology* 57:4-14.

Renner, S. 2004. Plant dispersal across the tropical atlantic by wind and sea currents. *Int. J. Plant Sci.* 165:23-33.

Ribeiro, J.F. & Walter, B.M.T. 1998. Fitofisionomias do bioma Cerrado. In: Sano, S.M. & Almeida, S.P. *Cerrado: ambiente e flora*. Planaltina: EMBRAPA CPAC. p.89-166.

Rizzini, C.T. 1979. Tratado de Fitogeografia do Brasil. 2 vols. HUCITEC/EDUSP, São Paulo.

Schaefer, C.E.; Hugo G. Cândido, Guilherme Resende Corrêa, Jaquelina A. Nunes and Daniel M. Arruda. 2016. Soils Associated with Rupestrian Grasslands In Ecology and Conservation of Mountaintop Grasslands in Brazil, G.W. Fernandes (ed.). Springer International Publishing Switzerland. DOI 10.1007/978-3-319-29808-5\_3

Species Link. 2014 Centro de Referência em Informação Ambiental, CRIA. http://splink.cria.org.br/project?&setlang=en

Simon, M. F. & Proença, C. Phytogeographic patterns of *Mimosa* (Mimosoideae, Leguminosae) in the Cerrado biome of Brazil: an indicator genus of high-altitude centers of endemism? *Biological Conservation* 96:279-296.

Simon, M.; Grether, R.; Queiroz, L.P.; Skema, C. Pennington, R.T. & Hughes, C.E. 2009. Recent assembly of the Cerrado, a neotropical plant diversity hotspot, by in situ evolution of adaptations to fire. *Proceedings of the National Academy of Sciences* 106: 1-6.

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

Souza ER, Lewis GP, Forest F, Schnadelbach AS, van den Berg C, Queiroz LP. 2013. Phylogeny of Calliandra (Leguminosae: Mimosoideae) based on nuclear and plastid molecular markers. Taxon 62: 1200–1219.

Thorne, R. 2004. Tropical Plant Disjunctions: a personal reflection. *International Plant Sciences* 165:137-138.

Wesselingh, F.P., Hoorn, C., Kroonenberg, S.B., Antonelli, A., Lundberg, J.G., Vonhof, H.B. & Kooghiemstra, H. 2010. On the origin of Amazon landscapes and biodiversity: a synthesis. In Hoorn, C. & Wesselingh, F. (eds.) *Amazonia: landscape and species evolution. A look into the past.* Wiley-Blackwell, Chichester, UK, p. 421- 431.

Winkworth, R.C., Wagstaff, S.J., Glenny, D. & Lockhart, P.J. 2005. Evolution of the New Zealand mountain flora: origins, diversification and dispersal. *Organisms, Diversity and Evolution* 5: 237-247.

Vasconcelos, M.F. 2008. Mountaintop endemism in eastern Brazil: why some bird species from campos rupestres of the Espinhaço Range are not endemic to the Cerrado region? *Revista Brasileira de Ornitologia*, 16(4):348-362.

Vasconcelos, M.F. 2011. O que são campos rupestres e campos de altitude nos topos de montanha do Leste do Brasil? *Revista Brasileira de Botânica* 34: 241-246.

Veloso, H.P. 1992. *Manual técnico da vegetação brasileira* (Manuais técnicos de geociências no1). Rio de Janeiro, IBGE. 92p.

Vilhena, D., Antonelli, A. (2015). A network approach for identifying and delimiting biogeographical regions. Nature Communications 6, 6848.

Vivo, M. 1997. Mammalian evidence of historical ecological change in the Caatinga semiarid vegetation of northeastern Brazil. *Journal of Comparative Biology* 2:65-73.

Yu, Y., Harris, A.J., He, X.J. 2014 . *RASP (Reconstruct Ancestral State in Phylogenies) 3.0.* *beta*. Avaliable at http://mnh.scu.edu.cn/soft/blog/RASP/

Zizka, A. et al. (2019). CoordinateCleaner: Standardized cleaning of occurrence records from biological collection databases. Methods in Ecology and Evolution 10: 744–751.